



# > Manual del usuario de SPSS Base 16.0



Si desea obtener más información sobre los productos de software de SPSS®, visite nuestra página Web en <http://www.spss.com> o póngase en contacto con

SPSS Inc.  
233 South Wacker Drive, 11th Floor  
Chicago, IL 60606-6412 EE.UU.  
Tel: (312) 651-3000  
Fax: (312) 651-3668

SPSS es una marca registrada; los demás nombres de productos son marcas comerciales de SPSS Inc. para los programas de software de su propiedad. El material descrito en este software no puede ser reproducido ni distribuido sin la autorización expresa por escrito por parte de los propietarios de la marca registrada y de los derechos de la licencia en el software y en los copyrights de los materiales publicados.

El SOFTWARE y la documentación se proporcionan con DERECHOS LIMITADOS. Su uso, duplicación o revelación por parte del Gobierno están sujetos a las restricciones establecidas en la subdivisión (c)(1)(ii) de la cláusula Rights in Technical Data and Computer Software en 52.227-7013. El fabricante es SPSS Inc., 233 South Wacker Drive, 11th Floor, Chicago, IL 60606-6412, EE.UU.  
Nº de patente 7,023,453

Aviso general: El resto de los nombres de productos mencionados en este documento se utilizan sólo con fines identificativos y pueden ser marcas comerciales de sus respectivas empresas.

Windows es una marca comercial registrada de Microsoft Corporation.

Apple, Mac y el logotipo de Mac son marcas comerciales de Apple Computer, Inc., registradas en Estados Unidos y en otros países.

Este producto utiliza WinWrap Basic, Copyright 1993-2007, Polar Engineering and Consulting, <http://www.winwrap.com>.

Manual del usuario de SPSS Base 16.0  
Copyright © 2007 de SPSS Inc.  
Reservados todos los derechos.

Queda prohibida la reproducción, el almacenamiento en sistemas de recuperación o la transmisión de cualquier parte de esta publicación en cualquier forma y por cualquier medio (electrónico o mecánico, fotocopia, grabación o cualquier otro) sin previa autorización expresa y por escrito de parte del editor.

ISBN-13: 978-0-13-603600-5  
ISBN-10: 0-13-603600-7

1 2 3 4 5 6 7 8 9 0 10 09 08 07

---

# Prefacio

## SPSS 16.0

SPSS 16.0 es un sistema global para el análisis de datos. SPSS puede adquirir datos de casi cualquier tipo de archivo y utilizarlos para generar informes tabulares, gráficos y diagramas de distribuciones y tendencias, estadísticos descriptivos y análisis estadísticos complejos.

Este manual, *Manual del usuario de SPSS Base 16.0*, proporciona la documentación sobre la interfaz gráfica de usuario de SPSS. Los ejemplos que utilizan procedimientos estadísticos que se encuentran en SPSS Base 16.0 se suministran en el sistema de ayuda que se instala con el software. Los algoritmos utilizados en los procedimientos estadísticos se proporcionan en formato PDF y están disponibles en el menú Ayuda.

Asimismo, además de los menús y los cuadros de diálogos, SPSS utiliza un lenguaje de comandos. Algunas de las funciones avanzadas del sistema sólo son accesibles a través de la sintaxis de comandos. (Dichas funciones no están disponibles en la versión para estudiantes.) La información de referencia detallada sobre la sintaxis de comandos está disponible en dos formatos: integrada en el sistema de ayuda global y como un documento independiente en formato PDF en la referencia de sintaxis de comandos (*SPSS 16.0 Command Syntax Reference*), también disponible en el menú Ayuda.

## SPSS Opciones

Las siguientes opciones están disponibles como mejoras adicionales de la versión completa (no la versión para estudiantes) del sistema SPSS Base:

**SPSS Regression Models™** proporciona técnicas para analizar los datos que no se ajusten a los tradicionales modelos estadísticos. Incluye procedimientos para el análisis probit, la regresión logística, la estimación ponderada, la regresión de mínimos cuadrados en dos fases y la regresión no lineal general.

**SPSS Advanced Models™** se centra en las técnicas utilizadas con más frecuencia en la investigación experimental y biomédica sofisticada. Incluye procedimientos para los modelos lineales generales (MLG), los modelos lineales mixtos, los modelos lineales generalizados (GZLM), las ecuaciones de estimación generalizadas (GEE), el análisis de componentes de la varianza, el análisis loglineal, las tablas de mortalidad actuariales, el análisis de supervivencia de Kaplan-Meier y las regresiones de Cox básica y extendida.

**SPSS Tables™** crea distintos informes tabulares de gran calidad, como por ejemplo tablas sofisticadas y presenta datos de respuestas múltiples.

**SPSS Trends™** realiza análisis de predicción y de series temporales muy completos, con diversos modelos de ajuste de curvas, modelos de suavizado y métodos para la estimación de funciones autorregresivas.

**SPSS Categories®** ejecuta procedimientos de escalamiento óptimo, incluidos los análisis de correspondencias.

**SPSS Conjoint™** ofrece una manera realista de medir el modo en que los diferentes atributos del producto afectan a las preferencias de los consumidores y los ciudadanos. Con Análisis conjunto de SPSS se puede medir con facilidad el efecto sobre la compensación de cada uno de los atributos de un producto dentro del contexto del conjunto de atributos del producto, tal como hacen los consumidores cuando deciden lo que van a comprar.

**SPSS Exact Tests™** calcula los valores  $p$  exactos (valores de significación) para las pruebas estadísticas en aquellos casos en los que las muestras son pequeñas o se distribuyen de forma poco uniforme y puedan hacer que las pruebas habituales resulten poco precisas.

**SPSS Missing Value Analysis™** describe los patrones de los datos perdidos, realiza una estimación de las medias y otros estadísticos y permite imputar los valores a las observaciones perdidas.

**SPSS Maps™** transforma los datos distribuidos geográficamente en mapas de alta calidad, con símbolos, colores, gráficos de barras, gráficos de sectores y combinaciones de temas, para representar no sólo lo que está ocurriendo sino también dónde tiene lugar.

**SPSS Complex Samples™** permite a los analistas de encuestas, mercado, salud y opinión pública, así como a los sociólogos que utilizan una metodología de encuesta de ejemplo, incorporar los diseños de muestras complejas al análisis de datos.

**SPSS Classification Trees™** crea un modelo de clasificación basado en árboles. y clasifica casos en grupos o pronostica valores de una variable (criterio) dependiente basada en valores de variables independientes (predictores). El procedimiento proporciona herramientas de validación para análisis de clasificación exploratorios y confirmatorios.

**SPSS Data Preparation™** ofrece una rápida instantánea visual de los datos. Ofrece la posibilidad de aplicar reglas de validación que identifiquen valores de los datos no válidos. Puede crear reglas que marquen los valores fuera de rango, valores perdidos o valores en blanco. También puede guardar variables que registren cada una de las violaciones de las reglas y el número total de violaciones de reglas por cada caso. También se incluye un conjunto limitado de reglas predefinidas que puede copiar o modificar.

**Amos™** (del inglés **analysis of moment structures**, análisis de estructuras de momento) utiliza el modelado de ecuaciones estructurales para confirmar y explicar los modelos conceptuales que tratan las actitudes, percepciones y otros factores que determinan el comportamiento.

La familia de productos de SPSS también incluye aplicaciones para la introducción de datos, análisis de textos, clasificación, redes neuronales y servicios predictivos para empresas.

### ***Instalación***

Para instalar SPSS Basesistema, ejecute el Asistente para autorización de licencia utilizando el código de autorización que le envió SPSS Inc. Si desea obtener más información, consulte las instrucciones de instalación proporcionadas con SPSS Basesistema.

### ***Compatibilidad***

SPSS está diseñado para ejecutarse en gran cantidad de sistemas de ordenadores. Consulte las instrucciones de instalación entregadas con su sistema para obtener información específica acerca de los requisitos mínimos y los recomendados.

### ***Números de serie***

El número de serie es su número de identificación con SPSS Inc. Necesitará este número cuando se ponga en contacto con SPSS Inc. para recibir información sobre asistencia, formas de pago o actualización del sistema. El número de serie se incluye en el sistema Base de SPSS.

### ***Servicio al cliente***

Si tiene cualquier duda referente a la forma de envío o pago, póngase en contacto con su oficina local, que encontrará en la página Web de SPSS en <http://www.spss.com/worldwide>. Tenga preparado su número de serie para identificarse.

### ***Cursos de preparación***

SPSS Inc. ofrece cursos de preparación, tanto públicos como in situ. En todos los cursos habrá talleres prácticos. Estos cursos tendrán lugar periódicamente en las principales capitales. Si desea obtener más información sobre estos cursos, póngase en contacto con su oficina local que encontrará en la página Web de SPSS en <http://www.spss.com/worldwide>.

### ***Asistencia técnica***

El servicio de asistencia técnica de SPSS está a disposición de todos los clientes de mantenimiento. Los clientes podrán ponerse en contacto con este servicio de asistencia técnica si desean recibir ayuda sobre el uso de SPSS o sobre la instalación en alguno de los entornos de hardware admitidos. Para ponerse en contacto con el servicio de asistencia técnica, consulte la página Web de SPSS en <http://www.spss.com>, o póngase en contacto con la oficina más cercana, que encontrará en la página Web de SPSS en <http://www.spss.com/worldwide>. Tenga preparada la información necesaria para identificarse personalmente, a su organización y el número de serie de su sistema.

### ***Publicaciones adicionales***

Puede adquirir copias adicionales de los manuales de los productos directamente de SPSS Inc. Visite la sección Store de la página Web de SPSS en <http://www.spss.com/estore> o póngase en contacto con su oficina de SPSS local que encontrará en la página Web de SPSS en <http://www.spss.com/worldwide>. Para pedidos telefónicos en Estados Unidos y Canadá, llame a SPSS Inc. al 800-543-2185. Para pedidos telefónicos desde otros países, póngase en contacto con la oficina más cercana que encontrará en la página Web de SPSS.

El libro *SPSS Statistical Procedures Companion*, de Marija Noruis, ha sido publicado por Prentice Hall. Está prevista una nueva versión de este libro, actualizado para SPSS 16.0. El libro *SPSS Advanced Statistical Procedures Companion*, que también se basa en SPSS 16.0, se publicará muy pronto. El libro *SPSS Guide to Data Analysis* para SPSS 16.0 también está en proceso de desarrollo. Las publicaciones anunciadas de forma exclusiva por Prentice Hall estarán

disponibles en la página Web de SPSS en <http://www.spss.com/estore> (seleccione su país de origen y pulse en Books).

### ***Díganos su opinión***

Sus comentarios son importantes. Háganos saber su experiencia con los productos SPSS. Nos interesa especialmente recibir noticias sobre aplicaciones nuevas e interesantes para el sistema SPSS Basesistema. Envíenos un correo electrónico a [suggest@spss.com](mailto:suggest@spss.com) o escriba a SPSS Inc., Attn.: Director of Product Planning, 233 South Wacker Drive, 11th Floor, Chicago, IL 60606-6412, EE.UU.

### ***Acerca de este manual***

Este manual es la documentación de la interfaz gráfica del usuario para los procedimientos incluidos en el módulo SPSS Basesistema. Las ilustraciones de los cuadros de diálogo están tomadas de SPSS . La información detallada sobre la sintaxis de comandos para las características de SPSS Basesistema está disponible en dos formatos: integrada en el sistema de ayuda global y como un documento independiente en formato PDF en *SPSS 16.0 Command Syntax Reference*, disponible en el menú Ayuda.

### ***Cómo ponerse en contacto con SPSS***

Si desea que le incluyamos en nuestra lista de correo, póngase en contacto con nuestras oficinas que encontrará en la página Web en <http://www.spss.com/worldwide>.

---

# Contenido

<b>1</b>	<b>Resumen</b>	<b>1</b>
	Novedades de la versión 16.0 . . . . .	2
	Windows . . . . .	4
	Ventana designada frente a ventana activa . . . . .	6
	Barra de estado . . . . .	6
	Cuadros de diálogo . . . . .	7
	Nombres y etiquetas de variable en las listas de los cuadros de diálogo . . . . .	7
	Ajuste del tamaño de los cuadros de diálogo . . . . .	8
	Controles del cuadro de diálogo . . . . .	8
	Selección de variables . . . . .	8
	Iconos de tipo de datos, nivel de medida y lista de variables . . . . .	9
	Información sobre las variables en un cuadro de diálogo . . . . .	9
	Pasos básicos en el análisis de datos . . . . .	10
	Asesor estadístico . . . . .	10
	Información adicional . . . . .	11
<b>2</b>	<b>Obtención de ayuda</b>	<b>12</b>
	Obtención de ayuda sobre los términos de los resultados . . . . .	13
<b>3</b>	<b>Archivos de datos</b>	<b>15</b>
	Apertura de archivos de datos . . . . .	15
	Para abrir archivos de datos . . . . .	15
	Tipos de archivos de datos . . . . .	16
	Opciones de apertura de archivos . . . . .	17
	Lectura de archivos de Excel 95 o versiones posteriores . . . . .	17
	Lectura de archivos de Excel y otras hojas de cálculo de versiones anteriores . . . . .	17
	Lectura de archivos de dBASE . . . . .	18
	Lectura de archivos de Stata . . . . .	18
	Lectura de archivos de bases de datos . . . . .	18
	Asistente para la importación de texto . . . . .	32
	Lectura de datos de Dimensions . . . . .	41
	Información sobre el archivo . . . . .	44

Almacenamiento de archivos de datos . . . . .	45
Almacenamiento de archivos de datos modificados. . . . .	45
Almacenamiento de archivos de datos en formatos externos. . . . .	45
Almacenamiento de archivos de datos en formato de Excel. . . . .	48
Almacenamiento de archivos de datos en formato de SAS. . . . .	49
Almacenamiento de archivos de datos en formato de Stata . . . . .	50
Almacenamiento de subconjuntos de variables . . . . .	52
Exportación a base de datos . . . . .	52
Exportación a Dimensions . . . . .	65
Protección de datos originales . . . . .	66
Archivo activo virtual . . . . .	67
Creación de una caché de datos . . . . .	68

## **4 *Análisis en modo distribuido* 71**

Acceso al servidor . . . . .	71
Adición y edición de la configuración de acceso al servidor . . . . .	72
Para seleccionar, cambiar o añadir servidores. . . . .	73
Búsqueda de servidores disponibles . . . . .	74
Apertura de archivos de datos desde un servidor remoto . . . . .	75
Acceso a archivo en análisis en modo local y distribuido . . . . .	75
Disponibilidad de procedimientos en análisis en modo distribuido . . . . .	77
Especificaciones de rutas absolutas frente a rutas relativas . . . . .	77

## **5 *Editor de datos* 79**

Vista de datos. . . . .	79
Vista de variables. . . . .	80
Para visualizar o definir los atributos de las variables . . . . .	81
Nombres de variable . . . . .	82
Nivel de medida de variable. . . . .	83
Tipo de variable . . . . .	83
Etiquetas de variable . . . . .	85
Etiquetas de valor . . . . .	86
Inserción de saltos de línea en etiquetas . . . . .	86
Valores perdidos . . . . .	87
Ancho de columna. . . . .	88
Alineación de la variable . . . . .	88
Aplicación de atributos de definición de variables a varias variables . . . . .	88
Atributos de variable personalizados . . . . .	89



Personalización de la Vista de variables . . . . .	93
Revisión ortográfica de las etiquetas de los valores y las variables . . . . .	94
Introducción de datos. . . . .	94
Para introducir datos numéricos . . . . .	94
Para introducir datos no numéricos . . . . .	95
Para utilizar etiquetas de valor en la introducción de datos . . . . .	95
Restricciones de los valores de datos en el Editor de datos . . . . .	95
Edición de datos. . . . .	95
Para reemplazar o modificar un valor de datos . . . . .	96
Cortar, copiar y pegar valores de datos . . . . .	96
Inserción de nuevos casos . . . . .	97
Inserción de nuevas variables . . . . .	97
Para cambiar el tipo de datos . . . . .	98
Búsqueda de casos o variables . . . . .	98
Búsqueda y sustitución de datos y valores de atributo . . . . .	99
Estado de selección de casos en el Editor de datos . . . . .	100
Editor de datos: Opciones de presentación. . . . .	100
Impresión en el Editor de datos. . . . .	101
Para imprimir el contenido del Editor de datos . . . . .	101

## **6 Trabajo con varios orígenes de datos 102**

Tratamiento básico de varios orígenes de datos . . . . .	103
Trabajo con varios conjuntos de datos en la sintaxis de comandos . . . . .	104
Copia y pegado de información entre conjuntos de datos . . . . .	105
Cambio del nombre de los conjuntos de datos . . . . .	105
Supresión de varios conjuntos de datos . . . . .	106

## **7 Preparación de datos 107**

Propiedades de variables . . . . .	107
Definición de propiedades de variables . . . . .	108
Para definir propiedades de variables . . . . .	108
Definición de etiquetas de valor y otras propiedades de las variables . . . . .	109
Asignación del nivel de medida . . . . .	111
Atributos de variable personalizados . . . . .	112
Copia de propiedades de variables . . . . .	113
Conjuntos de respuestas múltiples . . . . .	114
Para definir conjuntos de respuestas múltiples . . . . .	114

Copia de propiedades de datos . . . . .	117
Para copiar propiedades de datos . . . . .	118
Selección de las variables de origen y de destino . . . . .	118
Selección de propiedades de variable para copiar. . . . .	120
Copia de propiedades (de archivo) de conjunto de datos . . . . .	121
Resultados . . . . .	124
Identificación de casos duplicados. . . . .	124
Agrupación visual. . . . .	127
Para agrupar variables . . . . .	128
Agrupación de variables . . . . .	128
Generación automática de categorías agrupadas . . . . .	131
Copia de categorías agrupadas . . . . .	133
Valores perdidos definidos por el usuario en la agrupación visual . . . . .	135

## **8 Transformaciones de los datos 136**

Cálculo de variables . . . . .	136
Calcular variable: Si los casos. . . . .	138
Calcular variable: Tipo y etiqueta. . . . .	138
Funciones. . . . .	139
Valores perdidos en funciones . . . . .	139
Generadores de números aleatorios. . . . .	140
Contar apariciones de valores dentro de los casos . . . . .	141
Contar valores dentro de los casos: Contar los valores . . . . .	142
Contar apariciones: Si los casos . . . . .	142
Recodificación de valores . . . . .	143
Recodificar en las mismas variables. . . . .	143
Recodificar en las mismas variables: Valores antiguos y nuevos . . . . .	144
Recodificar en distintas variables. . . . .	146
Recodificar en distintas variables: Valores antiguos y nuevos . . . . .	147
Asignar rangos a los casos. . . . .	148
Asignar rangos a los casos: Tipos . . . . .	149
Asignar rangos a los casos: Empates. . . . .	150
Recodificación automática . . . . .	151
Asistente para fecha y hora . . . . .	154
Fechas y horas en SPSS . . . . .	156
Creación de una variable de fecha/hora a partir de una cadena. . . . .	157
Creación de una variable de fecha/hora a partir de un conjunto de variables. . . . .	158
Adición o sustracción de valores a partir de variables de fecha/hora. . . . .	160
Extracción de parte de una variable de fecha/hora . . . . .	167

Transformaciones de los datos de serie temporal . . . . .	169
Definir fechas . . . . .	170
Crear serie temporal . . . . .	171
Reemplazar los valores perdidos . . . . .	174
Puntuación de datos con modelos predictivos . . . . .	175
Carga de un modelo guardado . . . . .	176
Visualización de una lista de modelos cargados . . . . .	179
Funciones adicionales con sintaxis de comandos . . . . .	179

## **9   *Gestión y transformación de los archivos* 180**

Ordenar casos . . . . .	180
Ordenación de variables . . . . .	181
Transponer . . . . .	182
Fusión de archivos de datos . . . . .	183
Añadir casos . . . . .	184
Añadir casos: Cambiar nombre . . . . .	186
Añadir casos: Información del diccionario . . . . .	186
Fusión de más de dos orígenes de datos . . . . .	187
Añadir variables . . . . .	187
Añadir variables: Cambiar nombre . . . . .	189
Fusión de más de dos orígenes de datos . . . . .	189
Agregar datos . . . . .	189
Agregar datos: Función de agregación . . . . .	192
Agregar datos: Nombre y etiqueta de variable . . . . .	192
Segmentar archivo . . . . .	193
Seleccionar casos . . . . .	194
Seleccionar casos: Si la opción . . . . .	196
Seleccionar casos: Muestra aleatoria . . . . .	196
Seleccionar casos: Amplitud . . . . .	197
Ponderar casos . . . . .	198
Reestructuración de los datos . . . . .	199
Para reestructurar datos . . . . .	199
Asistente de reestructuración de datos: Seleccionar tipo . . . . .	199
Asistente de reestructuración de datos (variables a casos): Número de grupos de variables . . . . .	203
Asistente de reestructuración de datos (variables a casos): Seleccionar Variables . . . . .	204
Asistente de reestructuración de datos (variables a casos): Crear variables de Variables . . . . .	206
Asistente de reestructuración de datos (variables a casos): Crear una variable de índices . . . . .	208
Asistente de reestructuración de datos (variables a casos): Crear varias variables de índices . . . . .	209
Asistente de reestructuración de datos (variables a casos): Opciones . . . . .	210

Asistente de reestructuración de datos (casos a variables): Seleccionar Variables . . . . .	211
Asistente de reestructuración de datos (casos a variables): Ordenar Datos . . . . .	213
Asistente de reestructuración de datos (casos a variables): Opciones. . . . .	214
Asistente de reestructuración de datos: Finalizar. . . . .	215

## **10 Trabajo con resultados 217**

Visor. . . . .	217
Mostrar y ocultar resultados . . . . .	218
Desplazamiento, eliminación y copia de resultados . . . . .	218
Cambio de la alineación inicial. . . . .	219
Cambio de la alineación de los elementos de resultados . . . . .	219
Resultados del Visor . . . . .	219
Adición de elementos al Visor . . . . .	221
Búsqueda y sustitución de información en el Visor. . . . .	222
Copia de resultados a otras aplicaciones . . . . .	223
Para copiar y pegar elementos de resultados en otra aplicación . . . . .	224
Exportar resultados . . . . .	224
Opciones de HTML, Word/RTF y Excel . . . . .	226
Opciones de PowerPoint . . . . .	227
Opciones de PDF. . . . .	227
Opciones del texto. . . . .	229
Opciones para la exportación de gráficos . . . . .	230
Impresión de documentos del Visor . . . . .	232
Para imprimir resultados y gráficos . . . . .	232
Presentación preliminar. . . . .	232
Atributos de página: Encabezados y pies. . . . .	233
Atributos de página: Opciones. . . . .	235
Almacenamiento de resultados. . . . .	236
Para guardar un documento del Visor . . . . .	236

## **11 Tablas pivote 237**

Gestión de una tabla pivote. . . . .	237
Activación de una tabla pivote. . . . .	237
Pivotado de una tabla . . . . .	237
Cambio del orden de presentación de los elementos dentro de una dimensión. . . . .	238
Movimiento de filas y columnas dentro de un elemento de una dimensión . . . . .	238
Para transponer filas y columnas. . . . .	239
Agrupación de filas o columnas . . . . .	239

Desagrupación de filas o columnas . . . . .	239
Rotación de etiquetas de filas o columnas . . . . .	239
Trabajo con capas . . . . .	240
Creación y presentación de capas . . . . .	240
Ir a la categoría de capa . . . . .	242
Visualización y ocultación de elementos . . . . .	243
Ocultación de filas y columnas de una tabla . . . . .	243
Visualización de las filas y columnas ocultas de una tabla . . . . .	243
Ocultación y visualización de las etiquetas de dimensión . . . . .	244
Ocultación y presentación de los títulos de tabla . . . . .	244
Aspectos de tabla . . . . .	244
Para aplicar o guardar un aspecto de tabla . . . . .	244
Para editar o crear un aspecto de tabla . . . . .	245
Propiedades de tabla . . . . .	246
Para modificar las propiedades de las tablas pivote . . . . .	246
Propiedades de tabla: General . . . . .	246
Propiedades de tabla: Notas al pie . . . . .	247
Propiedades de tabla: Formatos de casilla . . . . .	248
Propiedades de tabla: Bordes . . . . .	250
Propiedades de tabla: Imprimiendo . . . . .	251
Propiedades de casilla . . . . .	252
Fuente y fondo . . . . .	253
Formato de valor . . . . .	253
Alineación y márgenes . . . . .	254
Texto y notas al pie . . . . .	255
Adición de texto y notas al pie . . . . .	255
Para mostrar u ocultar el texto al pie . . . . .	256
Para ocultar o mostrar una nota al pie de una tabla . . . . .	256
Marcador de nota al pie . . . . .	256
Para volver a numerar las notas al pie . . . . .	257
Anchos de casillas de datos . . . . .	257
Cambio del ancho de columna . . . . .	257
Presentación de los bordes ocultos de una tabla pivote . . . . .	257
Selección de filas y columnas de una tabla pivote . . . . .	258
Impresión de tablas pivote . . . . .	259
Control de las rupturas de tabla para tablas anchas y largas . . . . .	259
Creación de un gráfico a partir de una tabla pivote . . . . .	260

## **12 Trabajar con sintaxis de comandos**

**261**

Reglas de la sintaxis . . . . .	261
---------------------------------	-----

Pegar sintaxis desde cuadros de diálogo . . . . .	263
Para pegar sintaxis desde cuadros de diálogo . . . . .	263
Copia de la sintaxis desde las anotaciones de los resultados . . . . .	264
Para copiar la sintaxis desde las anotaciones de los resultados. . . . .	264
Para ejecutar la sintaxis de comandos . . . . .	265
Archivos de sintaxis Unicode . . . . .	265
Varios comandos Ejecutar . . . . .	266
<b>13 Frecuencias</b>	<b>267</b>
Frecuencias: Estadísticos. . . . .	268
Frecuencias: Gráficos . . . . .	270
Frecuencias: Formato. . . . .	271
<b>14 Descriptivos</b>	<b>272</b>
Descriptivos: Opciones. . . . .	273
Funciones adicionales del comando DESCRIPTIVES . . . . .	274
<b>15 Explorar</b>	<b>276</b>
Explorar: Estadísticos . . . . .	277
Explorar: Gráficos. . . . .	278
Explorar: Transformaciones de potencia . . . . .	279
Explorar: Opciones . . . . .	279
Funciones adicionales del comando EXAMINE. . . . .	280
<b>16 Tablas de contingencia</b>	<b>281</b>
Capas de las tablas de contingencia. . . . .	283
Gráficos de barras agrupadas de las tablas de contingencia . . . . .	283
Tablas de contingencia: Estadísticos . . . . .	283
Tablas de contingencia: Mostrar en las casillas . . . . .	286
Tablas de contingencia: Formato de tabla . . . . .	287

<b>17</b>	<b><i>Resumir</i></b>	<b>288</b>
	Resumir: Opciones . . . . .	290
	Resumir: Estadísticos . . . . .	290
<b>18</b>	<b><i>Medias</i></b>	<b>293</b>
	Medias: Opciones . . . . .	295
<b>19</b>	<b><i>Cubos OLAP</i></b>	<b>298</b>
	Cubos OLAP: Estadísticos . . . . .	299
	Cubos OLAP: Diferencias . . . . .	302
	Cubos OLAP: Título . . . . .	303
<b>20</b>	<b><i>Pruebas T</i></b>	<b>304</b>
	Prueba T para muestras independientes. . . . .	304
	Definición de grupos en la prueba T para muestras independientes. . . . .	306
	Prueba T para muestras independientes: Opciones . . . . .	306
	Prueba T para muestras relacionadas . . . . .	307
	Prueba T para muestras relacionadas: Opciones . . . . .	308
	Prueba T para una muestra. . . . .	309
	Prueba T para una muestra: Opciones . . . . .	310
	Funciones adicionales del comando T-TEST . . . . .	310
<b>21</b>	<b><i>ANOVA de un factor</i></b>	<b>311</b>
	ANOVA de un factor: Contrastes. . . . .	312
	ANOVA de un factor: Contrastes post hoc. . . . .	313
	ANOVA de un factor: Opciones. . . . .	316
	Funciones adicionales del comando ONEWAY . . . . .	317

## **22 MLG Análisis univariante** **318**

MLG: Modelo . . . . .	320
Construir términos . . . . .	320
Suma de cuadrados . . . . .	321
MLG: Contrastes . . . . .	322
Tipos de contrastes . . . . .	322
MLG: Gráficos de perfil . . . . .	323
MLG: Comparaciones post hoc . . . . .	324
MLG: Guardar . . . . .	326
Opciones MLG . . . . .	328
Funciones adicionales de los comandos UNIANOVA . . . . .	329

## **23 Correlaciones bivariadas** **331**

Correlaciones bivariadas: Opciones . . . . .	333
Funciones adicionales de los comandos CORRELATIONS y NONPAR CORR. . . . .	333

## **24 Correlaciones parciales** **335**

Correlaciones parciales: Opciones . . . . .	336
Funciones adicionales del comando PARTIAL CORR. . . . .	337

## **25 Distancias** **338**

Distancias: Medidas de disimilaridad . . . . .	340
Distancias: Medidas de similaridad . . . . .	341
Funciones adicionales del comando PROXIMITIES . . . . .	342

## **26 Regresión lineal** **343**

Métodos de selección de variables en el análisis de regresión lineal . . . . .	344
Regresión lineal: Establecer regla . . . . .	346
Regresión lineal: Gráficos . . . . .	346
Regresión lineal: Para guardar variables nuevas . . . . .	347



Regresión lineal: Estadísticos . . . . .	350
Regresión lineal: Opciones . . . . .	351
Funciones adicionales del comando REGRESSION . . . . .	352
<b>27 Regresión ordinal</b>	<b>353</b>
Regresión ordinal: Opciones . . . . .	354
Resultados de la regresión ordinal . . . . .	356
Modelo de ubicación de la regresión ordinal . . . . .	357
Construir términos . . . . .	359
Modelo de escala de la regresión ordinal . . . . .	358
Construir términos . . . . .	359
Funciones adicionales del comando PLUM . . . . .	359
<b>28 Estimación curvilínea</b>	<b>360</b>
Modelos del procedimiento Estimación curvilínea . . . . .	362
Estimación curvilínea: Guardar . . . . .	362
<b>29 Regresión de mínimos cuadrados parciales</b>	<b>364</b>
Modelo . . . . .	366
Opciones . . . . .	367
<b>30 Análisis discriminante</b>	<b>369</b>
Análisis discriminante: Definir rango . . . . .	371
Análisis discriminante: Seleccionar casos . . . . .	371
Análisis discriminante: Estadísticos . . . . .	372
Análisis discriminante: Método de inclusión por pasos . . . . .	373
Análisis discriminante: Clasificar . . . . .	374
Análisis discriminante: Guardar . . . . .	376
Funciones adicionales del comando DISCRIMINANT . . . . .	376

<b>31</b>	<b><i>Análisis factorial</i></b>	<b>377</b>
	Selección de casos en el análisis factorial . . . . .	378
	Análisis factorial: Descriptivos . . . . .	379
	Análisis factorial: Extracción . . . . .	380
	Análisis factorial: Rotación . . . . .	381
	Análisis factorial: Puntuaciones factoriales . . . . .	382
	Análisis factorial: Opciones . . . . .	383
	Funciones adicionales del comando FACTOR . . . . .	383
<b>32</b>	<b><i>Selección de procedimientos para la conglomeración</i></b>	<b>385</b>
<b>33</b>	<b><i>Análisis de conglomerados en dos fases</i></b>	<b>387</b>
	Opciones del análisis de conglomerados en dos fases . . . . .	390
	Gráficos de análisis de conglomerados en dos fases . . . . .	392
	Resultados de análisis de conglomerados en dos fases . . . . .	393
<b>34</b>	<b><i>Análisis de conglomerados jerárquico</i></b>	<b>395</b>
	Análisis de conglomerados jerárquico: Método . . . . .	397
	Análisis de conglomerados jerárquico: Estadísticos . . . . .	398
	Análisis de conglomerados jerárquico: Gráficos . . . . .	399
	Análisis de conglomerados jerárquico: Guardar variables nuevas . . . . .	399
	Funciones adicionales de la sintaxis de comandos CLUSTER . . . . .	400
<b>35</b>	<b><i>Análisis de conglomerados de K-medias</i></b>	<b>401</b>
	Eficacia del análisis de conglomerados de K-medias . . . . .	403
	Análisis de conglomerados de K-medias: Iterar . . . . .	403
	Análisis de conglomerados de K-medias: Guardar . . . . .	404
	Análisis de conglomerados de K-medias: Opciones . . . . .	404
	Funciones adicionales del comando QUICK CLUSTER . . . . .	405

## **36 Pruebas no paramétricas**

**406**

Prueba de chi-cuadrado . . . . .	406
Prueba de chi-cuadrado: Rango y valores esperados . . . . .	408
Prueba de chi-cuadrado: Opciones . . . . .	408
Funciones adicionales del comando NPAR TESTS (prueba de chi-cuadrado) . . . . .	409
Prueba binomial . . . . .	409
Prueba binomial: Opciones . . . . .	410
Funciones adicionales del comando NPAR TESTS (Prueba binomial) . . . . .	411
Prueba de rachas . . . . .	411
Prueba de rachas: punto de corte . . . . .	412
Prueba de rachas: Opciones . . . . .	412
Funciones adicionales del comando NPAR TESTS (Prueba de rachas) . . . . .	413
Prueba de Kolmogorov-Smirnov para una muestra . . . . .	413
Prueba de Kolmogorov-Smirnov para una muestra: Opciones . . . . .	414
Funciones adicionales del comando NPAR TESTS (Prueba de Kolmogorov-Smirnov para una muestra) . . . . .	415
Pruebas para dos muestras independientes . . . . .	415
Dos muestras independientes: Tipos de pruebas . . . . .	416
Pruebas para dos muestras independientes: Definir grupos . . . . .	417
Pruebas para dos muestras independientes: Opciones . . . . .	417
Funciones adicionales del comando NPAR TESTS (Dos muestras independientes) . . . . .	418
Pruebas para dos muestras relacionadas . . . . .	418
Dos muestras relacionadas: Tipos de pruebas . . . . .	419
Pruebas para dos muestras relacionadas: Opciones . . . . .	420
Funciones adicionales del comando NPAR TESTS (Dos muestras relacionadas) . . . . .	420
Pruebas para varias muestras independientes . . . . .	420
Tipos de pruebas para varias muestras independientes . . . . .	421
Varias muestras independientes: Definir rango . . . . .	422
Varias muestras independientes: Opciones . . . . .	422
Funciones adicionales del comando NPAR TESTS (K muestras independientes) . . . . .	422
Pruebas para varias muestras relacionadas . . . . .	423
Tipos de prueba en el procedimiento Pruebas para varias muestras relacionadas . . . . .	423
Pruebas para varias muestras relacionadas: Estadísticos . . . . .	424
Funciones adicionales del comando NPAR TESTS (K pruebas relacionadas) . . . . .	424

## **37 Análisis de respuestas múltiples**

**425**

Definir conjuntos de respuestas múltiples . . . . .	426
Frecuencias de respuestas múltiples . . . . .	427

Tablas de contingencia de respuestas múltiples . . . . .	429
Tablas de respuestas múltiples: Definir rangos de las variables . . . . .	430
Tablas de contingencia de respuestas múltiples: Opciones. . . . .	431
Funciones adicionales del comando MULT RESPONSE. . . . .	432

## **38 Informes de los resultados 433**

Informe de estadísticos en filas . . . . .	433
Para obtener un informe de resumen: Estadísticos en filas . . . . .	434
Formato de las columnas de datos y de la ruptura de columnas del informe. . . . .	435
Líneas de resumen finales y Líneas de resumen del informe. . . . .	435
Opciones de ruptura del informe . . . . .	436
Opciones del informe. . . . .	436
Diseño del informe. . . . .	437
Títulos del informe . . . . .	438
Informe de estadísticos en columnas . . . . .	438
Para obtener un informe de resumen: Estadísticos en columnas . . . . .	439
Función Columna de resumen total . . . . .	440
Columna de resumen total . . . . .	441
Formato de columna del informe . . . . .	442
Opciones de la ruptura de columnas para los estadísticos en el informe . . . . .	442
Opciones de columnas para los estadísticos en el informe. . . . .	442
Diseño del informe para los estadísticos en columnas . . . . .	443
Funciones adicionales del comando REPORT . . . . .	443

## **39 Análisis de fiabilidad 444**

Análisis de fiabilidad: Estadísticos . . . . .	446
Funciones adicionales del comando RELIABILITY . . . . .	448

## **40 Escalamiento multidimensional 449**

Escalamiento multidimensional: Forma de los datos . . . . .	451
Escalamiento multidimensional: Crear la medida a partir de los datos. . . . .	451
Escalamiento multidimensional: Modelo . . . . .	452
Escalamiento multidimensional: Opciones . . . . .	453
Funciones adicionales del comando de ALSCAL . . . . .	454

<b>41 Estadísticos de la razón</b>	<b>455</b>
Estadísticos de la razón .....	457
<b>42 Curvas COR</b>	<b>459</b>
Curvas COR: Opciones .....	460
<b>43 Conceptos básicos de la utilidad de gráficos</b>	<b>462</b>
Creación y modificación de gráficos .....	462
Generación de gráficos .....	462
Edición de gráficos .....	466
Opciones de definición de gráfico .....	469
Adición y edición de títulos y notas al pie .....	469
Para establecer las opciones generales .....	470
<b>44 Utilidades</b>	<b>473</b>
Información sobre la variable .....	473
Comentarios del archivo de datos .....	474
Conjuntos de variables .....	474
Definir conjuntos de variables .....	474
Utilizar conjuntos de variables .....	475
Reordenación de listas de variables de destino .....	477
<b>45 Opciones</b>	<b>478</b>
General: Opciones .....	479
Opciones del Visor .....	481
Datos: Opciones .....	482
Cambio de la Vista de variables por defecto .....	484
Moneda: Opciones .....	484
Para crear formatos de moneda personalizados .....	485

Etiqueta de los resultados: Opciones . . . . .	485
Opciones de gráfico . . . . .	487
Colores de los elementos de datos . . . . .	488
Líneas de los elementos de datos . . . . .	488
Marcadores de los elementos de datos . . . . .	489
Rellenos de los elementos de datos . . . . .	489
Tabla pivote: Opciones . . . . .	490
Opciones de ubicaciones de archivos. . . . .	492
Procesos: Opciones . . . . .	493

## **46 Personalización de menús y barras de herramientas 496**

Editor de menús . . . . .	496
Personalización de las barras de herramientas. . . . .	497
Mostrar barras de herramientas. . . . .	497
Para personalizar las barras de herramientas. . . . .	498
Propiedades de la barra de herramientas . . . . .	499
Barra de herramientas de edición . . . . .	499
Crear nueva herramienta . . . . .	500

## **47 Trabajos de producción 502**

Opciones HTML . . . . .	505
Opciones de PowerPoint . . . . .	506
Opciones de PDF . . . . .	506
Opciones del texto . . . . .	506
Valores de tiempo de ejecución . . . . .	507
Entradas del usuario. . . . .	508
Ejecución de trabajos de producción desde una línea de comandos. . . . .	509
Conversión de los archivos de la unidad de producción . . . . .	510

## **48 Sistema de gestión de resultados 511**

Tipos de objetos de resultados . . . . .	514
Identificadores de comandos y subtipos de tabla . . . . .	515
Etiquetas . . . . .	516
Opciones de SGR . . . . .	517

Registro . . . . .	522
Exclusión de presentación de resultados del Visor . . . . .	522
Envío de resultados a archivos de datos de SPSS . . . . .	523
Ejemplo: Tabla sencilla de dos dimensiones . . . . .	523
Ejemplo: Tablas con capas . . . . .	524
Archivos de datos creados a partir de varias tablas . . . . .	525
Control de elementos de columna para las variables de control del archivo de datos . . . . .	528
Nombres de variable en los archivos de datos generados por SGR . . . . .	530
Estructura de tablas OXML . . . . .	531
Identificadores SGR . . . . .	535
Copia de identificadores SGR desde los titulares del Visor . . . . .	536

## ***Apéndice***

### ***Índice***

**537**





# Resumen

SPSS proporciona un poderoso sistema de análisis estadístico y gestión de datos en un entorno gráfico, utilizando menús descriptivos y cuadros de diálogo sencillos que realizan la mayor parte del trabajo. La mayoría de las tareas se pueden llevar a cabo simplemente situando el puntero del ratón en el lugar deseado y pulsando en el botón.

Además de la cómoda interfaz que permite realizar análisis estadísticos con sólo situar el puntero y pulsar, SPSS proporciona:

**Editor de datos.** El Editor de datos es un sistema versátil, similar a una hoja de cálculo, para definir, introducir, editar y presentar datos.

**Visor.** El Visor permite examinar los resultados, mostrarlos y ocultarlos de forma selectiva, modificar el orden de presentación en la pantalla y desplazar tablas y gráficos de gran calidad a otras aplicaciones y desde ellas.

**Tablas de pivote multidimensionales.** Sus resultados cobrarán vida en las tablas pivote multidimensionales. Explore las tablas reorganizando las filas, las columnas y las capas. Haga importantes descubrimientos que suelen quedar velados en los informes convencionales. Compare fácilmente los grupos dividiendo la tabla de manera que aparezca solamente un grupo cada vez.

**Gráficos de alta resolución.** Como funciones básicas se incluyen gráficos de sectores, gráficos de barras, histogramas, diagramas de dispersión y gráficos 3-D de alta resolución y a todo color, entre muchos otros.

**Acceso a bases de datos.** Obtenga información de bases de datos mediante el Asistente para bases de datos en lugar de utilizar consultas SQL de gran complejidad.

**Transformaciones de los datos.** Las funciones de transformación le ayudarán a preparar sus datos para el análisis. Puede crear fácilmente subconjuntos de datos, combinar categorías, añadir, agregar, fusionar, segmentar y transponer archivos, entre muchas otras posibilidades.

**Ayuda en pantalla.** Los tutoriales le ofrecerán una introducción global ampliamente detallada; los temas de la Ayuda sensible al contexto de los cuadros de diálogo le guiarán a través de la ejecución de tareas específicas; las definiciones de las ventanas emergentes para los resultados de las tablas pivote le explicarán los términos estadísticos y Estudios de casos le proporcionará ejemplos prácticos sobre cómo utilizar los procedimientos estadísticos y cómo interpretar los resultados.

**Lenguaje de comandos.** Aunque la mayoría de las tareas se pueden llevar a cabo simplemente situando el puntero del ratón en el lugar deseado y pulsando en el botón, SPSS proporciona además un potente lenguaje de comandos que permite guardar y automatizar muchas tareas habituales. El lenguaje de comandos también proporciona algunas funcionalidades no incluidas en los menús y cuadros de diálogo.

La documentación completa sobre la sintaxis de comandos está integrada en el sistema de ayuda global y está disponible como un documento en formato PDF independiente, la referencia de sintaxis de comandos de SPSS (*Command Syntax Reference*), que también está disponible en el menú Ayuda.

## ***Novedades de la versión 16.0***

**Mejoras de la interfaz de usuario.** La interfaz interactiva incluye las siguientes mejoras:

- Ahora se puede ajustar el tamaño de todos los cuadros de diálogo. La posibilidad de ensanchar los cuadros de diálogo permite ensanchar las listas de variables, lo que permite ver una mayor longitud de los nombres de variables y/o sus etiquetas descriptivas. La posibilidad de alargar los cuadros de diálogo permite alargar las listas de variables, lo que permite ver más variables sin necesidad de desplazarse.
- Ahora todos los cuadros de diálogo admiten la selección de variables mediante la función de arrastrar y colocar.
- El orden y las características de presentación de la lista de variables se puede cambiar directamente en todos los cuadros de diálogo. Cambie el orden de clasificación (alfabético, orden de archivos, nivel de medida) y/o cambie entre la presentación de nombres de variables y etiquetas de variable cuando lo desee. Si desea obtener más información, consulte [Nombres y etiquetas de variable en las listas de los cuadros de diálogo](#) en p. 7.

**Gestión de datos y resultados.** La gestión de datos y resultados incluye las siguientes mejoras:

- Lectura y escritura de archivos de Excel 2007.
- Elija entre trabajar con varios conjuntos de datos o con un único conjunto de datos a la vez. Si desea obtener más información, consulte [General: Opciones](#) en Capítulo 45 en p. 479.
- Busque y sustituya información en los documentos del visor, incluidos elementos ocultos y capas de tablas pivote multidimensionales. Si desea obtener más información, consulte [Búsqueda y sustitución de información en el Visor](#) en Capítulo 10 en p. 222.
- Asigne valores perdidos y etiquetas de valor a cualquier variable de cadena, independientemente del ancho de cadena definido (antes estaba limitado a cadenas con un ancho definido de 8 bytes como máximo).
- Nuevas funciones de cadena basadas en caracteres.
- Compatible con el Sistema de gestión de resultados (SGR) para los gráficos y mapas de imagen con formato de archivo del Visor (.spv) y formato VML con información emergente sobre el gráfico para documentos HTML. Si desea obtener más información, consulte [Sistema de gestión de resultados](#) en Capítulo 48 en p. 511.
- Personalice la Vista de variables en el Editor de datos. Cambie el orden de presentación de las columnas de atributos y controle qué columnas de atributos se muestran. Si desea obtener más información, consulte [Personalización de la Vista de variables](#) en Capítulo 5 en p. 93.
- Ordene las variables del conjunto de datos activo alfabéticamente o por los valores de atributo (diccionario). Si desea obtener más información, consulte [Ordenación de variables](#) en Capítulo 9 en p. 181.

- Revise la ortografía de las etiquetas de variable y las etiquetas de valor en la Vista de variables. Si desea obtener más información, consulte [Revisión ortográfica de las etiquetas de los valores y las variables](#) en Capítulo 5 en p. 94.
- Cambie el tipo de variable básico (numérica, de cadena), cambie el ancho definido de las variables de cadena y establezca de forma automática el ancho de las variables de cadena en el valor más largo observado para cada variable.
- Lea y escriba datos Unicode y archivos de sintaxis. Si desea obtener más información, consulte [General: Opciones](#) en Capítulo 45 en p. 479.
- Controle la ubicación del directorio por defecto para abrir y guardar archivos. Si desea obtener más información, consulte [Opciones de ubicaciones de archivos](#) en Capítulo 45 en p. 492.

**Rendimiento.** Para los ordenadores con varios procesadores o con procesadores con varios núcleos, ahora pueden ejecutarse algunos procedimientos mediante multiproceso, lo que ofrece un rendimiento más rápido.

**Mejoras en los estadísticos.** Los estadísticos incluyen las siguientes mejoras:

- Mínimos cuadrados parciales (PLS). Es una técnica de predicción alternativa a la regresión de mínimos cuadrados ordinarios (OLS), a la correlación canónica o al modelado de ecuaciones estructurales, y resulta particularmente útil cuando las variables predictoras están muy correlacionadas o cuando el número de predictores es superior al número de casos. Si desea obtener más información, consulte [Regresión de mínimos cuadrados parciales](#) en Capítulo 29 en p. 364.
- Perceptrón multicapa (MLP). El procedimiento MLP ajusta un tipo concreto de red neuronal denominado perceptrón multicapa. El perceptrón multicapa utiliza una arquitectura proactiva y puede tener varias capas ocultas. El perceptrón multicapa es muy flexible respecto a los tipos de modelos que puede ajustar. Es una de las arquitecturas de redes neuronales que se utilizan con mayor frecuencia. Este procedimiento está disponible en la nueva opción Redes neuronales.
- Función de base radial (RBF). Una red de función de base radial (RBF) es una red de aprendizaje proactiva y supervisada que sólo tiene una capa oculta, denominada capa de función de base radial. Al igual que la red de perceptrón multicapa (MLP), la red de función de base radial puede realizar predicciones y clasificaciones. Puede ser mucho más rápida que la MLP, sin embargo, no es tan flexible en el tipo de modelos a los que puede ajustarse. Este procedimiento está disponible en la nueva opción Redes neuronales.
- Los modelos lineales generalizados admiten numerosas características nuevas, incluidas las distribuciones multinomiales ordinales y las distribuciones Tweedie, la estimación de máxima verosimilitud del parámetro auxiliar binomial negativo y los estadísticos de la razón de verosimilitud. Este procedimiento está disponible en la opción Modelos avanzados.
- La regresión de Cox ahora ofrece la posibilidad de exportar la información del modelo a un archivo XML (PMML). Este procedimiento está disponible en la opción Modelos avanzados.
- Regresión de Cox de muestras complejas. Aplica la regresión de impactos proporcionales de Cox al análisis de los tiempos de supervivencia, es decir, el tiempo transcurrido antes de la aparición de un evento para muestras extraídas mediante métodos de muestreo complejo. Este procedimiento admite predictores continuos y categóricos, que pueden ser dependientes del tiempo. Proporciona un método sencillo para considerar las diferencias de

los subgrupos, además de analizar los efectos de un conjunto de predictores. Realiza una estimación de las varianzas teniendo en cuenta el diseño muestral empleado para seleccionar la muestra, incluidos los métodos de probabilidad igual y probabilidad proporcional al tamaño (PPS), y los procedimientos de muestreo con reposición (CR) o sin reposición (SR). Este procedimiento está disponible en la opción Muestras complejas.

**Ampliación mediante programación.** La ampliación mediante programación incluye las siguientes mejoras:

- R-Plugin. Combine la potencia de SPSS con la posibilidad de escribir sus propias rutinas estadísticas con R. Este complemento sólo se puede descargar de [www.spss.com/devcentral](http://www.spss.com/devcentral) (<http://www.spss.com/devcentral>).
- Estructuras anidadas del comando `Begin Program-End Program`.
- Posibilidad de crear y gestionar varios conjuntos de datos.

**Sintaxis de comandos.** Para obtener una lista completa de los cambios y adiciones de la sintaxis de comandos, consulte la sección *Historial de versiones* del capítulo *Introducción* en la referencia de sintaxis de comandos *Command Syntax Reference* (disponible en el menú Ayuda).

#### **Funciones ya no compatibles**

- Ya no existe un editor de gráficos independiente para los gráficos “interactivos”. Los gráficos creados a partir de los cuadros de diálogo de los gráficos “interactivos” existentes y de la sintaxis del comando `IGRAPH` se crean en el mismo formato que el resto de gráficos y se modifican en el mismo Editor de gráficos.
- Ya no están disponibles algunas de las funciones proporcionadas en los cuadros de diálogo de los gráficos “interactivos” existentes y la sintaxis del comando `IGRAPH`.
- Ya no está disponible el Visor de borrador.
- No es posible abrir los archivos del Visor creados en versiones anteriores de SPSS (archivos `.spo`) en SPSS 16.0. Para los sistemas operativos Windows, el CD de instalación incluye un Visor heredado que puede instalar para ver y editar archivos del Visor creados en versiones anteriores.
- Ya no está disponible la opción Mapas.
- Ya no están disponibles las interfaces de cuadro de diálogo para los procedimientos existentes en las opciones Tendencias y Tablas Para la opción Tendencias, esto incluye los siguientes comandos: `AREG`, `ARIMA` y `EXSMOOTH`. Para la opción Tablas, esto incluye el comando `TABLES`. Si tiene una licencia para cualquiera de estas opciones que incluya los procedimientos existentes, se seguirá admitiendo la sintaxis de comandos para estos comandos.

## **Windows**

Existen diversos tipos de ventanas en SPSS :

**Editor de datos.** El Editor de datos muestra el contenido del archivo de datos. Puede crear nuevos archivos de datos o modificar los existentes con el Editor de datos. Si tiene más de un archivo de datos abierto, habrá una ventana Editor de datos independiente para cada archivo.

**Visor.** Todas las tablas, los gráficos y los resultados estadísticos se muestran en el Visor. Puede editar los resultados y guardarlos para utilizarlos posteriormente. La ventana del Visor se abre automáticamente la primera vez que se ejecuta un procedimiento que genera resultados.

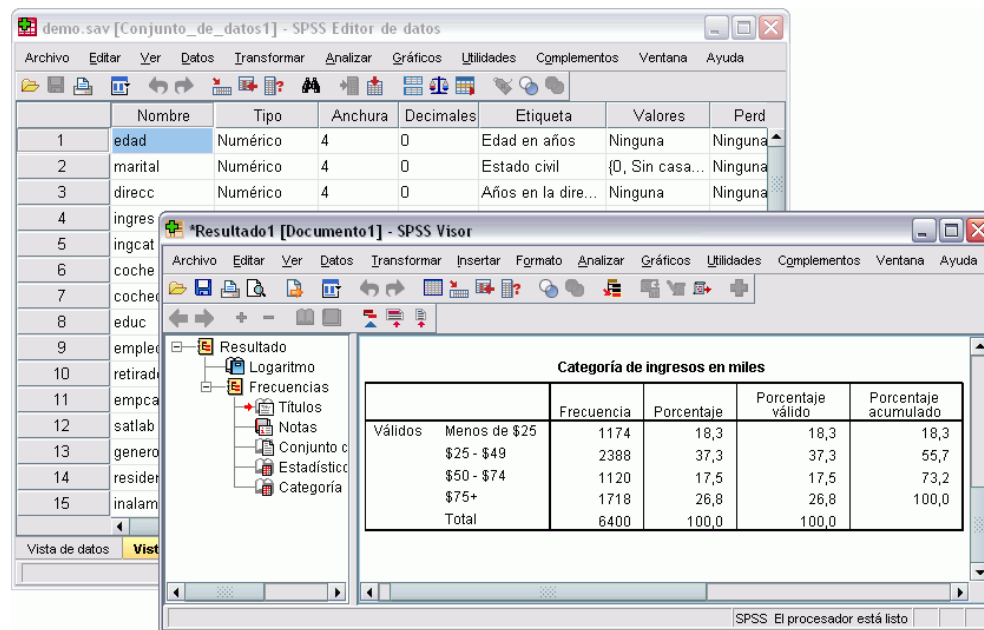
**Editor de tablas pivote.** Con el Editor de tablas pivote es posible modificar los resultados mostrados en este tipo de tablas de diversas maneras. Puede editar el texto, intercambiar los datos de las filas y las columnas, añadir colores, crear tablas multidimensionales y ocultar y mostrar los resultados de manera selectiva.

**Editor de gráficos.** Puede modificar los gráficos y diagramas de alta resolución en las ventanas de los gráficos. Es posible cambiar los colores, seleccionar diferentes tipos de fuentes y tamaños, intercambiar los ejes horizontal y vertical, rotar diagramas de dispersión 3-D e incluso cambiar el tipo de gráfico.

**Editor de resultados de texto.** Los resultados de texto que no aparecen en las tablas pivote pueden modificarse con el Editor de resultados de texto. Puede editar los resultados y cambiar las características de las fuentes (tipo, estilo, color y tamaño).

**Editor de sintaxis.** Puede pegar las selecciones del cuadro de diálogo en una ventana de sintaxis, donde aparecerán en forma de sintaxis de comandos. A continuación puede editar esta sintaxis de comandos para utilizar las funciones especiales que no se encuentran disponibles en los cuadros de diálogo. También puede guardar los comandos en un archivo para utilizarlos en sesiones posteriores.

Figura 1-1  
Editor de datos y Visor



## ***Ventana designada frente a ventana activa***

Si tiene abiertas varias ventanas del Visor, los resultados se dirigirán hacia la ventana **designada** del Visor. Si tiene abierta más de una ventana del Editor de sintaxis, la sintaxis de comandos se pegará en la ventana designada del Editor de sintaxis. Las ventanas designadas se indican con un signo más en el icono de la barra de título y es posible cambiarlas en cualquier momento.

La ventana designada no debe confundirse con la ventana **activa**, que es la ventana actualmente seleccionada. Si tiene ventanas superpuestas, la ventana activa es la que aparece en primer plano. Si abre una ventana, esa ventana se convertirá automáticamente en la ventana activa y en la ventana designada.

### ***Para cambiar la ventana designada del Visor o del Editor de sintaxis***

- ▶ Convierta la ventana que desee designar en la ventana activa (pulse en cualquier punto de la ventana).
- ▶ Pulse en el botón Designar ventana de la barra de herramientas (la que tiene el icono del signo más).  
*o*
- ▶ Elija en los menús:  
Utilidades  
Designar ventana

*Nota:* en cuanto a las ventanas Editor de datos, la ventana Editor de datos activa determina el conjunto de datos que se utiliza en análisis o cálculos posteriores. No hay ninguna ventana Editor de datos “designada”. Si desea obtener más información, consulte [Tratamiento básico de varios orígenes de datos](#) en Capítulo 6 en p. 103.

## ***Barra de estado***

La barra de estado que aparece en la parte inferior de cada ventana de SPSS proporciona la siguiente información:

**Estado del comando.** En cada procedimiento o comando que se ejecuta, un recuento de casos indica el número de casos procesados hasta el momento. En los procedimientos estadísticos que requieren procesamientos iterativos, se muestra el número de iteraciones.

**Estado del filtro.** Si ha seleccionado una muestra aleatoria o un subconjunto de casos para el análisis, el mensaje Filtrado indica que existe algún tipo de filtrado activado actualmente y por tanto en el análisis no se incluyen todos los casos del archivo de datos.

**Estado de ponderación.** El mensaje Ponderado indica que se está utilizando una variable de ponderación para ponderar los casos para el análisis.

**Estado de división del archivo.** El mensaje Segmentado indica que el archivo de datos se ha segmentado en diferentes grupos para su análisis en función de los valores de una o más variables de agrupación.

## Cuadros de diálogo

La mayoría de las opciones de los menús acceden a un cuadro de diálogo cuando se seleccionan. Los cuadros de diálogo se utilizan para seleccionar variables y opciones para el análisis.

Los cuadros de diálogo para los procedimientos estadísticos tienen normalmente dos componentes básicos:

**Lista de variables de origen.** Una lista de variables en el conjunto de datos activo. En la lista de origen sólo aparecen los tipos de variables que el procedimiento seleccionado permite. La utilización de variables de cadena corta y de cadena larga está restringida en muchos procedimientos.

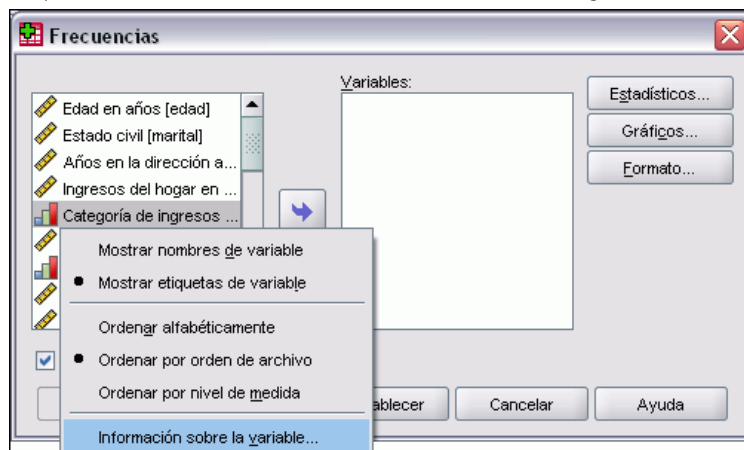
**Listas de variables de destino.** Una o varias listas que indican las variables elegidas para el análisis; un ejemplo son las listas de variables dependientes e independientes.

## Nombres y etiquetas de variable en las listas de los cuadros de diálogo

Puede mostrar tanto nombres como etiquetas de variable en las listas de los cuadros de diálogo y puede controlar el orden en el que aparecen las variables en las listas de variables de origen.

- Para controlar los atributos de presentación por defecto de las variables en las listas de origen, elija Opciones en el menú Editar. Si desea obtener más información, consulte [General: Opciones](#) en Capítulo 45 en p. 479.
- Para cambiar los atributos de presentación de las listas de variables de origen dentro de un cuadro de diálogo, pulse con el botón derecho del ratón en una variable de la lista de origen y seleccione los atributos de presentación en el menú contextual. Puede mostrar los nombres o las etiquetas de variable (los nombres se muestran para cualquier variable sin etiquetas definidas) y puede ordenar la lista de origen por orden de archivo, orden alfabético o nivel de medida. Si desea obtener más información sobre el nivel de medida, consulte [Iconos de tipo de datos, nivel de medida y lista de variables](#) en p. 9.

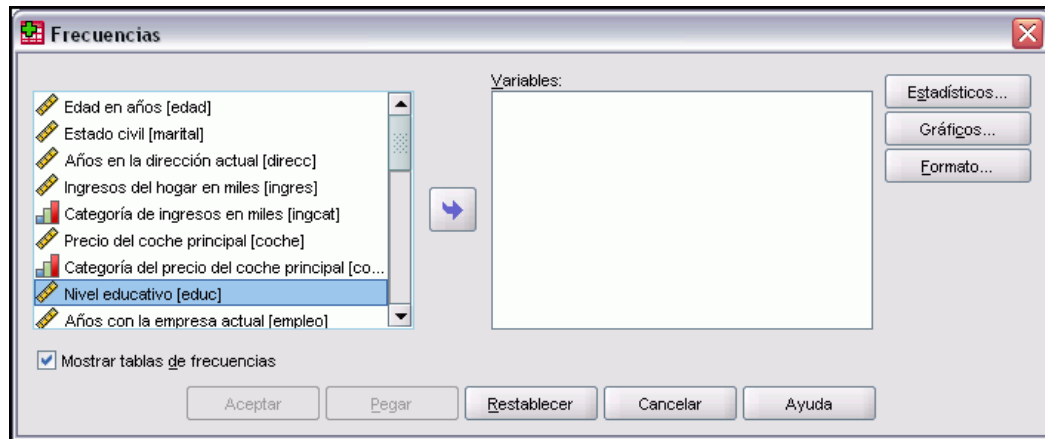
Figura 1-2  
Etiquetas de variable mostradas en un cuadro de diálogo



## Ajuste del tamaño de los cuadros de diálogo

Puede ajustar el tamaño de los cuadros de diálogo como si fueran ventanas, pulsando y arrastrando los bordes o esquinas exteriores. Por ejemplo, si aumenta el ancho del cuadro de diálogo, las listas de variables también serán más anchas.

Figura 1-3  
Cuadro de diálogo con tamaño ajustado



## Controles del cuadro de diálogo

Existen cinco controles estándares en la mayoría de los cuadros de diálogo:

**Aceptar.** Ejecuta el procedimiento. Después de seleccionar las variables y elegir las especificaciones adicionales, pulse en Aceptar para ejecutar el procedimiento y cerrar el cuadro de diálogo.

**Pegar.** Genera la sintaxis de comandos a partir de las selecciones del cuadro de diálogo y la pega en la ventana de sintaxis. A continuación, puede personalizar los comandos con funciones adicionales que no se encuentran disponibles en los cuadros de diálogo.

**Restablecer.** Desactiva las variables en las listas de variables seleccionadas y restablece todas las especificaciones del cuadro de diálogo y los subcuadros de diálogo al estado por defecto.

**Cancelar.** Cancela los cambios que se realizaron en las selecciones del cuadro de diálogo desde la última vez que se abrió y lo cierra. Durante una sesión se mantienen las selecciones del cuadro de diálogo. El cuadro de diálogo retiene el último conjunto de especificaciones hasta que se anulan.

**Ayuda.** Proporciona ayuda contextual. Este control le lleva a una ventana de Ayuda con información sobre el cuadro de diálogo actual.

## Selección de variables

Para seleccionar una única variable, simplemente selecciónela en la lista de variables de origen y arrástrela y colóquela en la lista de variables de destino. También puede utilizar el botón de dirección para mover las variables de la lista de origen a las listas de destino. Si sólo existe una



lista de variables de destino, puede pulsar dos veces en las variables individuales para desplazarlas desde la lista de origen hasta la de destino.

También pueden seleccionar varias variables:

- Para seleccionar varias variables que estén agrupadas en la lista de variables, pulse en la primera de ellas y, a continuación, en la última del grupo mientras mantiene pulsada la tecla Mayús.
- Para seleccionar varias variables que no estén agrupadas en la lista de variables, pulse en la primera de ellas; a continuación, pulse en la siguiente variable mientras mantiene pulsada la tecla Ctrl, y así sucesivamente (en Macintosh, mantenga pulsada la tecla Comando).

## ***Iconos de tipo de datos, nivel de medida y lista de variables***

Los iconos que se muestran junto a las variables en las listas de los cuadros de diálogo proporcionan información acerca del tipo de variable y el nivel de medidas.

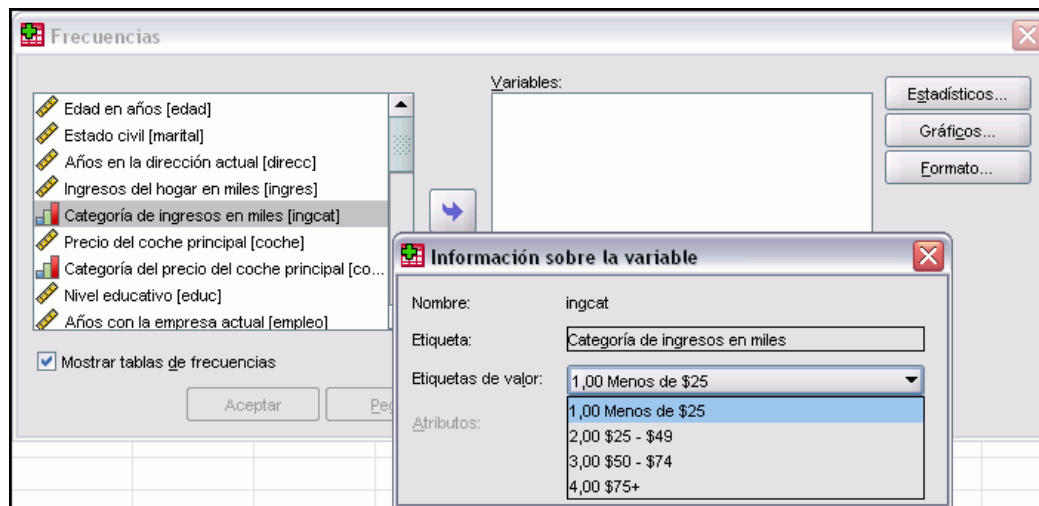
Nivel de medida	Tipo de datos			
	Numérico	Cadena	Fecha	Time
Escala		n/a		
Ordinal				
Nominal				

- Si desea obtener más información sobre el nivel de medida, consulte [Nivel de medida de variable en p. 83](#).
- Si desea obtener más información sobre los tipos de dato numérico, cadena, fecha y hora, consulte [Tipo de variable en p. 83](#).

## ***Información sobre las variables en un cuadro de diálogo***

- ▶ Pulse con el botón derecho del ratón en la lista de variables de origen o de destino.
- ▶ Seleccione Información sobre la variable.

Figura 1-4  
Información sobre la variable



## ***Pasos básicos en el análisis de datos***

Analizar datos con SPSS es fácil. Simplemente tiene que:

**Introducir los datos en SPSS .** Es posible abrir un archivo de datos de SPSS previamente guardado, leer una hoja de cálculo, una base de datos o un archivo de datos de texto, o introducir los datos directamente en el Editor de datos.

**Seleccionar un procedimiento.** Seleccione un procedimiento de los menús para calcular estadísticos o crear un gráfico.

**Seleccionar las variables para el análisis.** Las variables del archivo de datos se muestran en un cuadro de diálogo para el procedimiento.

**Ejecute el procedimiento y observe los resultados.** Los resultados se muestran en el Visor.

## ***Asesor estadístico***

Si no está familiarizado con SPSS o con los procedimientos estadísticos disponibles, el Asesor estadístico puede ayudarle solicitándole información mediante preguntas simples, utilizando un lenguaje no técnico y ejemplos visuales que le ayudarán a seleccionar las funciones estadísticas y los gráficos más apropiados para sus datos.

Para utilizar el Asesor estadístico, elija en los menús de cualquier ventana de SPSS :

?

Asesor estadístico

El Asesor estadístico cubre sólo un subconjunto selecto de procedimientos del sistema Base. Está diseñado para proporcionar una asistencia general para muchas de las técnicas estadísticas básicas que se utilizan habitualmente.

## ***Información adicional***

Si desea obtener una introducción global más detallada a los conceptos básicos, consulte el tutorial en pantalla. En cualquier menú de SPSS , elija:

?

Tutorial

## ***Obtención de ayuda***

La ayuda se proporciona de diversas formas:

**Menú Ayuda.** En la mayoría de las ventanas de SPSS, el menú Ayuda proporciona acceso al sistema de ayuda principal además de a los tutoriales y al material de referencia técnica.

- **Temas.** Proporciona acceso a las pestañas Contenido, Índice y Buscar, que pueden usarse para buscar temas específicos de la Ayuda.
- **Tutorial.** Instrucciones ilustradas paso a paso sobre cómo utilizar muchas de las funciones básicas de SPSS. No necesita consultar el tutorial de principio a fin. Puede elegir los temas que desea consultar, pasar de un tema a otro, ver los temas en cualquier orden y utilizar el índice o la tabla de contenido para buscar temas concretos.
- **Estudios de casos.** Ejemplos prácticos sobre cómo crear diferentes tipos de análisis estadísticos y cómo interpretar los resultados. También se proporcionan los archivos de datos de muestra utilizados en estos ejemplos para que pueda trabajar en dichos ejemplos y observar con exactitud cómo se generaron los resultados. Puede elegir los procedimientos concretos que desee aprender en la tabla de contenido o buscar los temas correspondientes en el índice.
- **Asesor estadístico.** Método de asistencia para orientarle en el proceso de búsqueda del procedimiento que desea utilizar. Tras realizar una serie de selecciones, el Asesor estadístico abre el cuadro de diálogo para el procedimiento estadístico, de generación de informes o de creación de gráficos que cumple los criterios seleccionados. El Asesor estadístico proporciona acceso a la mayoría de los procedimientos estadísticos y de generación de informes en el sistema Base y en los procedimientos de creación de gráficos.
- **Referencia de sintaxis de comandos** La información de referencia detallada sobre la sintaxis de comandos está disponible en dos formatos: integrada en el sistema de ayuda global y como un documento independiente en formato PDF en *SPSS Command Syntax Reference*, disponible en el menú Ayuda.
- **Algoritmos de estadísticos.** Los algoritmos que se utilizan para la mayoría de los procedimientos están disponibles en dos formatos: integrada en el sistema de ayuda global y como un documento independiente en formato PDF en *SPSS Algorithms*, disponible en el CD de manuales. Para los vínculos a algoritmos específicos en el sistema de ayuda, seleccione Algoritmos en el menú Ayuda.

**Ayuda contextual.** En muchos puntos de la interfaz de usuario, puede obtener ayuda sensible al contexto.

- **Botón Ayuda en los cuadros de diálogo.** La mayoría de los cuadros de diálogo disponen de un botón Ayuda que permite acceder directamente al tema de ayuda correspondiente. Este tema proporciona información general y vínculos a los temas relacionados.

- **Ayuda en el menú contextual de una tabla pivote.** Pulse con el botón derecho del ratón en los términos de una tabla pivote activada en el Visor y seleccione *¿Qué es esto?* en el menú contextual para ver las definiciones de los términos.
- **Sintaxis de comandos.** En una ventana de sintaxis de comandos, coloque el cursor en cualquier punto de un bloque de sintaxis para un comando y pulse F1 en el teclado. Se muestra un esquema de sintaxis de comandos completo para dicho comando. La documentación completa sobre la sintaxis de comandos está disponible en los vínculos de la lista de temas relacionados y en la pestaña Contenido de la ayuda.

### **Otros recursos**

**Sitio Web del servicio técnico.** Se pueden encontrar respuestas a muchos problemas comunes en <http://support.spss.com>. (El sitio Web de servicio técnico requiere un ID de inicio de sesión y una contraseña. La información sobre cómo obtener el ID y la contraseña se facilita en la dirección URL mencionada anteriormente.)

**SPSS Developer Central.** Developer Central incluye recursos para todos los niveles de usuarios y desarrolladores de la aplicación SPSS. Descargue utilidades, ejemplos de gráficos, nuevos módulos estadísticos y artículos sobre la tecnología SPSS. Utilice los foros para interactuar con SPSS y la comunidad de usuarios de SPSS. Visite Developer Central de SPSS en <http://www.spss.com/devcentral>.

## **Obtención de ayuda sobre los términos de los resultados**

Para ver la definición de un término de los resultados de la tabla pivote en el Visor:

- ▶ Pulse dos veces en la tabla pivote para activarla.
- ▶ Pulse con el botón derecho del ratón en el término del que desee obtener información.
- ▶ Seleccione *¿Qué es esto?* en el menú contextual.

Aparecerá una definición del término en una ventana emergente.

Figura 2-1

Ayuda de glosario de la tabla pivote activada con el botón derecho del ratón

viewertut2 - Visor SPSS

Archivo Edición Ver Insertar Pivotar Formato Analizar Gráficos Utilidades Ventana ?

**Tabla de contingencia Tiene PDA \* Sexo \* Internet**

Internet			Sexo		Total
			Hombre	Mujer	
Sí	Tiene PDA	Recuento	1962	1897	3859
		Frecuencia esperada	1955,6	1903,4	3859,0
	% de Tiene PDA	50,8%	49,2%	100,0%	
		323	327	650	
No	Tiene PDA	Recuento	29,4	320,6	650,0
		Frecuencia esperada	29,4	320,6	650,0
	% de Tiene PDA	49,7%	50,3%	100,0%	
		323	327	650	
Total	Tiene PDA	Recuento	2285	2224	4509
		Frecuencia esperada	2285,0	2224,0	4509,0
	% de Tiene PDA	50,7%	49,3%	100,0%	
		512	513	1025	
No	Tiene PDA	Recuento	512	513	1025
		Frecuencia esperada	512,5	512,5	1025,0
	% de Tiene PDA	50,0%	50,0%	100,0%	
		306	305	611	
Total	No	Recuento	306	305	611
		Frecuencia esperada	305,5	305,5	611,0
	% de Tiene PDA	50,1%	49,9%	100,0%	
		306	305	611	

Muestra el porcentaje de la variable que cae en la casilla definida.

Preparado SPSS El procesador está preparado

# Archivos de datos

Los archivos de datos pueden tener formatos muy diversos, y este programa se ha sido diseñado para trabajar con muchos de ellos, incluyendo:

- Hojas de cálculo creadas con Excel y Lotus
- Tablas de bases de datos de diferentes orígenes de bases de datos, incluido Oracle, SQLServer, Access, dBASE, entre otros
- Archivos de texto delimitados por tabuladores y otros tipos de archivos de sólo texto
- Archivos de datos con formato SPSS creados en otros sistemas operativos.
- Archivos de datos de SYSTAT.
- archivos de datos de SAS
- archivos de datos de Stata

## Apertura de archivos de datos

Además de los archivos guardados en formato SPSS, puede abrir archivos de Excel, SAS, Stata, archivos delimitados por tabuladores y otros archivos sin necesidad de convertirlos a un formato intermedio ni de introducir información sobre la definición de los datos.

- Abre un archivo de datos y lo convierte en el conjunto de datos activo. Si ya ha abierto uno o más archivos de datos, permanecerán abiertos y disponibles para su uso posterior durante la sesión. Al pulsar en cualquier punto de la ventana Editor de datos de un archivo de datos abierto lo convertirá en el conjunto de datos activo. Si desea obtener más información, consulte [Trabajo con varios orígenes de datos](#) en Capítulo 6 en p. 102.
- En el análisis en modo distribuido donde un servidor remoto procesa los comandos y ejecuta los procedimientos, las unidades, carpetas y archivos de datos disponibles dependen de lo que esté disponible en el servidor remoto. En la parte superior del cuadro de diálogo se indica el nombre del servidor actual. Sólo tendrá acceso a los archivos de datos del equipo local si especifica la unidad como un dispositivo compartido y las carpetas que contienen los archivos de datos como carpetas compartidas. Si desea obtener más información, consulte [Análisis en modo distribuido](#) en Capítulo 4 en p. 71.

## Para abrir archivos de datos

- ▶ Elija en los menús:  
Archivo  
Abrir  
Datos...

- ▶ En el cuadro de diálogo Abrir datos, seleccione el archivo que desea abrir.
- ▶ Pulse en Abrir.

Si lo desea, puede:

- Establecer de forma automática la longitud de cada variable de cadena en el valor más largo observado para dicha variable mediante Minimizar longitudes de cadena en función de los valores observados. Esto es especialmente útil cuando se leen archivos de datos de página de código en modo Unicode. Si desea obtener más información, consulte [General: Opciones](#) en Capítulo 45 en p. 479.
- Leer los nombres de las variables de la primera fila de los archivos de hoja de cálculo.
- Especificar el rango de casillas que desee leer en los archivos de hojas de cálculo.
- Especificar una hoja de trabajo dentro de un archivo de Excel que desee leer (Excel 95 o versiones posteriores).

Si desea obtener información sobre la lectura de datos de bases de datos, consulte [Lectura de archivos de bases de datos en p. 18](#). Si desea obtener información sobre la lectura de datos de archivos de datos de texto, consulte [Asistente para la importación de texto en p. 32](#).

## ***Tipos de archivos de datos***

**SPSS.** Abre archivos de datos guardados con formato SPSS y también el producto SPSS/PC+ para DOS.

**SPSS/PC+.** Abre archivos de datos de SPSS/PC+.

**SYSTAT.** Abre archivos de datos de SYSTAT.

**SPSS portátil.** Abre archivos de datos guardados con formato portátil. El almacenamiento de archivos en este formato lleva mucho más tiempo que guardarlos en formato SPSS.

**Excel.** Abre archivos de Excel.

**Lotus 1-2-3.** Abre archivos de datos guardados en formato 1-2-3 en las versiones 3.0, 2.0 o 1A de Lotus.

**SYLK.** Abre archivos de datos guardados en formato SYLK (vínculo simbólico), un formato utilizado por algunas aplicaciones de hoja de cálculo.

**dBASE.** Abre archivos con formato dBASE para dBASE IV, dBASE III o III PLUS, o dBASE II. Cada caso es un registro. Las etiquetas de valor y de variable y las especificaciones de valores perdidos se pierden si se guarda un archivo en este formato.

**SAS.** versiones 6-9 de SAS y archivos de transporte SAS.

**Stata.** Versiones 4-8 de Stata.



## ***Opciones de apertura de archivos***

**Leer los nombres de variable.** En las hojas de cálculo, puede leer los nombres de variable de la primera fila del archivo o de la primera fila del rango definido. Los valores se convertirán según sea preciso para crear nombres de variables válidos, incluyendo la conversión de espacios en subrayados.

**Hoja de trabajo.** Los archivos de Excel 95 o de versiones posteriores pueden contener varias hojas de trabajo. El Editor de datos lee por defecto la primera hoja. Para leer una diferente, seleccione la que desee en la lista desplegable.

**Rango.** En los archivos de hoja de cálculo, también puede leer un rango de casillas. Para especificar rangos de casillas utilice el mismo método que empleará en la aplicación de hoja de cálculo.

## ***Lectura de archivos de Excel 95 o versiones posteriores***

Siga las siguientes normas para leer archivos de Excel 95 o de versiones posteriores:

**Tipo y ancho de datos.** Cada columna es una variable. El tipo de datos y el ancho de cada variable está determinado por el tipo de datos y el ancho en el archivo de Excel. Si la columna contiene más de un tipo de datos (por ejemplo, fecha y numérico), el tipo de datos se define como cadena y todos los valores se leen como valores de cadena válidos.

**Casillas en blanco.** En las variables numéricas, las casillas en blanco se convierten en el valor perdido del sistema indicado por un punto (o una coma). En las variables de cadena, los espacios en blanco son valores de cadena válidos y las casillas en blanco se tratan como valores de cadena válidos.

**Nombres de variable.** Si lee la primera fila del archivo de Excel (o la primera fila del rango especificado) como nombres de variable, los valores que no cumplan las normas de denominación de variables se convertirán en nombres de variables válidos y los nombres originales se utilizarán como etiquetas de variable. Si no lee nombres de variable del archivo de Excel, se asignarán nombres de variable por defecto.

## ***Lectura de archivos de Excel y otras hojas de cálculo de versiones anteriores***

Las siguientes normas se aplican para leer archivos de Excel de versiones anteriores a Excel 95 y otros datos de hoja de cálculo:

**Tipo y ancho de datos.** El tipo y el ancho de los datos para cada variable se determinan según el ancho de la columna y el tipo de datos de la primera casilla de la columna. Los valores de otro tipo se convierten en valor perdido por el sistema. Si la primera casilla de datos de la columna está en blanco, se utiliza el tipo de datos global por defecto para la hoja de cálculo (normalmente numérico).

**Casillas en blanco.** En las variables numéricas, las casillas en blanco se convierten en el valor perdido del sistema indicado por un punto (o una coma). En las variables de cadena, los espacios en blanco son valores de cadena válidos y las casillas en blanco se tratan como valores de cadena válidos.

**Nombres de variable.** Si no se leen los nombres de variable de la hoja de cálculo, se utilizan las letras de las columnas (*A, B, C, ...*) como nombres de variable de los archivos de Excel y de Lotus. Para los archivos de SYLK y de Excel guardados en el formato de presentación R1C1, el programa utiliza para los nombres de variable el número de la columna precedido por la letra *C* (*C1, C2, C3, ...*).

### ***Lectura de archivos de dBASE***

Los archivos de bases de datos son, lógicamente, muy similares a los archivos de datos con formato SPSS. Las siguientes normas generales se aplican a los archivos de dBASE:

- Los nombres de campo se convierten en nombres de variable válidos.
- Los dos puntos en los nombres de campo de dBASE se convierten en subrayado.
- Se incluyen los registros marcados para ser eliminados que aún no se han purgado. El programa crea una nueva variable de cadena, *D\_R*, que incluye un asterisco en los casos marcados para su eliminación.

### ***Lectura de archivos de Stata***

Las siguientes normas generales se aplican a los archivos de Stata:

- **Nombres de variables.** Los nombres de variable de Stata se convierten en nombres de variable de SPSS en formato que distingue entre mayúsculas y minúsculas. Los nombres de variable de Stata que sólo se diferencian en el uso de las mayúsculas y minúsculas se convierten en nombres de variable válidos añadiendo un subrayado y una letra secuencial (*\_A, \_B, \_C, ..., \_Z, \_AA, \_AB, ..., etc.*).
- **Etiquetas de variable.** Las etiquetas de variable de Stata se convierten en etiquetas de variable de SPSS .
- **Etiquetas de valor.** Las etiquetas de valor de Stata se convierten en etiquetas de valor de SPSS , excepto las etiquetas de valor de Stata asignadas a valores perdidos “extendidos”.
- **Valores perdidos.** Los valores perdidos “extendidos” de Stata se convierten en valores perdidos del sistema.
- **Conversión de fechas.** Los valores de formato de fecha de Stata se convierten en valores con formato `DATE` (d-m-a) de SPSS . Los valores de formato de fecha de “serie temporal” de Stata (semanas, meses, trimestres, etc.) se convierten a formato numérico simple (F), conservando el valor entero interno original, que es el número de semanas, meses, trimestres, etc., desde el inicio de 1960.

### ***Lectura de archivos de bases de datos***

Podrá leer los datos desde cualquier formato de base de datos para los que disponga de un controlador de base de datos. En el análisis en modo local, los controladores necesarios deben estar instalados en el ordenador local. En el análisis en modo distribuido (disponible con SPSS Server), los controladores deben estar instalados en el servidor remoto. Si desea obtener más información, consulte [Análisis en modo distribuido](#) en Capítulo 4 en p. 71.

**Para leer archivos de base de datos**

- ▶ Elija en los menús:
  - Archivo
  - Abrir base de datos
  - Nueva consulta...
- ▶ Seleccione el origen de datos.
- ▶ Si es necesario (según el origen de datos), seleccione el archivo de base de datos y/o escriba un nombre de acceso, contraseña y demás información.
- ▶ Seleccione las tablas y los campos. Para los orígenes de datos OLE DB (sólo disponibles en los sistemas operativos Windows), únicamente puede seleccionar una tabla.
- ▶ Especifique cualquier relación existente entre las tablas.
- ▶ Si lo desea:
  - Especifique cualquier criterio de selección para los datos.
  - Añada un mensaje solicitando al usuario que introduzca datos para crear una consulta con parámetros.
  - Guarde la consulta creada antes de ejecutarla.

**Para editar una consulta de base de datos guardada**

- ▶ Elija en los menús:
  - Archivo
  - Abrir base de datos
  - Editar consulta...
- ▶ Seleccione el archivo de consulta (\*.spq) que desee editar.
- ▶ Siga las instrucciones para crear una consulta.

**Para leer archivos de bases de datos con una consulta ODBC guardada**

- ▶ Elija en los menús:
  - Archivo
  - Abrir base de datos
  - Ejecutar consulta...
- ▶ Seleccione el archivo de consulta (\*.spq) que desee ejecutar.
- ▶ Si es necesario (según el archivo de base de datos), introduzca un nombre de acceso y una contraseña.
- ▶ Si la consulta tiene una solicitud incrustada, introduzca otra información necesaria (por ejemplo, el trimestre para el que desee obtener cifras de ventas).

### Selección de un origen de datos

Utilice la primera pantalla del Asistente para bases de datos para seleccionar el tipo de origen de datos que se leerá.

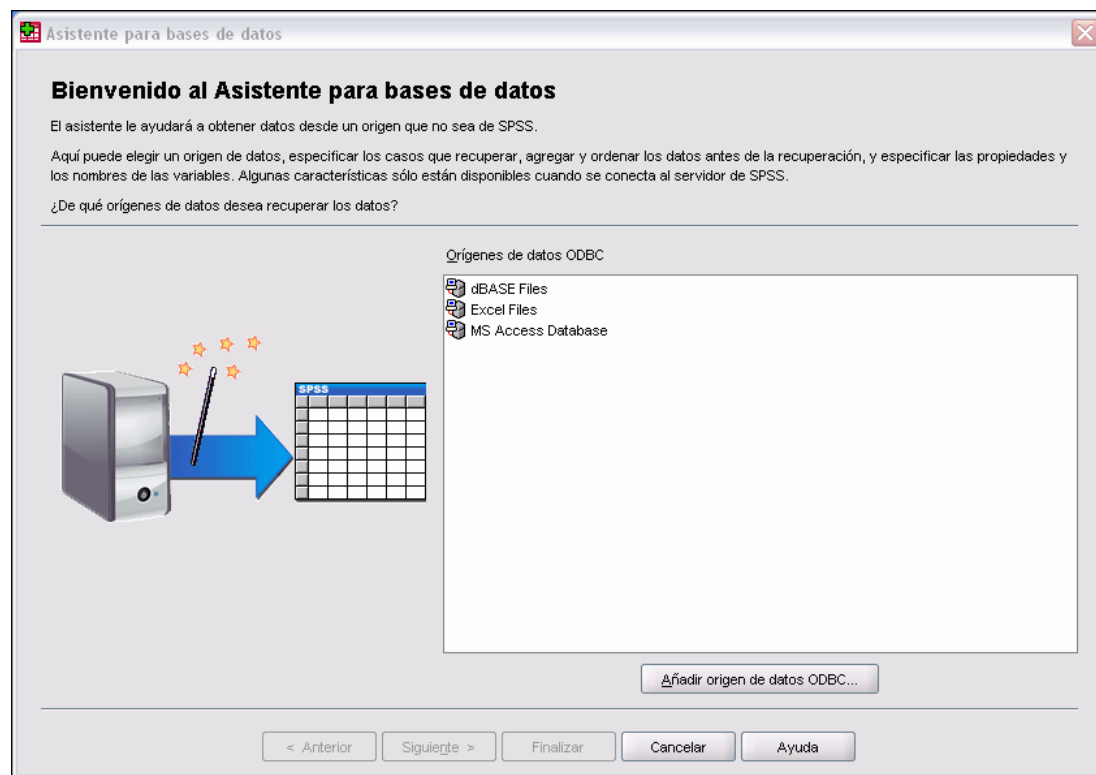
#### Orígenes de datos ODBC

Si no tiene configurado ningún origen de datos ODBC o si desea añadir uno nuevo, pulse en Añadir origen de datos ODBC.

- En los sistemas operativos Linux, este botón no está disponible. Los orígenes de datos ODBC se especifican en *odbc.ini* y es necesario especificar las variables de entorno *ODBCINI* con la ubicación de dicho archivo. Si desea obtener más información, consulte la documentación de los controladores de la base de datos.
- En el análisis en modo distribuido (disponible con SPSS Server), este botón no está disponible. Para añadir orígenes de datos en el análisis en modo distribuido, consulte con el administrador del sistema.

Un origen de datos ODBC está compuesto por dos partes esenciales de información: el controlador que se utilizará para acceder a los datos y la ubicación de la base de datos a la que se desea acceder. Para especificar los orígenes de datos, deberán estar instalados los controladores adecuados. Para el análisis en modo local, puede instalar los controladores de varios formatos de bases de datos desde el CD de instalación de SPSS .

Figura 3-1  
Asistente para bases de datos



### Orígenes de datos OLE DB

Para acceder a los orígenes de datos OLE DB (sólo disponibles en los sistemas operativos Windows), debe tener instalados los siguientes elementos:

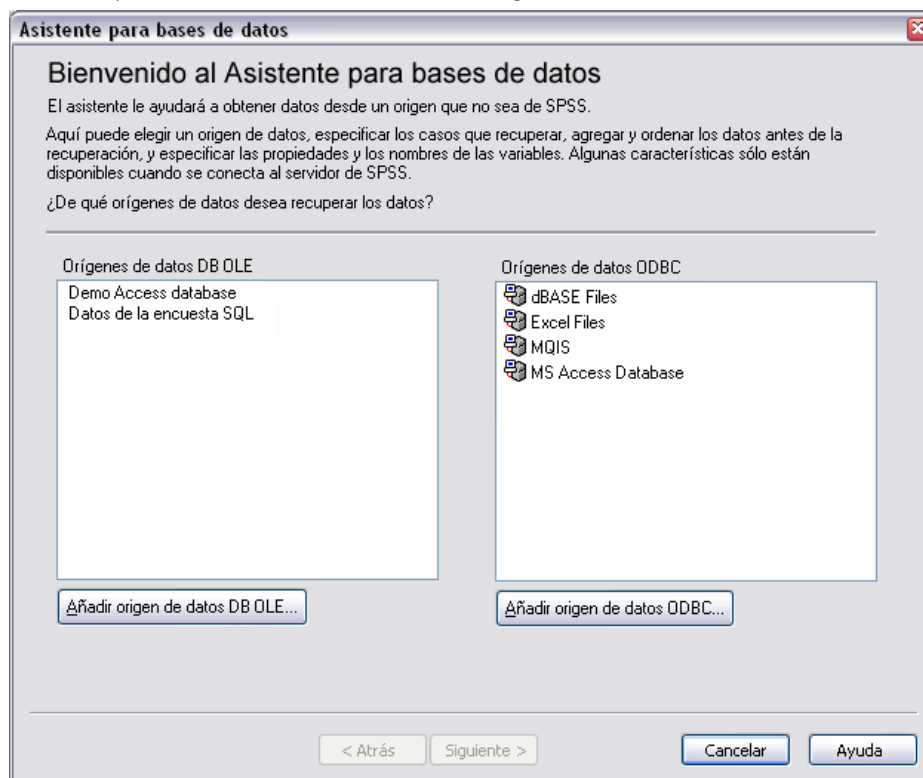
- .NET framework
- Dimensions Data Model y OLE DB Access

Las versiones de estos componentes compatibles con esta versión se pueden instalar desde el CD de instalación y están disponibles en el menú de reproducción automática.

- Las uniones entre tablas no están disponibles para los orígenes de datos OLE DB. Sólo se puede leer una tabla al mismo tiempo.
- Se pueden añadir orígenes de datos OLE DB en análisis en modo local. Para añadir orígenes de datos OLE DB en el análisis en modo distribuido en un servidor Windows, consulte con el administrador del sistema.
- En el análisis en modo distribuido (disponible con SPSS Server), los orígenes de datos OLE DB sólo están disponibles en servidores Windows, y debe tener instalado en el servidor .NET, Dimensions Data Model y OLE DB Access.

Figura 3-2

Asistente para bases de datos con acceso a orígenes de datos OLE DB



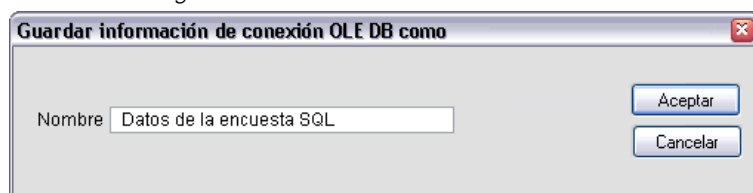
Para añadir un origen de datos OLE DB:

- Pulse en Añadir origen de datos OLE DB.

- ▶ En las propiedades del vínculo de datos, pulse en la pestaña Proveedor y seleccione el proveedor OLE DB.
- ▶ Pulse en Siguiente o en la pestaña Conexión.
- ▶ Seleccione la base de datos introduciendo la ubicación del directorio y el nombre de base de datos o pulsando en el botón para desplazarse hasta una base de datos. (Puede que también sean necesarios un nombre de usuario y una contraseña.)
- ▶ Pulse en Aceptar una vez que haya escrito toda la información necesaria. (Puede comprobar si la base de datos especificada está disponible pulsando en el botón Probar conexión.)
- ▶ Escriba un nombre para la información de conexión de base de datos. (Este nombre se mostrará en la lista de orígenes de datos OLE DB disponibles.)

Figura 3-3

Cuadro de diálogo Guardar información de conexión OLE DB como



- ▶ Pulse en Aceptar.

Volverá a la primera pantalla del Asistente para bases de datos, donde puede seleccionar el nombre guardado de la lista de orígenes de datos OLE DB y continuar con el siguiente paso del asistente.

### **Eliminación de orígenes de datos OLE DB**

Para eliminar nombres de orígenes de datos de la lista de orígenes de datos OLE DB, elimine el archivo UDL que contiene el nombre del origen de datos:

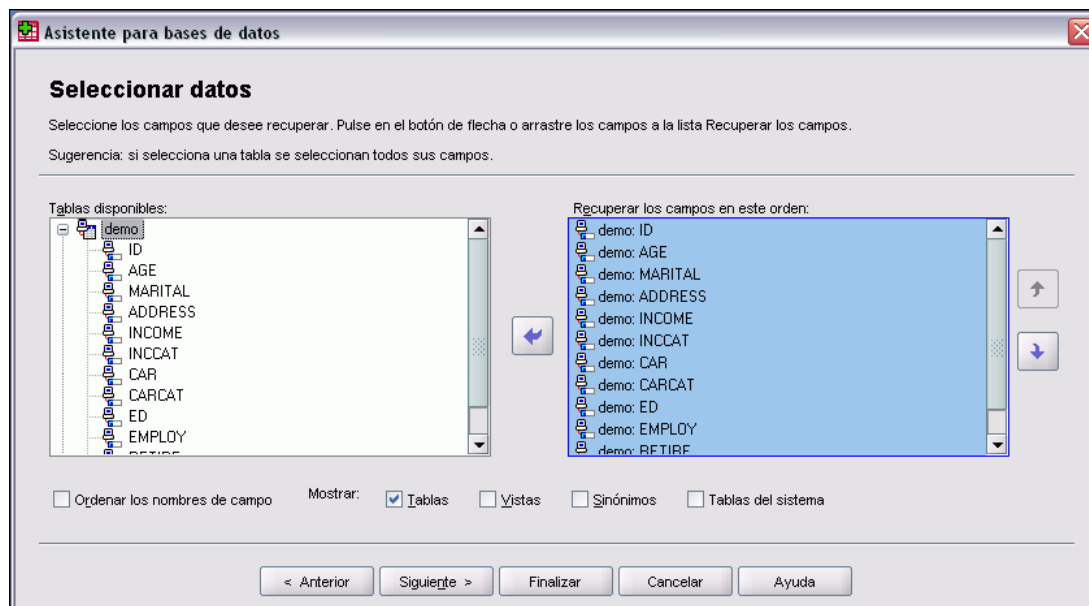
```
[unidad]:\Documents and Settings\[nombre de usuario]\Local Settings\Application Data\SPSS\UDL
```

### **Selección de campos de datos**

El paso de selección de datos controla las tablas y los campos que se deben leer. Los campos (las columnas) de la base de datos se leen como variables.

Si una tabla tiene un campo cualquiera seleccionado, todos sus campos serán visibles en las ventanas subsiguientes del Asistente para bases de datos; sin embargo, sólo se importarán como variables los campos seleccionados en este paso. Esto le permitirá crear uniones entre tablas y especificar criterios empleando los campos que no esté importando.

Figura 3-4  
Asistente para bases de datos, selección de datos



**Presentación de los nombres de los campos.** Para ver los campos de la tabla, pulse en el signo más (+) situado a la izquierda del nombre de una tabla. Para ocultar los campos pulse en el signo menos (–) situado a la izquierda del nombre de una tabla.

**Para añadir un campo.** Pulse dos veces en cualquier campo de la lista Tablas disponibles o arrástrelo hasta la lista Recuperar los campos en este orden. Los campos se pueden volver a ordenar arrastrándolos y colocándolos dentro de la lista de campos.

**Para eliminar un campo.** Pulse dos veces en cualquier campo de la lista Recuperar los campos en este orden, o bien arrástrelo hasta la lista Tablas disponibles.

**Ordenar los nombres de campo.** Si se selecciona, el Asistente para bases de datos mostrará los campos disponibles en orden alfabético.

Por defecto, la lista muestra sólo las tablas disponibles de bases de datos estándar. Puede controlar el tipo de elementos que se muestran en la lista:

- **Tablas.** Tablas de base de datos estándar.
- **Vistas.** Las vistas son “tablas” virtuales o dinámicas definidas por consultas. Estas tablas pueden incluir uniones de varias tablas y/o campos derivados de cálculos basados en los valores de otros campos.
- **Sinónimos.** Un sinónimo es un alias para una tabla o vista que suele estar definido en una consulta.
- **Tablas del sistema.** Las tablas del sistema definen propiedades de la base de datos. En algunos casos, las tablas de base de datos estándar pueden estar clasificadas como tablas del sistema y sólo se mostrarán si se selecciona esta opción. El acceso a tablas del sistema reales suele estar limitado a los administradores de la base de datos.

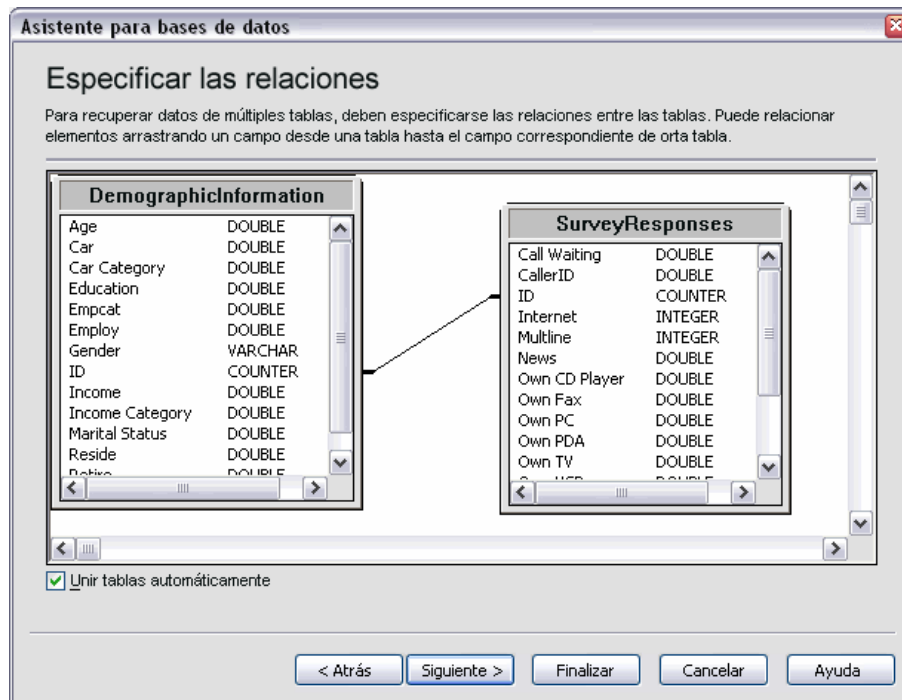
*Nota:* para los orígenes de datos OLE DB (sólo disponibles en los sistemas operativos Windows), únicamente puede seleccionar los campos de una sola tabla. Las uniones entre varias tablas no son compatibles con los orígenes de datos OLE DB.

### Creación de una relación entre tablas

El paso Especificar relaciones permite definir relaciones entre las tablas para orígenes de datos ODBC. Si selecciona campos de más de una tabla, deberá definir al menos una unión.

Figura 3-5

Asistente para bases de datos, especificación de relaciones



**Establecimiento de relaciones.** Para crear una relación, arrastre un campo desde cualquier tabla hasta el campo con el que quiera unirlo. El Asistente para bases de datos dibujará una **línea de unión** entre los dos campos que indica su relación. Estos campos deben ser del mismo tipo de datos.

**Unir tablas automáticamente.** Intenta unir las tablas automáticamente en función de las claves primarias/externas o de los nombres de campo y tipos de datos coincidentes.

**Tipo de unión.** Si el controlador permite uniones exteriores, podrá especificar uniones interiores, uniones exteriores izquierdas o uniones exteriores derechas.

- **Uniones interiores.** Una unión interior incluye sólo las filas donde los campos relacionados son iguales. En este ejemplo, se incluirán todas las filas con los mismos valores de ID en las dos tablas.
- **Uniones exteriores.** Además de las coincidencias de uno a uno con uniones interiores, también puede utilizar uniones exteriores para fusionar tablas con un esquema de coincidencia de uno a varios. Por ejemplo, puede hacer una coincidencia con una tabla donde sólo hay algunos



registros que representan los valores de datos y las etiquetas descriptivas asociadas, con valores en una tabla que contiene cientos o miles de registros que representan los encuestados. Una unión exterior izquierda incluye todos los registros de la tabla izquierda y sólo aquellos registros de la tabla derecha en los que los campos relacionados son iguales. En una unión exterior derecha, se importan todos los registros de la tabla derecha y sólo aquellos registros de la tabla izquierda en los que los campos relacionados son iguales.

### Limitar la recuperación de casos

Este paso permite especificar el criterio para seleccionar subconjuntos de casos (filas). La limitación de los casos consiste generalmente en rellenar la rejilla de criterios con uno o varios criterios. Los criterios constan de dos expresiones y de alguna relación entre ellas, y devuelven un valor *verdadero*, *falso* o *perdido* para cada caso.

- Si el resultado es *verdadero*, se selecciona el caso.
- Si el resultado es *falso* o *perdido*, no se selecciona el caso.
- La mayoría de los criterios utiliza al menos uno de los seis operadores de relación (<, >, <=, >=, = y <>).
- Las expresiones pueden incluir nombres de campo, constantes, operadores aritméticos, funciones numéricas y de otros tipos, y variables lógicas. Puede utilizar como variables los campos que no vaya a importar.

Figura 3-6  
Asistente para bases de datos, limitación de recuperación de casos

**Asistente para bases de datos**

**Limitar la recuperación de casos**

Es posible limitar el número de casos recuperados especificando criterios de selección.  
Sugerencia: los campos y funciones se pueden añadir a una expresión arrastrándolos y colocándolos en una casilla.  
Consulte la ayuda sobre cómo expresar los valores de fecha y hora.

**Campos:**

- demo
  - ID
  - AGE
  - MARITAL
  - ADDRESS
  - INCOME
  - INCCAT
  - CAP

**Funciones**

- ASCII(strexpr)
- CHAR(asciiode)
- CONCAT(strexpr, strexpr)
- INSERT(strexpr, pos, len, new)
- LEFT(strexpr, count)
- LTRIM(strexpr)
- LENGTH(strexpr)
- LOCATE(haystack, needle)

Pedir el valor al usuario...

**Criterios:**

	Conector	Expresión 1	Relación	Expresión 2
1				
2				
3				
4				
5				
6				

Casos de muestra aleatorios

**Método de muestreo**

- Recuperar casos y seleccionar aleatoriamente en SPSS
- Seleccionar aleatoriamente en la base de datos y recuperar en SPSS

Sugerencia: si recupera los casos y los selecciona en SPSS, la función de agregación de la base de datos no está disponible.

**Tamaño de la muestra**

- Porcentaje aproximado de todos los casos  
Porcentaje: 10
- Número exacto de casos  
Número de casos: 100  
Desde primer: 1000 Casos

< Anterior    Siguiente >    Finalizar    Cancelar    Ayuda

Para crear sus criterios necesita por lo menos dos expresiones y una relación para conectarlas.

- ▶ Para crear una expresión, seleccione uno de los siguientes métodos:
  - En una celda Expresión, puede escribir nombres de campo, constantes, operadores aritméticos, funciones numéricas y de otro tipo, y variables lógicas.
  - Pulse dos veces en el campo de la lista Campos.
  - Arrastre el campo de la lista Campos hasta la celda Expresión.
  - Seleccione un campo del menú desplegable en una celda Expresión activa.
- ▶ Para elegir el operador relacional (como = o >), sitúe el cursor en la casilla Relación y escriba el tipo de operador o selecciónelo en el menú desplegable.

Si SQL contiene las cláusulas `WHERE` con expresiones para la selección de casos, las fechas y las horas de las expresiones deberán especificarse de un modo especial (incluidas las llaves que se muestran en los ejemplos:)

- Los literales de fecha se deben especificar usando el formato general `{d 'aaaa-mm-dd'}`.
- Los literales de hora deben especificarse usando el formato general `{t 'hh:mm:ss'}`.
- Los literales de fecha y hora (marcas de hora) se deben especificar usando el formato general `{ts 'aaaa-mm-dd hh:mm:ss'}`.
- El valor completo de fecha y/o hora debe ir entre comillas simples. Los años se deben expresar en formato de cuatro dígitos y las fechas y horas deben contener dos dígitos para cada parte del valor. Por ejemplo, el 1 de enero de 2005, 1:05 AM se expresaría como:

```
{ts '2005-01-01 01:05:00'}
```

**Funciones.** SPSS ofrece una selección de funciones preincorporadas SQL aritméticas, lógicas, de cadena, de fecha y de hora. Puede arrastrar una función de la lista hasta la expresión, o introducir una función SQL válida. Consulte la documentación de la base de datos para obtener funciones SQL válidas. Hay una lista de funciones estándar disponibles en:

<http://msdn2.microsoft.com/en-us/library/ms711813.aspx>

**Utilizar muestreo aleatorio.** Esta opción selecciona una muestra aleatoria de casos del origen de datos. Para grandes orígenes de datos, es posible que desee limitar el número de casos a una pequeña y representativa muestra, lo que reduce considerablemente el tiempo de ejecución de procesos. Si el muestreo aleatorio original se encuentra disponible para el origen de datos, resulta más rápido que el muestreo aleatorio de SPSS dado que este último aún debe leer todo el origen de datos para extraer una muestra aleatoria.

- **Aproximadamente.** Genera una muestra aleatoria con el porcentaje aproximado de casos indicado. Dado que esta rutina toma una decisión pseudo-aleatoria para cada caso, el porcentaje de casos seleccionados sólo se puede aproximar al especificado. Cuantos más casos contenga el archivo de datos, más se acercará el porcentaje de casos seleccionados al porcentaje especificado.
- **Exactamente.** Selecciona una muestra aleatoria con el número de casos especificado a partir del número total de casos especificado. Si el número total de casos especificado excede el número total de casos presentes en el archivo de datos, la muestra contendrá un número menor de casos proporcional al número solicitado.

*Nota:* Si utiliza el muestreo aleatorio, agregación (disponible en el modo distribuido con el servidor de SPSS) no estará disponible.

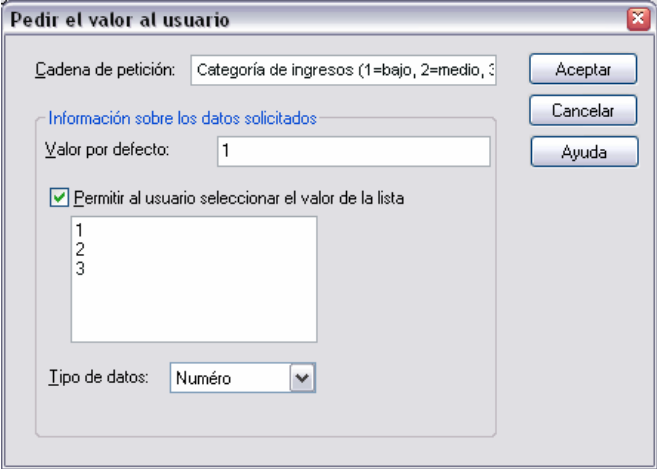
**Pedir el valor al usuario.** Permite insertar una solicitud en la consulta para crear una **consulta con parámetros**. Cuando un usuario ejecute la consulta, se le solicitará que introduzca los datos (según lo que se haya especificado aquí). Puede interesarle esta opción si necesita obtener diferentes vistas de los mismos datos. Por ejemplo, es posible que desee ejecutar la misma consulta para ver las cifras de ventas de diversos trimestres fiscales.

- Sitúe el cursor en cualquier casilla de expresión y pulse en Pedir el valor al usuario para crear una petición.

### **Creación de una consulta con parámetros**

Utilice el paso Pedir el valor al usuario para crear un cuadro de diálogo que solicite información al usuario cada vez que ejecute su consulta. Esta característica resulta útil para realizar consultas de un mismo origen de datos empleando criterios diferentes.

Figura 3-7  
*Pedir el valor al usuario*



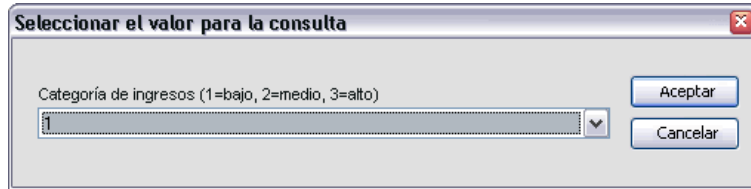
Para crear una solicitud, introduzca una cadena de petición y un valor por defecto. Esta cadena aparecerá cada vez que un usuario ejecute la consulta. La cadena especificará el tipo de información que debe introducir. Si la información no se ofrece en una lista, la cadena sugerirá el formato que debe aplicarse a la información. A continuación se muestra un ejemplo: Introduzca un trimestre (Q1, Q2, Q3, ...).

**Permitir al usuario seleccionar el valor de la lista.** Si selecciona esta casilla de verificación, puede limitar las elecciones del usuario a los valores que incluya en esta lista. Asegúrese de que los valores se separan por retornos de carro.

**Tipo de datos.** Seleccione aquí el tipo de datos (Número, Cadena o Fecha).

El resultado final presenta este aspecto:

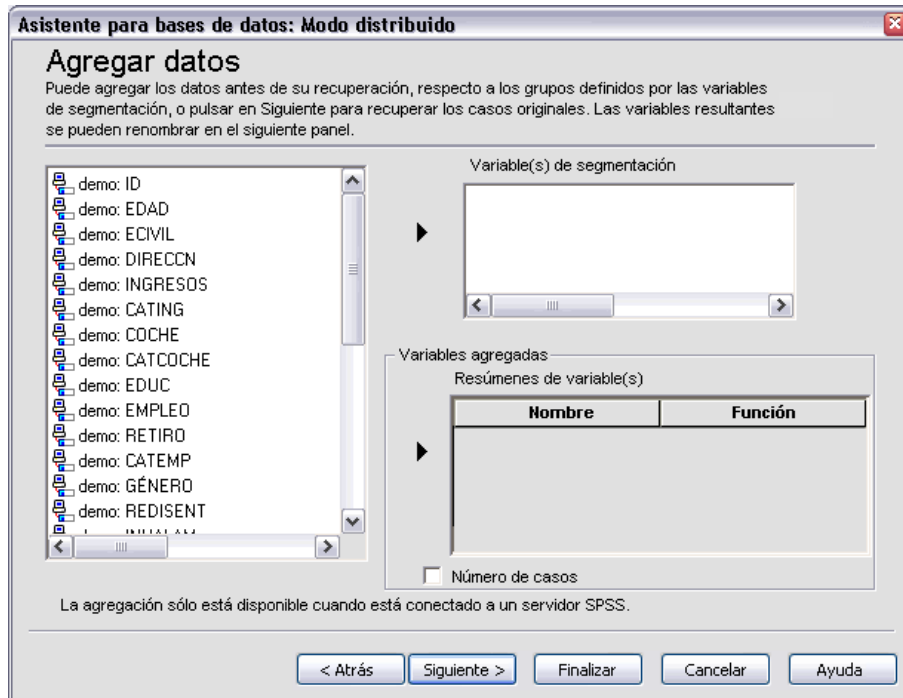
Figura 3-8  
Petición de datos definida por el usuario



### Adición de Datos

Si se encuentra en modo distribuido, conectado a un servidor remoto (disponible con el servidor SPSS), podrá agregar los datos antes de leerlos en SPSS .

Figura 3-9  
Asistente para bases de datos, agregación de datos



También se pueden agregar los datos después de leerlos en SPSS , pero si lo hace antes ahorrará tiempo en el caso de grandes orígenes de datos.

- ▶ Para crear datos agregados, seleccione una o más variables de segmentación que definan cómo deben agruparse los casos.
- ▶ Seleccione una o varias variables agregadas.
- ▶ Seleccione una función de agregación para cada variable agregada.
- ▶ Si lo desea, cree una variable que contenga el número de casos en cada grupo de segmentación.

---

*Nota:* Si utiliza el muestreo aleatorio de SPSS , la agregación no estará disponible.

### **Definición de variables**

**Nombres y etiquetas de variables.** El nombre completo del campo (columna) de la base de datos se utiliza como etiqueta de la variable. A menos que modifique el nombre de la variable, el Asistente para bases de datos asignará nombres de variable a cada columna de la base de datos de una de las siguientes formas:

- Si el nombre del campo de la base de datos forma un nombre de variable válido y único, se usará como el nombre de la variable.
- Si el nombre del campo de la base de datos no es un nombre de variable válido y único, se generará automáticamente un nombre único.

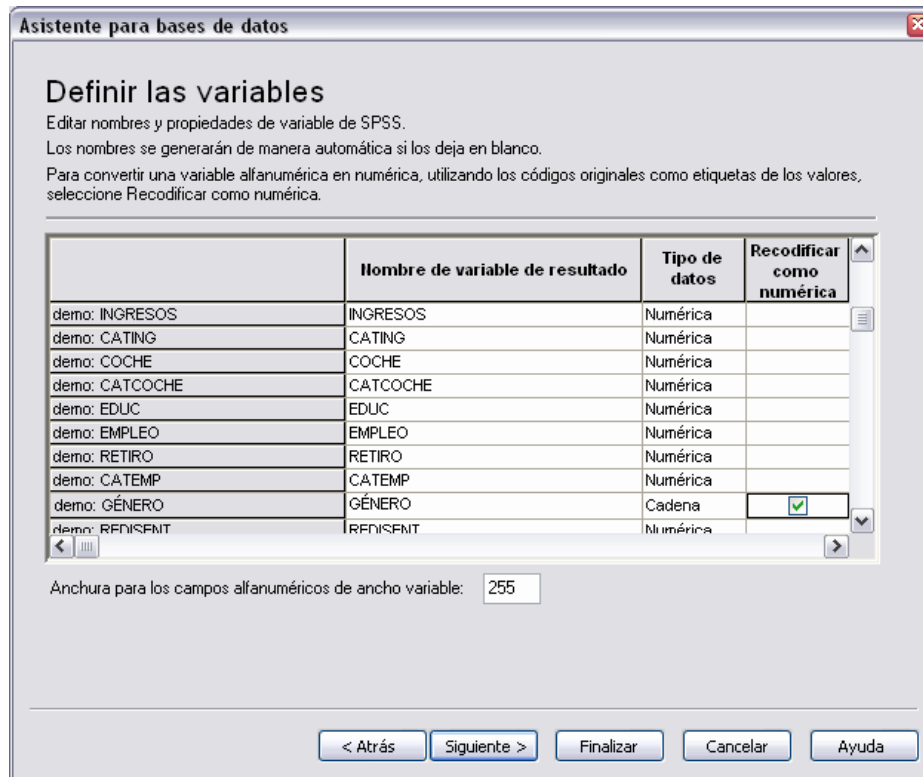
Pulse en cualquier casilla para editar el nombre de la variable.

**Conversión de cadenas en variables numéricas.** Seleccione la casilla Recodificar como numérica para convertir automáticamente una variable de cadena en una variable numérica. Los valores de cadena se convierten en valores enteros consecutivos en función del orden alfabético de los valores originales. Los valores originales se mantienen como etiquetas de valor para las nuevas variables.

**Anchura para los campos de ancho variable.** Esta opción controla la anchura de los valores de las cadenas de anchura variable. Por defecto, la anchura es de 255 bytes y sólo se leen los primeros 255 bytes (generalmente 255 caracteres en idiomas de un solo byte). El valor máximo que se puede asignar a este parámetro es de 32.767 bytes. Aunque posiblemente no desee truncar los valores de cadena, tampoco deseará especificar un valor innecesariamente alto, ya que produciría que el procesamiento fuera ineficaz.

**Minimizar las longitudes de cadena en función de los valores observados.** Establece automáticamente el ancho de cada variable de cadena al valor observado más largo.

Figura 3-10  
Asistente para bases de datos, definición de variables



### Ordenación de casos

Si se encuentra en modo distribuido, conectado a un servidor remoto (disponible con el servidor SPSS), podrá ordenar los datos antes de leerlos en SPSS .

Figura 3-11  
Asistente para bases de datos, ordenación de casos



También se pueden ordenar los datos después de leerlos en SPSS , pero si lo hace antes ahorrará tiempo en el caso de grandes orígenes de datos.

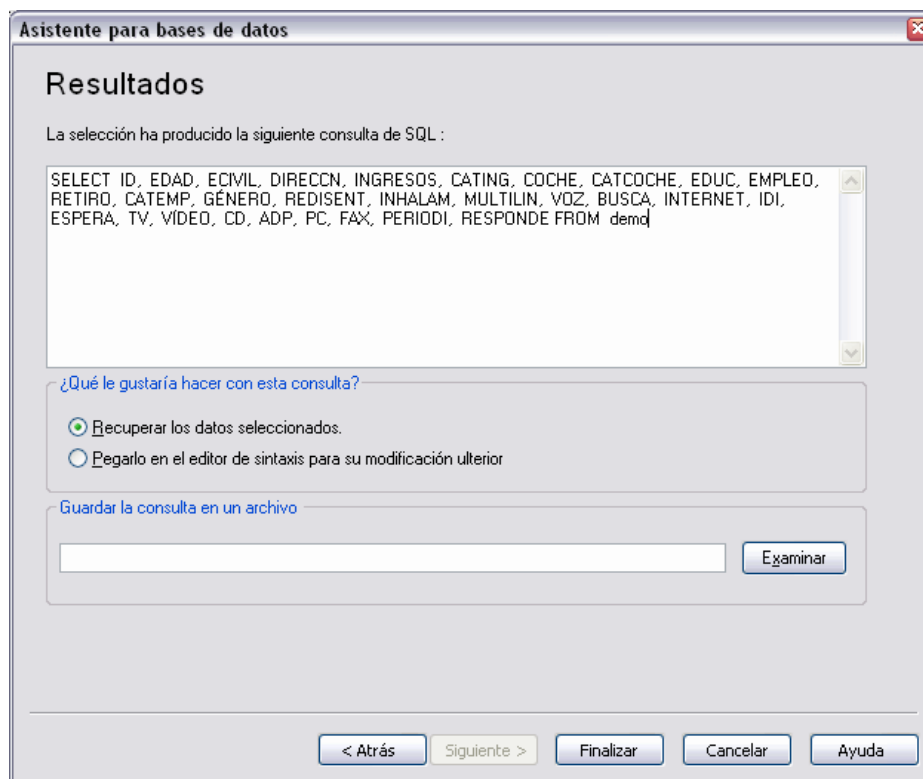
## Resultados

El paso Resultados muestra la instrucción Select de SQL para la consulta.

- Se puede editar la instrucción Select de SQL antes de ejecutar la consulta, pero si pulsa el botón Anterior para introducir cambios en pasos anteriores, se perderán los cambios realizados en la instrucción Select.
- Para guardar la consulta para utilizarla más adelante, utilice la sección Guardar la consulta en un archivo.
- Para pegar la sintaxis GET DATA completa en una ventana de sintaxis, seleccione Pegarlo en el editor de sintaxis para su modificación ulterior. Copiar y pegar la instrucción Select de la ventana Resultados no pegará la sintaxis de comandos necesaria.

*Nota:* La sintaxis pegada contiene un espacio en blanco delante de las comillas de cierre en cada línea de SQL generada por el asistente. Estos espacios no son superfluos. Cuando se procesa el comando, todas las líneas de la instrucción SQL se fusionan de un modo muy literal. Si esos espacios, los caracteres último y primero de cada línea se unirían.

Figura 3-12  
Asistente para bases de datos, panel de resultados



### **Asistente para la importación de texto**

El Asistente para la importación de texto puede leer archivos de datos de texto de diversos formatos:

- Archivos delimitados por tabuladores
- Archivos delimitados por espacios
- Archivos delimitados por comas
- Archivos con formato de campos fijos

En los archivos delimitados, también se pueden especificar otros caracteres como delimitadores entre valores, o bien especificar varios delimitadores diferentes.

#### **Para leer archivos de datos de texto**

- ▶ Elija en los menús:
  - Archivo
  - Leer datos de texto...
- ▶ Seleccione el archivo de texto en el cuadro de diálogo Abrir datos.

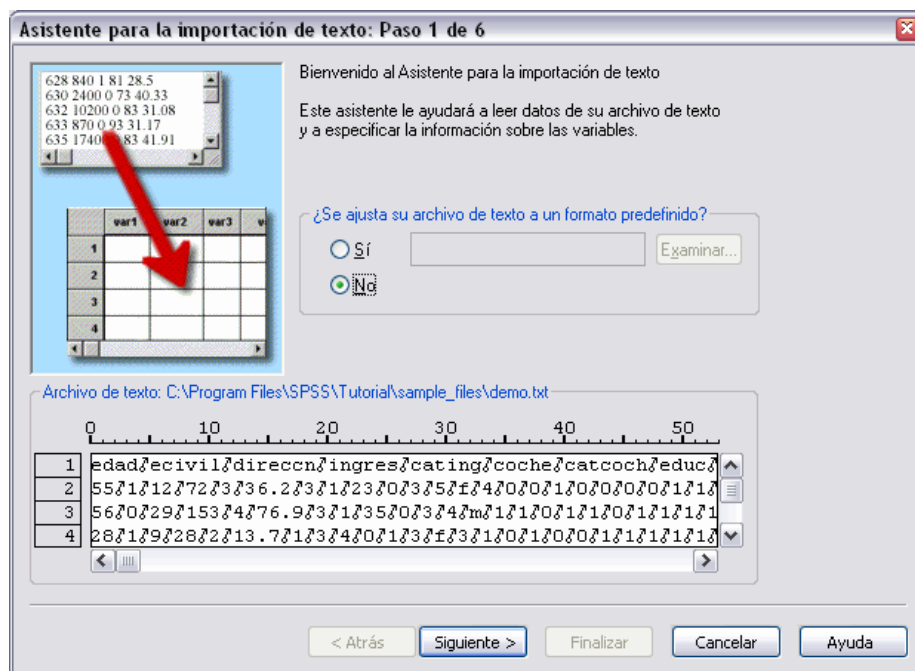


- Siga los pasos indicados en el Asistente para la importación de texto para definir cómo desea leer el archivo de datos de texto.

### Asistente para la importación de texto: Paso 1

Figura 3-13

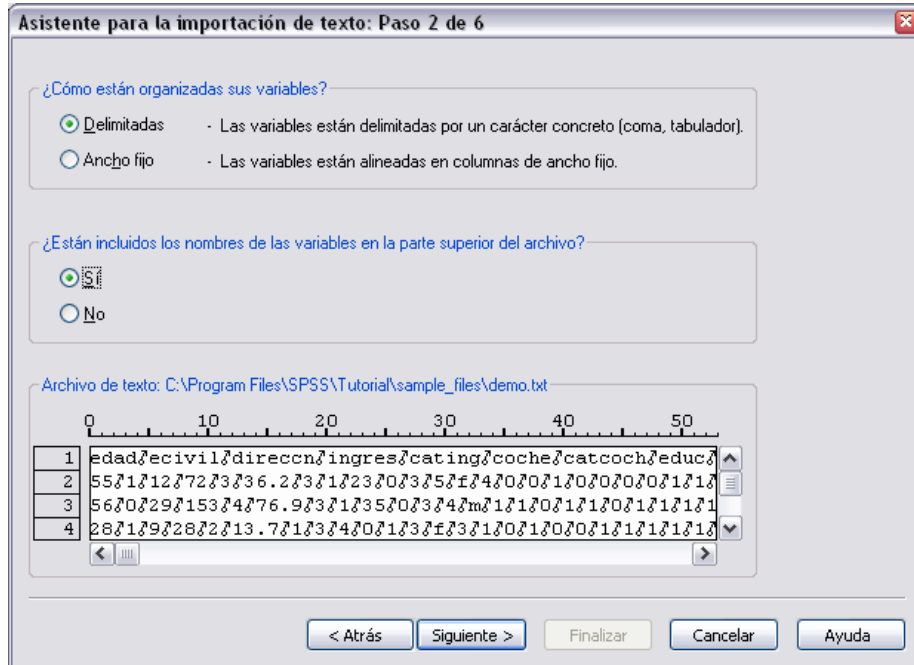
Asistente para la importación de texto: Paso 1



El archivo de texto se mostrará en una ventana de vista previa. Puede aplicar un formato predefinido (guardado con anterioridad desde el Asistente para la importación de texto) o seguir los pasos del asistente para especificar cómo desea que se lean los datos.

## Asistente para la importación de texto: Paso 2

Figura 3-14  
Asistente para la importación de texto: Paso 2



Este paso ofrece información sobre las variables. Una de las variables es similar a uno de los campos de la base de datos. Por ejemplo, cada elemento de un cuestionario es una variable.

**¿Cómo están organizadas sus variables?** Para leer los datos adecuadamente, el Asistente para la importación de texto necesita saber cómo determinar el lugar en el que terminan los valores de datos de una variable y comienzan los valores de datos de la variable siguiente. La organización de las variables define el método utilizado para diferenciar una variable de la siguiente.

- **Delimitado.** Se utilizan espacios, comas, tabulaciones u otros caracteres para separar variables. Las variables quedan registradas en el mismo orden para cada caso, pero no necesariamente conservando la misma ubicación para las columnas.
- **Ancho fijo.** Cada variable se registra en la misma posición de columna en el mismo registro (línea) para cada caso del archivo de datos. No se requiere delimitador entre variables. De hecho, en muchos archivos de datos de texto generados por programas de ordenador, podría parecer que los valores de los datos se suceden, sin espacios que los separen. La ubicación de la columna determina qué variable se está leyendo.

**¿Están incluidos los nombres de las variables en la parte superior del archivo?** Si la primera fila del archivo de datos contiene etiquetas descriptivas para cada variable, podrá utilizar dichas etiquetas como nombres de las variables. Los valores que no cumplan las normas de denominación de variables se convertirán en nombres de variables válidos.

### Asistente para la importación de texto: Paso 3 (Archivos delimitados)

Figura 3-15

Asistente para la importación de texto: Paso 3 (para archivos delimitados)

Este paso ofrece información sobre los casos. Un caso es similar a un registro de una base de datos. Por ejemplo, cada persona que responde a un cuestionario es un caso.

**¿En qué número de línea comienza el primer caso de datos?** Indica la primera línea del archivo de datos que contiene valores de datos. Si la línea o líneas superiores del archivo de datos contienen etiquetas descriptivas o cualquier otro texto que no represente valores de datos, dicha línea o líneas *no* serán la línea 1.

**¿Cómo se representan sus casos?** Controla la manera en que el Asistente para la importación de texto determina dónde finaliza cada caso y comienza el siguiente.

- **Cada línea representa un caso.** Cada línea contiene un sólo caso. Es bastante común que cada línea (fila) contenga un sólo caso, aunque dicha línea puede ser muy larga para un archivo de datos con un gran número de variables. Si no todas las líneas contienen el mismo número de valores de datos, el número de variables para cada caso quedará determinado por la línea que tenga el mayor número de valores de datos. A los casos con menos valores de datos se les asignarán valores perdidos para las variables adicionales.
- **Un número concreto de variables representa un caso.** El número de variables especificado para cada caso informa al Asistente para la importación de texto de dónde detener la lectura de un caso y comenzar la del siguiente. Una misma línea puede contener varios casos y los casos pueden empezar en medio de una línea y continuar en la línea siguiente. El Asistente para la importación de texto determina el final de cada caso basándose en el número de valores leídos, independientemente del número de líneas. Cada caso debe contener valores de datos (o valores perdidos indicados por delimitadores) para todas las variables; de otra forma, el archivo de datos no se leerá correctamente.

**¿Cuántos casos desea importar?** Puede importar todos los casos del archivo de datos, los primeros  $n$  casos (siendo  $n$  un número especificado por el usuario) o una muestra aleatoria a partir de un porcentaje especificado. Dado que esta rutina de muestreo aleatorio toma una decisión pseudo-aleatoria para cada caso, el porcentaje de casos seleccionados sólo se puede aproximar al porcentaje especificado. Cuantos más casos contenga el archivo de datos, más se acercará el porcentaje de casos seleccionados al porcentaje especificado.

### Asistente para la importación de texto: Paso 3 (Archivos de ancho fijo)

Figura 3-16

Asistente para la importación de texto: Paso 3 (para archivos de ancho fijo)

Asistente para la importación de texto: Ancho fijo: Paso 3 de 6

¿En qué número de línea comienza el primer caso de los datos? 1

¿Cuántas líneas representan un caso? 1

¿Cuántos casos desea importar?

Todos los casos

Los primeros 1000 casos.

Un porcentaje de los casos: 10 %

Vista previa de datos

	edad	ecivil	direccn	ingres	cating	coche	d
1							
2	55	1	12	72	3	36.2	3
3	56	0	29	153	4	76.9	3
4	28	1	9	28	2	13.7	1

< Atrás Siguiente > Finalizar Cancelar Ayuda

Este paso ofrece información sobre los casos. Un caso es similar a un registro de una base de datos. Por ejemplo, cada encuestado es un caso.

**¿En qué número de línea comienza el primer caso de datos?** Indica la primera línea del archivo de datos que contiene valores de datos. Si la línea o líneas superiores del archivo de datos contienen etiquetas descriptivas o cualquier otro texto que no represente valores de datos, dicha línea o líneas *no* serán la línea 1.

**¿Cuántas líneas representan un caso?** Controla la manera en que el Asistente para la importación de texto determina dónde finaliza cada caso y comienza el siguiente. Cada variable queda definida por su número de línea dentro del caso y por la ubicación de su columna. Para leer los datos correctamente, deberá especificar el número de líneas de cada caso.

**¿Cuántos casos desea importar?** Puede importar todos los casos del archivo de datos, los primeros  $n$  casos (siendo  $n$  un número especificado por el usuario) o una muestra aleatoria a partir de un porcentaje especificado. Dado que esta rutina de muestreo aleatorio toma una decisión pseudo-aleatoria para cada caso, el porcentaje de casos seleccionados sólo se puede aproximar

al porcentaje especificado. Cuantos más casos contenga el archivo de datos, más se acercará el porcentaje de casos seleccionados al porcentaje especificado.

### Asistente para la importación de texto: Paso 4 (Archivos delimitados)

Figura 3-17

Asistente para la importación de texto: Paso 4 (para archivos delimitados)

Asistente para la importación de texto delimitado - Paso 4 de 6

¿Qué delimitador se encuentra entre las variables?

Tabulador       Espacio

Coma       Punto y coma

Otro:

¿Cuál es el calificador de texto?

Ninguno

Comillas simples

Comillas dobles

Otro:

Vista previa de datos

edad	ecivil	direccn	ingres	cating	coche	cat
55	1	12	72	3	36.2	3
56	0	29	153	4	76.9	3
28	1	9	28	2	13.7	1
24	1	4	26	2	12.5	1

< Atrás      Siguiente >      Finalizar      Cancelar      Ayuda

Este paso muestra la mejor opción, según el Asistente para la importación de texto, para leer el archivo de datos y le permite modificar la manera en que el asistente leerá las variables del archivo de datos.

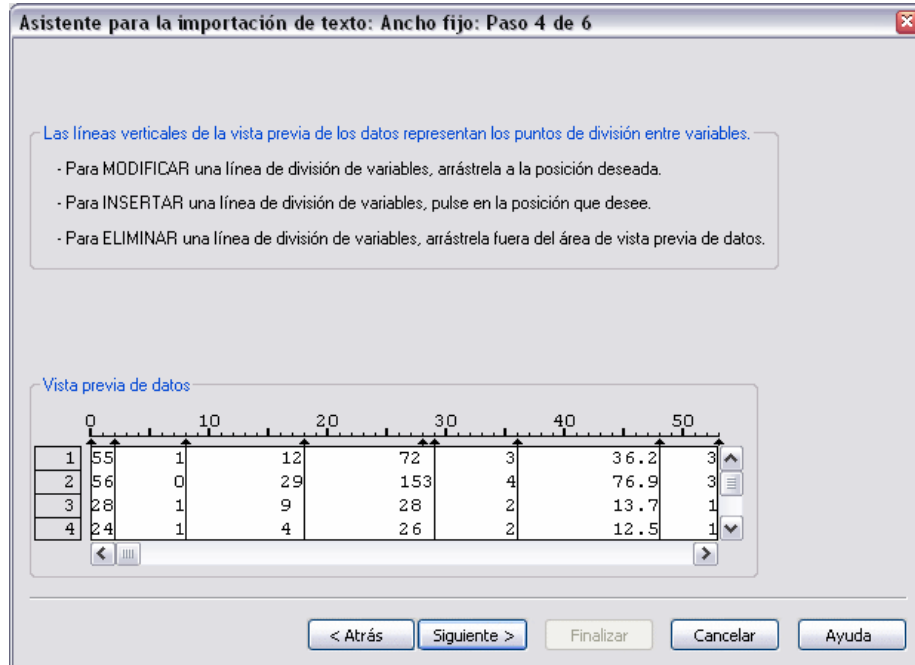
**¿Qué delimitador desea para la separación entre variables?** Indica los caracteres o símbolos que separan los valores de datos. Puede seleccionar cualquier combinación de espacios, comas, signos de punto y coma, tabulaciones o cualquier otro carácter. En caso de existir varios delimitadores consecutivos sin valores de datos, dichos delimitadores serán considerados valores perdidos.

**¿Cuál es el calificador de texto?** Caracteres utilizados para encerrar valores que contienen caracteres delimitadores. Por ejemplo, si una coma es el delimitador, los valores que contengan comas se leerán incorrectamente a menos que estos valores se encierre en un calificador de texto, impidiendo que las comas del valor se interpreten como delimitadores entre los valores. Los archivos de datos con formato CSV de Excel utilizan las comillas dobles (") como calificador de texto. El calificador de texto aparece tanto al comienzo como al final del valor, encerrándolo completamente.

### Asistente para la importación de texto: Paso 4 (Archivos de ancho fijo)

Figura 3-18

Asistente para la importación de texto: Paso 4 (para archivos de ancho fijo)



Este paso muestra la mejor opción, según el Asistente para la importación de texto, para leer el archivo de datos y le permite modificar la manera en que el asistente leerá las variables del archivo de datos. Las líneas verticales de la ventana de vista previa indican el lugar en el que en ese momento el Asistente para la importación de texto piensa que cada variable comienza en el archivo.

Inserte, mueva y elimine líneas de ruptura de variable según convenga para separar variables. Si se utilizan varias líneas para cada caso, los datos aparecerán como una línea para cada caso y las líneas posteriores se adjuntarán al final de la línea.

#### Notas:

En archivos de datos generados por ordenador que producen un flujo continuo de valores de datos sin espacios ni otras características distintivas, puede resultar difícil determinar el lugar en el que comienza cada variable. Los archivos de datos del tipo citado anteriormente suelen depender de un archivo de definición de datos u otro tipo de descripción escrita que especifique la ubicación por líneas y columnas de cada variable.

### Asistente para la importación de texto: Paso 5

Figura 3-19  
Asistente para la importación de texto: Paso 5

Asistente para la importación de texto: Paso 5 de 6

Especificaciones para la(s) variable(s) seleccionada(s) en la vista previa de datos

Nombre de variable:       Nombre original:

Formato de datos:

Vista previa de datos

edad	ecivil	direccn	ingres	cating	coche	cat
55	1	12	72	3	36.2	3
56	0	29	153	4	76.9	3
28	1	9	28	2	13.7	1
24	1	4	26	2	12.5	1

< Atrás    Siguiente >    Finalizar    Cancelar    Ayuda

Este paso controla el nombre de la variable y el formato de datos que el Asistente para la importación de texto utilizará para leer cada variable, así como las que se incluirán en el archivo de datos definitivo.

**Nombre de variable.** Puede sobrescribir los nombres de variable por defecto y sustituirlos por otros diferentes. Si lee nombres de variable desde el archivo de datos, el Asistente para la importación de texto modificará de manera automática los nombres de variable que no cumplan las normas de denominación de variables. Seleccione una variable en la ventana de vista previa e introduzca un nombre de variable.

**Formato de datos.** Seleccione una variable en la ventana de vista previa y, a continuación, seleccione un formato de la lista desplegable. Pulse el botón del ratón con la tecla MAYÚS presionada para seleccionar una serie de variables consecutivas, o bien con la tecla CTRL presionada para seleccionar una serie de variables no consecutivas.

### Opciones para el formato del Asistente para la importación de texto

Entre las opciones de formato para la lectura de variables con el Asistente para la importación de texto se encuentran:

**No importar.** Omite la variable o variables seleccionadas del archivo de datos importado.

**Numérico.** Los valores válidos incluyen números, los signos más y menos iniciales y un indicador decimal.

**Cadena.** Son valores válidos prácticamente todos los caracteres del teclado y los espacios en blanco incrustados. En los archivos delimitados, puede especificar hasta un máximo de 32.767 de caracteres para el valor. El Asistente para la importación de texto fija como valor por defecto para el número de caracteres el valor de cadena más largo que se haya encontrado para la variable o variables seleccionadas. Para los archivos de ancho fijo, el número de caracteres en los valores de cadena queda definido por la ubicación de las líneas de ruptura de variable en el paso 4.

**Fecha/hora.** Entre los valores válidos se encuentran las fechas con formato general: *dd-mm-aaaa*, *mm/dd/aaaa*, *dd.mm.aaaa*, *aaaa/mm/dd*, *hh:mm:ss*, así como una amplia variedad de formatos de hora y fecha. Los meses se pueden representar en dígitos, números romanos, abreviaturas de tres letras o con el nombre completo. Seleccione un formato de fecha de la lista.

**Dólar.** Los valores válidos son números con un signo dólar inicial optativo y puntos separadores de millares también optativos.

**Coma.** Entre los valores válidos se encuentran los números que utilizan un punto para separar los decimales y una coma para separar los millares.

**Punto.** Entre los valores válidos se encuentran los números que utilizan una coma para separar los decimales y un punto para separar los millares.

*Nota:* Los valores que contengan caracteres no válidos para el formato seleccionado serán considerados valores perdidos. Los valores que contengan uno cualquiera de los delimitadores especificados serán considerados como valores múltiples.

### Asistente para la importación de texto: Paso 6

Figura 3-20  
Asistente para la importación de texto: Paso 6

Asistente para la importación de texto: Paso 6 de 6

Ha definido satisfactoriamente el formato del archivo de texto.

¿Desea guardar este formato de archivo para su uso en el futuro?

Sí  No

¿Desea pegar la sintaxis?

Sí  No  Hacer caché de datos en equipo local

Pulse en el botón Finalizar para completar el Asistente para la importación de texto.

Vista previa de datos

edad	ecivil	direccn	ingres	cating	coche	catcoch	educ
55	1	12	72	3	36.2	3	1
56	0	29	153	4	76.9	3	1
28	1	9	28	2	13.7	1	3
24	1	4	26	2	12.5	1	4

< Atrás   Siguiente >   Finalizar   Cancelar   Ayuda



Este es el paso final del Asistente para la importación de texto. Puede guardar sus propias especificaciones en un archivo para hacer uso de ellas cuando importe archivos de datos de texto similares. También puede pegar la sintaxis generada por el Asistente para la importación de texto en una ventana de sintaxis. Así podrá personalizar y/o guardar dicha sintaxis para utilizarla en futuras sesiones o en trabajos de producción.

**Caché local de los datos.** Una caché de los datos es una copia completa del archivo de datos, almacenada en un espacio temporal del disco. La caché del archivo de datos puede mejorar el rendimiento.

## ***Lectura de datos de Dimensions***

En los sistemas operativos Microsoft Windows, puede leer datos de productos de SPSS Dimensions, incluyendo Quanvert, Quancept y mrInterview. (*Nota:* esta función sólo está disponible si se ha instalado SPSS en un sistema operativo Microsoft Windows.

Para leer los orígenes de datos de Dimensions, debe tener instalados los siguientes elementos:

- .NET framework
- Dimensions Data Model y OLE DB Access

Las versiones de estos componentes compatibles con esta versión se pueden instalar desde el CD de instalación y están disponibles en el menú de reproducción automática. Sólo puede añadir orígenes de datos de Dimensions en el análisis en modo local. Esta característica no está disponible en el análisis en modo distribuido con el servidor de SPSS.

Para leer datos de un origen de datos de Dimensions:

- ▶ En cualquiera de las ventanas de SPSS abiertas, elija en los menús:  
Archivo  
Abrir datos de Dimensions
- ▶ En la pestaña Propiedades de vínculo de datos: Conexión, especifique el archivo de metadatos, el tipo de datos de casos y el archivo de datos de casos.
- ▶ Pulse en Aceptar.
- ▶ En el cuadro de diálogo Importación de datos de Dimensions, seleccione las variables que desea incluir y seleccione cualquier criterio de selección de casos.
- ▶ Pulse en Aceptar para leer los datos.

### ***Pestaña Propiedades de vínculo de datos: Conexión***

Para leer un origen de datos de Dimensions, debe especificar:

**Ubicación de metadatos.** El archivo del documento de metadatos (*.mdd*) que contiene la información de definición del cuestionario.

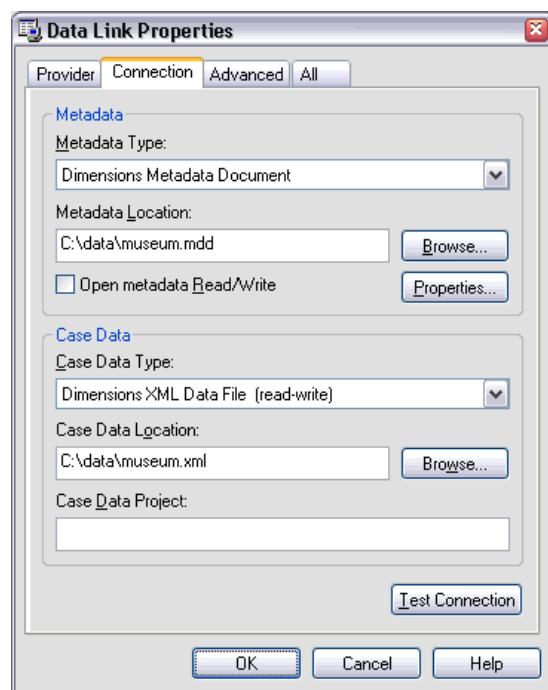
**Tipo de datos de casos.** El formato del archivo de datos de casos. Los formatos disponibles incluyen:

- Archivo de datos de Quancept (DRS). Datos del caso en un archivo Quancept *.drs*, *.drz* o *.dru*.
- Base de datos de Quanvert. Datos del caso en una base de datos de Quanvert.
- Base de datos de Dimensions (MS SQL Server). Datos de casos en una base de datos relacional de investigación de mercado de SPSS en SQL Server. Esta opción se puede utilizar para leer los datos recopilados mediante mrInterview.
- Archivo de datos XML de Dimensions. Datos de casos en un archivo XML.

**Ubicación de datos de casos.** El archivo que contiene los datos de casos. El formato de este archivo debe ser coherente con el tipo de datos de casos seleccionado.

Figura 3-21

Propiedades de vínculo de datos: Pestaña Conexión



*Nota:* no se sabe hasta qué punto el resto de opciones de la pestaña Conexión o cualquier opción del resto de pestañas Propiedades de vínculo de datos pueden afectar a la lectura de datos de Dimensions en SPSS, por lo que se recomienda no cambiar ninguna de ellas.

### **Pestaña Seleccionar variables**

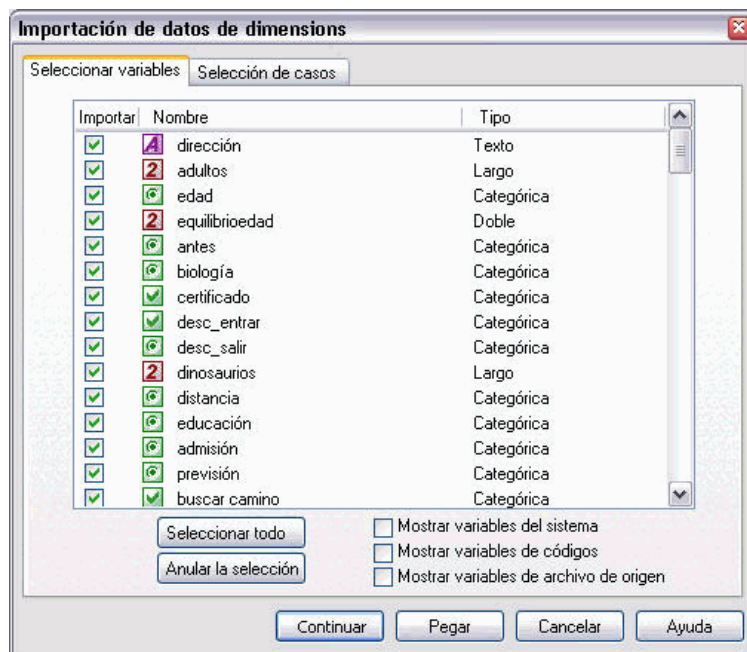
Puede seleccionar un subconjunto de variables que se van a leer. Por defecto, todas las variables típicas del origen de datos se muestran y aparecen seleccionadas.

- **Mostrar variables del sistema.** Muestra cualquier variable “de sistema”, incluidas las variables que indican el estado de encuesta (*en curso*, *finalizada*, *fecha de finalización*, etc.). A continuación, puede seleccionar cualquier variable de sistema que desee incluir. Por defecto, se excluyen todas las variables del sistema.

- **Mostrar variables de códigos.** Muestra cualquier variable que represente códigos que se utilizan para respuestas “Otros” abiertas para variables categóricas. A continuación, puede seleccionar cualquier variable de códigos que desee incluir. Por defecto, se excluyen todas las variables de códigos.
- **Mostrar variables de archivo de origen.** Muestra cualquier variable que contenga nombres de archivo de imágenes de respuestas exploradas. A continuación, puede seleccionar cualquier variable de archivo de origen que desee incluir. Por defecto, todas las variables de archivo de origen están excluidas.

Figura 3-22

Importación de datos de Dimensions: pestaña Seleccionar variables



### ***Pestaña Selección de casos***

Para los orígenes de datos de Dimensions que contienen variables del sistema, puede seleccionar casos basados en algunos criterios de variable de sistema. No es necesario incluir las variables del sistema correspondientes en la lista de variables que se van a leer, pero las variables del sistema necesarias deben existir en los datos de origen para aplicar los criterios de selección. Si las variables del sistema necesarias no existen en los datos de origen, se ignorarán los criterios de selección correspondientes.

**Estado de recopilación de datos.** Puede seleccionar datos de encuestados, datos de prueba o ambos. También puede seleccionar casos basados en cualquier combinación de los siguientes parámetros de estado de encuesta:

- Finalizada correctamente
- Activa / en curso
- Tiempo agotado

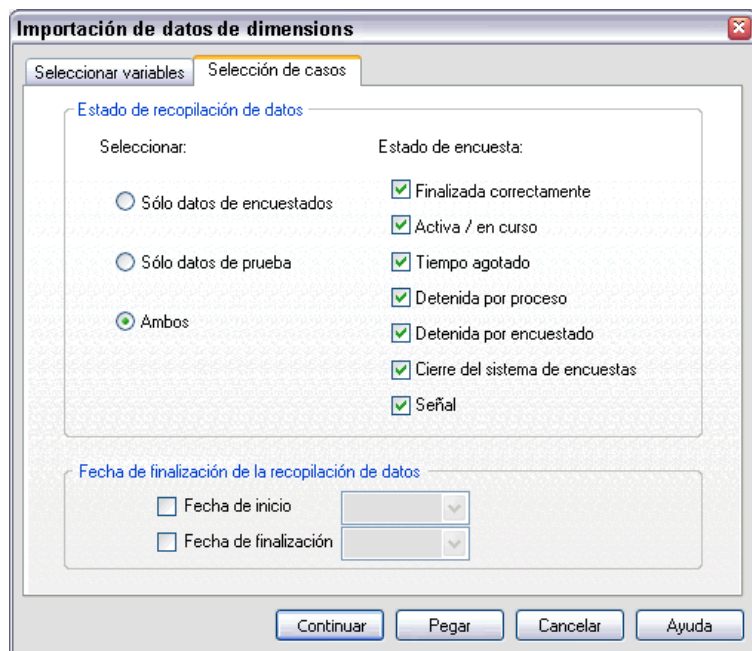
- Detenida por proceso
- Detenida por encuestado
- Cierre del sistema de encuestas
- Señal (el proceso se completa con una instrucción señalizadora)

**Fecha de finalización de la recopilación de datos.** Puede seleccionar casos basados en la fecha de finalización de la recopilación de datos.

- **Fecha de inicio.** Se incluyen los casos para los que se completó la recopilación de datos durante o después de la fecha especificada.
- **Fecha de finalización.** Se incluyen los casos para los que se completó la recopilación de datos antes de la fecha especificada. *No* se incluyen los casos para los que la recopilación de datos se completó en la fecha de finalización.
- Si especifica tanto una fecha de inicio como una fecha de finalización, se definirá un rango de fechas de finalización desde la fecha de inicio hasta la fecha de finalización (ésta última no incluida).

Figura 3-23

Importación de datos de Dimensions: pestaña Selección de casos



## Información sobre el archivo

Un archivo de datos contiene mucho más que datos sin formato. También contiene información sobre la definición de las variables, incluyendo:

- Los nombres de las variables
- Los formatos de las variables
- Las etiquetas descriptivas de variable y de valor

Esta información se almacena en la parte del diccionario sobre el archivo de datos. El Editor de datos proporciona una forma de presentar la información sobre la definición de la variable. También se puede mostrar la información completa del diccionario para el conjunto de datos activo o para cualquier otro archivo de datos.

#### ***Para mostrar información sobre los archivos de datos***

- ▶ Seleccione en los menús de la ventana Editor de datos:
    - Archivo
    - Mostrar información del archivo de datos
  - ▶ Para el archivo de datos abierto actualmente, elija Archivo de trabajo.
  - ▶ Para otros archivos de datos, elija Archivo externo y seleccione el archivo de datos.
- La información sobre el archivo de datos se muestra en el Visor.

## ***Almacenamiento de archivos de datos***

Además de guardar los archivos de datos en formato de SPSS, también puede guardarlos en una amplia variedad de formatos externos, entre ellos:

- Excel y otros formatos de hoja de cálculo
- Archivos de texto delimitado por tabuladores y CSV
- SAS
- Stata
- Tablas de base de datos

## ***Almacenamiento de archivos de datos modificados***

- ▶ Active la ventana Editor de datos (pulse en cualquier punto de la ventana para activarla).
- ▶ Elija en los menús:
  - Archivo
  - Guardar

El archivo de datos modificado se guarda y sobrescribe la versión anterior del archivo.

Nota: los archivos de datos guardados en modo Unicode no se pueden leer en versiones de SPSS anteriores a la 16.0. Para guardar un archivo de datos Unicode en un formato que pueda ser leído por versiones anteriores, abra el archivo en el modo de página de código y vuelva a guardarlo. El archivo se guardará en la codificación basada en la configuración regional actual. Es posible que se pierdan algunos datos si el archivo contiene caracteres que no reconoce la configuración regional actual. Si desea obtener información acerca de cómo cambiar de modo Unicode a modo de página de código, consulte [General: Opciones](#) en p. 479.

## ***Almacenamiento de archivos de datos en formatos externos***

- ▶ Active la ventana Editor de datos (pulse en cualquier punto de la ventana para activarla).

- ▶ Elija en los menús:
  - Archivo
  - Guardar como...
- ▶ Seleccione un tipo de archivo de la lista desplegable.
- ▶ Introduzca un nombre de archivo para el nuevo archivo de datos.

***Para escribir nombres de variable en la primera fila de una hoja de cálculo o de un archivo de datos delimitados por tabuladores:***

- ▶ Pulse en Escribir nombres de variable en hoja de cálculo en el cuadro de diálogo Guardar datos como.

***Para guardar las etiquetas de valor en lugar de los valores de los datos en archivos Excel:***

- ▶ Pulse en Guardar etiquetas de valor donde se hayan definido en vez de valores de datos en el cuadro de diálogo Guardar datos como.

***Para guardar etiquetas de valor en un archivo de sintaxis de SAS (esta opción sólo está activa si se ha seleccionado un tipo de archivo de SAS):***

- ▶ Pulse en Guardar etiquetas de valor en un archivo .sas en el cuadro de diálogo Guardar datos como.

Si desea obtener información sobre la exportación de datos a tablas de base de datos, consulte [Exportación a base de datos en p. 52](#).

Si desea obtener información sobre la exportación de datos para utilizarlos en aplicaciones de Dimensions, consulte [Exportación a Dimensions en p. 65](#).

### ***Almacenamiento de datos: Tipos de archivos de datos***

Puede guardar datos en los siguientes formatos:

**SPSS (\*.sav).** Formato SPSS.

- Los archivos de datos guardados con formato SPSS no se pueden leer en versiones anteriores a la 7.5. Los archivos de datos guardados en modo Unicode no se pueden leer en versiones de SPSS anteriores a la 16.0. Si desea obtener más información, consulte [General: Opciones](#) en Capítulo 45 en p. 479.
- Al utilizar archivos de datos con nombres de variable con longitud superior a ocho bytes en la versión 10.X u 11.X, se utilizan versiones únicas de ocho bytes de los nombres de variable, pero se mantienen los nombres originales para su utilización en la versión 12.0 o posteriores. En versiones anteriores a la 10.0, los nombres largos originales de las variables se pierden si se guarda el archivo de datos.
- Al utilizar archivos de datos con variables de cadena con más de 255 bytes en versiones anteriores a la versión 13.0, dichas variables de cadena se fragmentan en variables de cadena de 255 bytes.

**Versión 7.0 (\*.sav).** Formato de la versión 7.0. Los archivos de datos guardados con formato de la versión 7.0 se pueden leer en la versión 7.0 y en versiones anteriores, pero no incluyen los conjuntos de respuestas múltiples definidos ni la información sobre la introducción de datos para Windows.

**SPSS/PC+ (\*.sys).** Formato SPSS/PC+. Si el archivo de datos contiene más de 500 variables, sólo se guardarán las 500 primeras. Para las variables con más de un valor definido como perdido por el usuario, los valores definidos como perdidos por el usuario adicionales se recodificarán en el primero de estos valores.

**SPSS portátil (\*.por).** El formato portátil puede leerse en otras versiones de SPSS y en versiones para otros sistemas operativos. Los nombres de variable se limitan a ocho bytes, y se convertirán a nombres únicos de ocho bytes si es preciso. En la mayoría de los casos, ya no es necesario guardar los datos en formato portátil, ya que los archivos de datos en formato SPSS deberían ser independientes de la plataforma y del sistema operativo. No se puede guardar los archivos de datos en un archivo portátil en modo Unicode. Si desea obtener más información, consulte [General: Opciones](#) en Capítulo 45 en p. 479.

**Delimitado por tabulaciones (\*.dat).** Archivos de texto con valores separados por tabuladores. (*Nota:* los tabuladores incrustados en los valores de cadena se conservarán como tabuladores en el archivo delimitado por tabuladores. No se realiza ninguna distinción entre los tabuladores incrustados en los valores y los tabuladores que separan los valores.)

**Delimitado por comas (\*.csv).** Archivos de texto con valores separados por comas o puntos y coma. Si el indicador decimal actual de SPSS es un punto, los valores se separan mediante comas. Si el indicador decimal actual es una coma, los valores se separan mediante punto y coma.

**ASCII en formato fijo (\*.dat).** Archivos de texto con formato fijo, utilizando los formatos de escritura por defecto para todas las variables. No existen tabuladores ni espacios entre los campos de variable.

**Excel 2007 (\*.xlsx).** Libro de trabajo con formato XLSX de Microsoft Excel 2007. El número máximo de variables es 16.000, el resto de variables adicionales por encima de esa cifra se eliminan. Si el conjunto de datos contiene más de un millón de casos, se crean varias hojas en el libro de trabajo.

**Excel 97 hasta 2003 (\*.xls).** Libro de trabajo de Microsoft Excel 97. El número máximo de variables es 256, el resto de variables adicionales por encima de esa cifra se eliminan. Si el conjunto de datos contiene más de 65.356 casos, se crean varias hojas en el libro de trabajo.

**Excel 2.1(\*.xls).** Archivo de hoja de cálculo de Microsoft Excel 2,1. El número máximo de variables es de 256 y el número máximo de filas es de 16.384.

**1-2-3 versión 3.0 (\*.wk3).** Archivo de hoja de cálculo de Lotus 1-2-3, versión 3.0. El número máximo de variables que puede guardar es 256.

**1-2-3 versión 2.0 (\*.wk1).** Archivo de hoja de cálculo de Lotus 1-2-3, versión 2.0. El número máximo de variables que puede guardar es 256.

**1-2-3 versión 1.0 (\*.wks).** Archivo de hoja de cálculo de Lotus 1-2-3, versión 1A. El número máximo de variables que puede guardar es 256.

**SYLK (\*.slk).** Formato de vínculo simbólico para archivos de hojas de cálculo de Microsoft Excel y de Multiplan. El número máximo de variables que puede guardar es 256.

**dBASE IV (\*.dbf).** Formato dBASE IV.

**dBASE III (\*.dbf).** Formato dBASE III.

**dBASE II (\*.dbf).** Formato dBASE II.

**SAS v7+ Windows extensión corta (\*.sd7).** Versiones 7-8 de SAS para Windows con formato de nombre de archivo corto.

**SAS v7+ Windows extensión larga (\*.sas7bdat).** Versiones 7-8 de SAS para Windows con formato de nombre de archivo largo.

**SAS v7+ para UNIX (\*.ssd01).** SAS v8 para UNIX.

**SAS v6 para Windows (\*.sd2).** Formato de archivo de SAS v6 para Windows/OS2.

**SAS v6 para UNIX (\*.ssd01).** Formato de archivo de SAS v6 para UNIX (Sun, HP, IBM).

**SAS v6 para Alpha/OSF (\*.ssd04).** Formato de archivo de SAS v6 para Alpha/OSF (DEC UNIX).

**Transporte de SAS (\*.xpt).** Archivo de transporte de SAS.

**Versión 8 de Stata Intercooled (\*.dta).**

**Versión 8 de Stata SE (\*.dta).**

**Versión 7 de Stata Intercooled (\*.dta).**

**Versión 7 de Stata SE (\*.dta).**

**Versión 6 (\*.dta) de Stata.**

**Versiones 4-5 (\*.dta) de Stata.**

### ***Almacenamiento de archivos: Opciones***

En los archivos de hoja de cálculo, los delimitados por tabuladores y los delimitados por coma, se pueden escribir nombres de variable en la primera fila del archivo.

### ***Almacenamiento de archivos de datos en formato de Excel***

Puede guardar los datos en uno de los tres formatos de archivo de Microsoft Excel. Excel 2.1, Excel 97 y Excel 2007.

- Excel 2.1 y Excel 97 tienen un límite de 256 columnas; por lo tanto, sólo se incluyen las primeras 256 variables.
- Excel 2007 tiene un límite de 16.000 columnas; por lo tanto, sólo se incluyen las primeras 16.000 variables.
- Excel 2,1 tiene un límite de 16.384 filas; por lo tanto, sólo se incluyen los primeros 16.384 casos.
- Excel 97 y Excel 2007 tienen un número limitado de filas por hoja, pero como los libros de trabajo pueden tener múltiples hojas, se crean más cuando se excede el máximo de cada hoja.



### **Tipos de variables**

La siguiente tabla muestra la relación del tipo de las variables entre los datos originales de SPSS y los datos exportados a Excel.

<b>SPSS Tipo de variable</b>	<b>Formato de datos de Excel</b>
Numérico	0.00; #,##0.00; ...
Coma	0.00; #,##0.00; ...
Dólar	\$#,##0_); ...
Fecha	d-mmm-aaaa
Time	hh:mm:ss
Cadena	General

### **Almacenamiento de archivos de datos en formato de SAS**

Al guardar un archivo de SAS, se aplica un tratamiento especial a determinadas características de los datos. Entre estos casos se incluyen:

- Algunos caracteres que se permiten en los nombres de variables de SPSS no son válidos en SAS, como por ejemplo @, # y \$. Al exportar los datos, estos caracteres no válidos se reemplazan por un carácter de subrayado.
- Los nombres de variable de SPSS que contienen caracteres de varios bytes (por ejemplo, caracteres japoneses o chinos) se convierten en nombres de variable con formato general *Vnnn*, donde *nnn* es un valor entero.
- Las etiquetas de variable de SPSS que contienen más de 40 caracteres se truncan al exportarlas a un archivo de SAS v6.
- Si existen, las etiquetas de variable de SPSS se asignan a etiquetas de variable de SAS. Si no hay ninguna etiqueta de variable en los datos de SPSS, el nombre de variable se asigna a la etiqueta de variable de SAS.
- SAS sólo permite que exista un valor perdido del sistema, mientras que SPSS permite que haya varios valores perdidos del sistema. Por tanto, todos los valores perdidos del sistema en SPSS se asignan a un único valor perdido del sistema en el archivo SAS.

### **Almacenamiento de etiquetas de valor**

Existe la posibilidad de guardar los valores y las etiquetas de valor asociadas al archivo de datos en un archivo de sintaxis de SAS. Por ejemplo, al exportar las etiquetas de valor del archivo de datos *coches.sav*, el archivo de sintaxis generado contiene:

```
libname library '\spss\' ;

proc format library = library ;
    value ORIGIN /* País de origen */
        1 = 'EE.UU.'
        2 = 'Europa'
```

```

3 = 'Japón' ;

value CYLINDER /* Número de cilindros */
3 = '3 cilindros'
4 = '4 cilindros'
5 = '5 cilindros'
6 = '6 cilindros'
8 = '8 cilindros' ;

value FILTER__ /* cylrec = 1 | cylrec = 2 (FILTER) */
0 = 'No seleccionado'
1 = 'Seleccionado' ;

proc datasets library = library ;
modify cars;

format ORIGIN ORIGIN.;
format CYLINDER CYLINDER.;
format FILTER__ FILTER__.;

quit;

```

Esta función no se admite para el archivo de transporte de SAS.

### ***Tipos de variables***

La siguiente tabla muestra la relación del tipo de las variables entre los datos originales de SPSS y los datos exportados a SAS.

<b>SPSS Tipo de variable</b>	<b>Tipo de variable de SAS</b>	<b>Formato de datos de SAS</b>
Numérico	Numérico	12
Coma	Numérico	12
Puntos	Numérico	12
Notación científica	Numérico	12
Fecha	Numérico	(Fecha) p.ej., MMDDAA10,...
Fecha (Hora)	Numérico	Hora18
Dólar	Numérico	12
Moneda personalizada	Numérico	12
Cadena	Carácter	\$8

### ***Almacenamiento de archivos de datos en formato de Stata***

- Los datos pueden escribirse en formato Stata 5-8, así como en formato Intercooled y SE (sólo en las versiones 7 y 8).
- Los archivos de datos que se guardan en formato Stata 5 se pueden leer con Stata 4.
- Los primeros 80 bytes de etiquetas de variable se guardan como etiquetas de variable Stata.

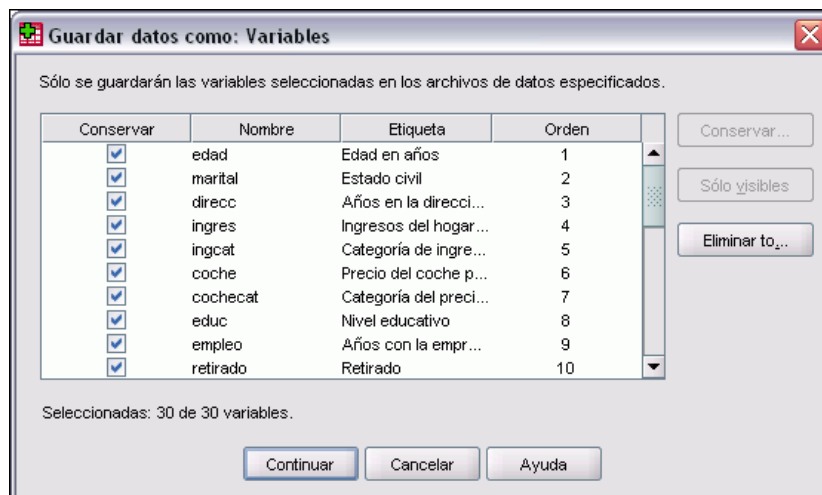
- Para variables numéricas, los primeros 80 bytes de etiquetas de valor se guardan como etiquetas de variable Stata. Para las variables de cadena, las etiquetas de valor se excluyen.
- Para las versiones 7 y 8, los primeros 32 bytes de nombres de variable en formato que distingue entre mayúsculas y minúsculas se guardan como nombres de variable Stata. Para versiones anteriores, los primeros ocho bytes de nombres de variable se guardan como nombres de variable Stata. Cualquier carácter distinto de letras, número y caracteres de subrayado se convierten en caracteres de subrayado.
- Los nombres de variable de SPSS que contienen caracteres de varios bytes (por ejemplo, caracteres japoneses o chinos) se convierten en nombres de variable con formato general *Vnnn*, donde *nnn* es un valor entero.
- Para las versiones 5-6 y las versiones Intercooled 7-8, se guardan los primeros 80 bytes de valores de cadena. Para Stata SE 7-8, se guardan los primeros 244 bytes de valores de cadena.
- Para las versiones 5-6 y las versiones Intercooled 7-8, sólo se guardan los primeros 2.047 bytes de valores de cadena. Para Stata SE 7-8, sólo se guardan las primeras 32.767 variables.

<b>Tipo de variable de SPSS</b>	<b>Tipo de variable de Stata</b>	<b>Formato de datos de Stata</b>
Numérico	Numérico	g
Coma	Numérico	g
Punto	Numérico	g
Notación científica	Numérico	g
Fecha*, Momento_fecha	Numérico	D_m_A
Tiempo, Tiempo_fecha	Numérico	g (número de segundos)
Día_semana	Numérico	g (1-7)
Mes	Numérico	g (1-12)
Dólar	Numérico	g
Moneda personalizada	Numérico	g
Cadena	Cadena	s

\*Date, Adate, Edate, SDate, Jdate, Qyr, Moyr, Wkyr

## Almacenamiento de subconjuntos de variables

Figura 3-24  
Cuadro de diálogo Guardar datos como: Variables



El cuadro de diálogo Guardar datos como: Variables permite seleccionar las variables que desea guardar en el nuevo archivo de datos. Por defecto, se almacenarán todas las variables. Anule la selección de las variables que no desea guardar o pulse en Eliminar todo y, a continuación, seleccione aquellas variables que desea guardar.

**Sólo visibles.** Selecciona sólo variables de conjuntos de variables que se usan actualmente. Si desea obtener más información, consulte [Utilizar conjuntos de variables](#) en Capítulo 44 en p. 475.

### Para guardar un subconjunto de variables

- ▶ Active la ventana Editor de datos (pulse en cualquier punto de la ventana para activarla).
- ▶ Elija en los menús:  
Archivo  
  Guardar como...
- ▶ Pulse en Variables.
- ▶ Seleccione las variables que desee almacenar.

## Exportación a base de datos

El Asistente para la exportación a base de datos permite:

- Reemplazar los valores de los campos (columnas) de la tabla de la base de datos existente o añadir nuevos campos a una tabla.
- Añadir nuevos registros (filas) a una tabla de base de datos.
- Reemplazar completamente una tabla de base de datos o crear una tabla nueva.

Para exportar datos a una base de datos:

- ▶ En los menús de la ventana del Editor de datos correspondientes al conjunto de datos que contiene los datos que se desean exportar, seleccione:
  - Archivo
  - Exportar a base de datos
- ▶ Seleccione el origen de base de datos.
- ▶ Siga las instrucciones del asistente para exportación para exportar los datos.

### **Creación de campos de base de datos a partir de variables de SPSS**

Al crear nuevos campos (añadiendo campos a una tabla de base de datos existente, creando una tabla nueva o reemplazando una tabla), puede especificar los nombres de campo, el tipo de datos y el ancho (donde corresponda).

**Nombre de campo.** Los nombres de campo por defecto son los mismos que los nombres de variable de SPSS . Puede cambiar los nombres de campo a cualquier nombre permitido por el formato de la base de datos. Por ejemplo, muchas bases de datos admiten que los nombres de los campos contengan caracteres que no se permiten en los nombres de variable, incluidos los espacios. Por tanto, un nombre de variable como *LlamadaEnEspera* se podría cambiar al nombre de campo *Llamada en espera*.

**Tipo.** El asistente para exportación realiza las asignaciones de los tipos de datos iniciales basados en los tipos de datos ODBC estándar o los tipos de datos permitidos por el formato de la base de datos seleccionada que sea más similar al formato de datos de SPSS , pero las bases de datos pueden realizar distinciones que no tengan equivalente directo en SPSS y viceversa. Por ejemplo, la mayoría de los valores numéricos de SPSS se almacenan como valores en punto flotante con doble precisión, mientras que los tipos de datos numéricos de las bases de datos incluyen números flotantes (doble), enteros, reales, etc. Además, muchas bases de datos no tienen equivalentes a los formatos de tiempo de SPSS . Puede cambiar el tipo de datos a cualquiera de los disponibles en la lista desplegable.

Como norma general, el tipo de datos básico (de cadena o numéricos) de la variable debe coincidir con el tipo de datos básico del campo de la base de datos. Si existe alguna discrepancia de tipo de datos que la base de datos no pueda resolver, se producirá un error y los datos no se exportarán a la base de datos. Por ejemplo, si exporta una variable de cadena a un campo de la base de datos con un tipo de datos numérico, se producirá un error si algún valor de la variable de cadena contiene caracteres no numéricos.

**Ancho.** Puede cambiar el ancho definido de los tipos de campo de cadena (char, varchar). Los anchos de campo numérico se definen por el tipo de datos.

Por defecto, los formatos de las variables de SPSS se asignan a tipos de campo de la base de datos en función del siguiente esquema general. Los tipos de campo de la base de datos reales pueden variar dependiendo de la base de datos.

<b>Formato de variable de SPSS</b>	<b>Tipo de campo de base de datos</b>
Numérico	Flotante o doble
Coma	Flotante o doble
Punto	Flotante o doble
Notación científica	Flotante o doble
Fecha	Fecha o Momento_fecha o marca de hora
Momento_fecha	Momento_fecha o marca de hora
Tiempo, Tiempo_fecha	Flotante o doble (número de segundos)
Día_semana	Entero (1-7)
Mes	Entero (1-12)
Dólar	Flotante o doble
Moneda personalizada	Flotante o doble
Cadena	Char or Varchar

### ***Valores definidos como perdidos por el usuario***

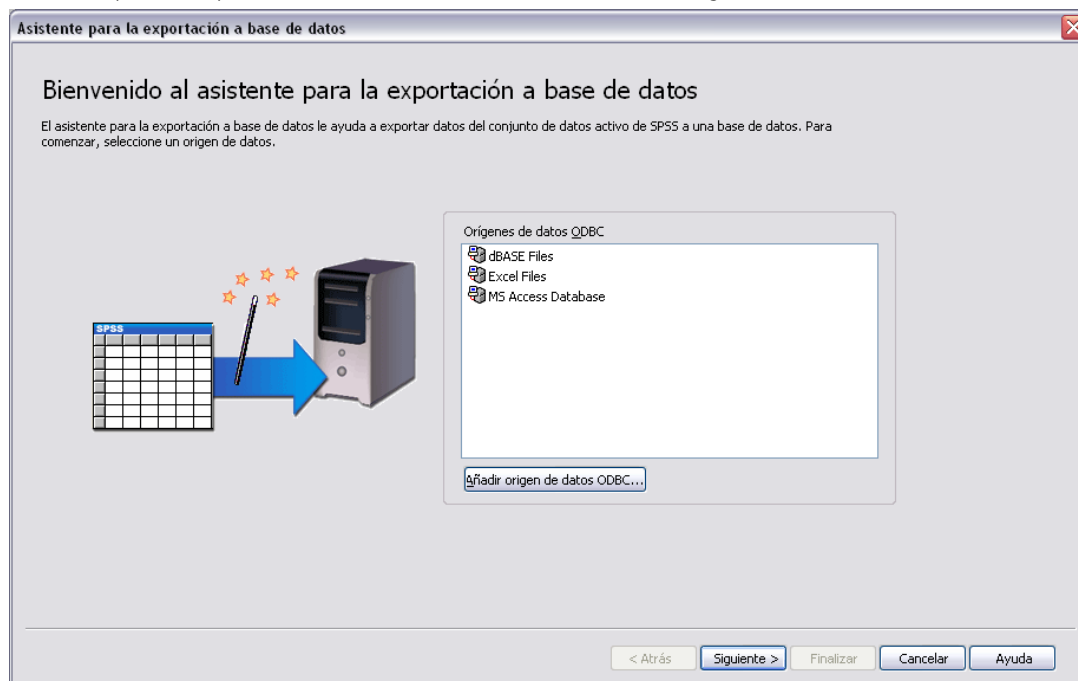
Existen dos opciones para el tratamiento de los valores perdidos definidos por el usuario cuando los datos de las variables se exportan a campos de bases de datos:

- **Exportar como valores válidos.** Los valores perdidos definidos por el usuario se tratan como valores no perdidos, válidos, regulares.
- **Exportar valores numéricos definidos como perdidos por el usuario como nulos y exportar valores de cadena definidos como perdidos por el usuario como espacios en blanco.** Los valores numéricos definidos como perdidos por el usuario reciben el mismo tratamiento que los valores perdidos del sistema. Los valores definidos como perdidos por el usuario se convierten en espacios en blanco (las cadenas no pueden ser valores perdidos del sistema).

### ***Selección de un origen de datos***

En el primer panel del Asistente para la exportación a base de datos, seleccione el origen de datos al que desea exportar los datos.

**Figura 3-25**  
Asistente para la exportación a base de datos, selección de un origen de datos



Puede exportar datos a cualquier origen de base de datos para el que tenga el controlador ODBC adecuado. (*Nota:* no se admite la exportación a orígenes de datos OLE DB.)

Si no tiene configurado ningún origen de datos ODBC o si desea añadir uno nuevo, pulse en Añadir origen de datos ODBC.

- En los sistemas operativos Linux, este botón no está disponible. Los orígenes de datos ODBC se especifican en *odbc.ini* y es necesario especificar las variables de entorno *ODBCINI* con la ubicación de dicho archivo. Si desea obtener más información, consulte la documentación de los controladores de la base de datos.
- En el análisis en modo distribuido (disponible con SPSS Server), este botón no está disponible. Para añadir orígenes de datos en el análisis en modo distribuido, consulte con el administrador del sistema.

Un origen de datos ODBC está compuesto por dos partes esenciales de información: el controlador que se utilizará para acceder a los datos y la ubicación de la base de datos a la que se desea acceder. Para especificar los orígenes de datos, deberán estar instalados los controladores adecuados. Para el análisis en modo local, puede instalar los controladores de varios formatos de bases de datos desde el CD de instalación de SPSS .

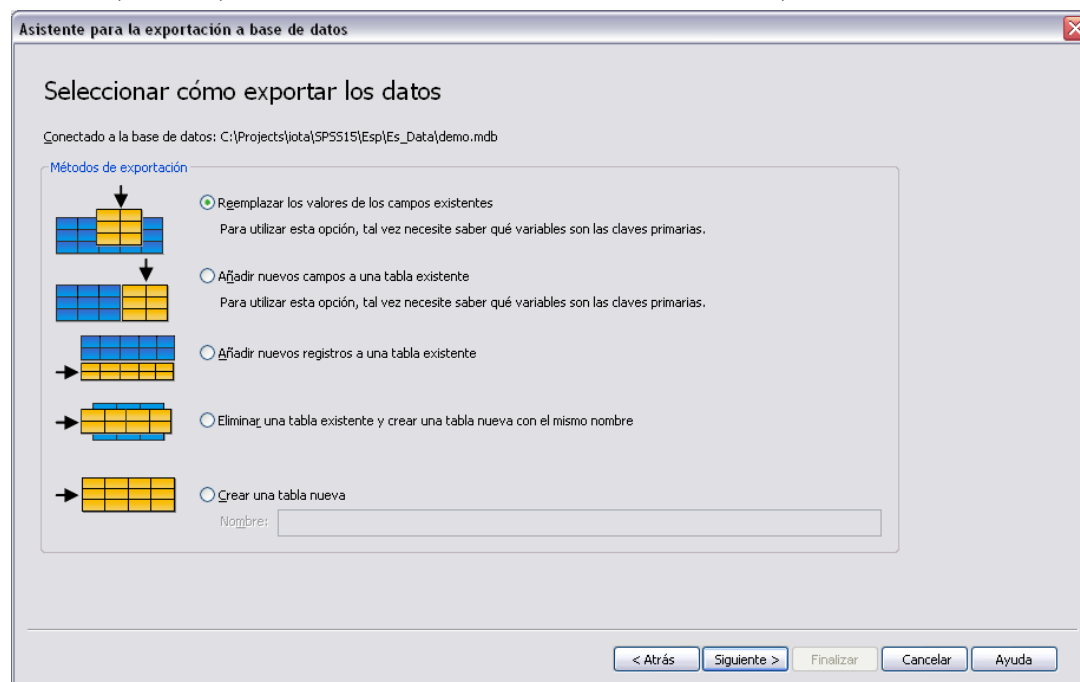
Algunos orígenes de datos pueden requerir un ID de acceso y una contraseña antes de poder continuar con el siguiente paso.

### Selección del modo de exportar los datos

Una vez seleccionado el origen de datos, se indica la forma en la que se desean exportar los datos.

Figura 3-26

Asistente para la exportación a base de datos, selección del modo de exportación



Las siguientes opciones están disponibles para exportar datos a una base de datos:

- **Reemplazar los valores de los campos existentes.** Reemplaza los valores de los campos seleccionados en una tabla existente con valores de las variables seleccionadas en el conjunto de datos activo. Si desea obtener más información, consulte [Sustitución de los valores de los campos existentes](#) en p. 60.
- **Añadir nuevos campos a una tabla existente.** Crea nuevos campos en una tabla existente que contiene los valores de las variables seleccionadas en el conjunto de datos activo. Si desea obtener más información, consulte [Adición de nuevos campos](#) en p. 61. Esta opción no está disponible para los archivos de Excel.
- **Añadir nuevos registros a una tabla existente.** Añade nuevos registros (filas) a una tabla existente que contiene los valores de los casos del conjunto de datos activo. Si desea obtener más información, consulte [Adición de nuevos registros \(casos\)](#) en p. 62.
- **Eliminar una tabla existente y crear una tabla nueva con el mismo nombre.** Elimina la tabla especificada y crea una nueva tabla con el mismo nombre que contiene variables seleccionadas del conjunto de datos activo. Toda la información de la tabla original, incluidas las definiciones de las propiedades del campo (como las claves primarias o los tipos de



datos) se pierde. Si desea obtener más información, consulte [Creación de una nueva tabla o sustitución de una tabla](#) en p. 63.

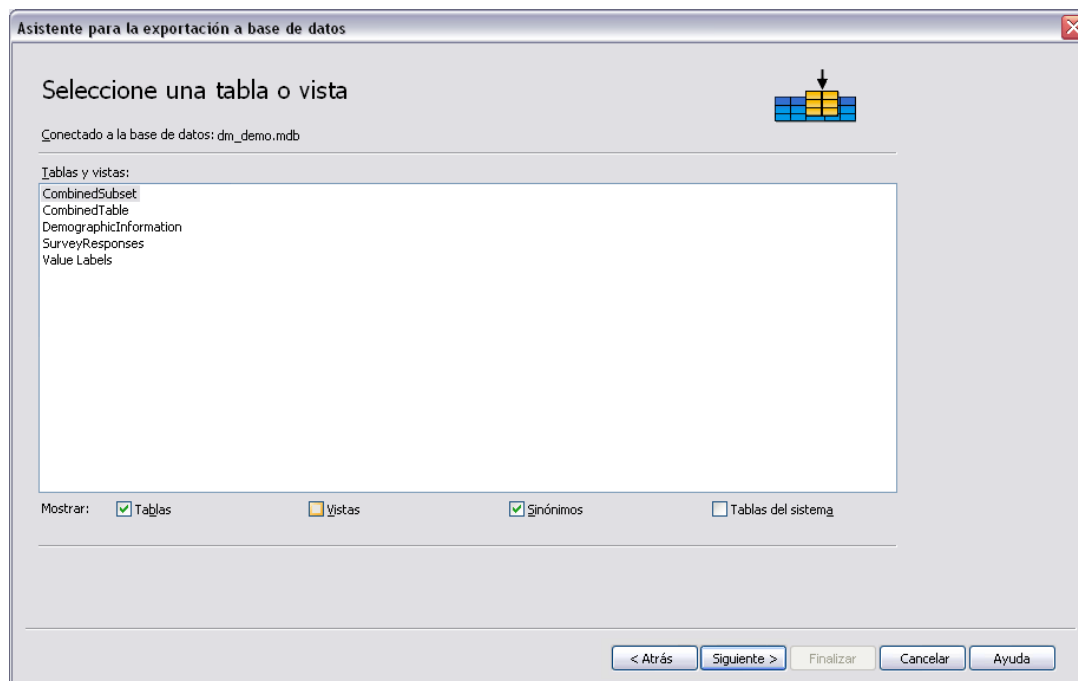
- **Crear una tabla nueva.** Crea una tabla nueva en la base de datos que contiene datos de las variables seleccionadas en el conjunto de datos activo. El nombre puede ser cualquier valor que esté permitido como nombre de tabla por el origen de datos. El nombre no puede coincidir con el nombre de una tabla o vista existentes en la base de datos. Si desea obtener más información, consulte [Creación de una nueva tabla o sustitución de una tabla](#) en p. 63.

### Selección de una tabla

Al modificar o reemplazar una tabla de la base de datos, es necesario seleccionar la tabla que desea modificar o reemplazar. Este panel del Asistente para la exportación a bases de datos muestra una lista de tablas y vistas de la base de datos seleccionada.

Figura 3-27

Asistente para la exportación a base de datos, selección de una tabla o vista



Por defecto, la lista muestra sólo las tablas de bases de datos estándar. Puede controlar el tipo de elementos que se muestran en la lista:

- **Tablas.** Tablas de base de datos estándar.
- **Vistas.** Las vistas son “tablas” virtuales o dinámicas definidas por consultas. Estas tablas pueden incluir uniones de varias tablas y/o campos derivados de cálculos basados en los valores de otros campos. Puede añadir registros o reemplazar valores de campos existentes en vistas, pero es posible que los campos que se pueden modificar estén limitados dependiendo de cómo esté estructurada la vista. Por ejemplo, no se puede modificar un campo derivado, añadir campos a una vista ni reemplazar una vista.

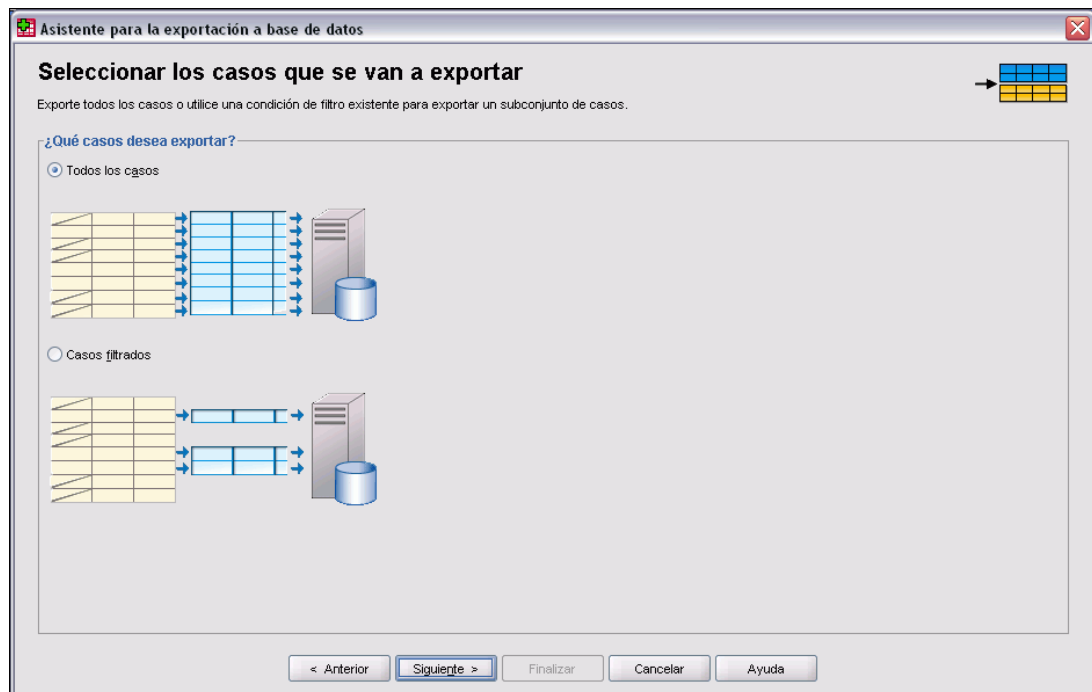
- **Sinónimos.** Un sinónimo es un alias para una tabla o vista que suele estar definido en una consulta.
- **Tablas del sistema.** Las tablas del sistema definen propiedades de la base de datos. En algunos casos, las tablas de base de datos estándar pueden estar clasificadas como tablas del sistema y sólo se mostrarán si se selecciona esta opción. El acceso a tablas del sistema reales suele estar limitado a los administradores de la base de datos.

### **Selección de casos para exportar**

La selección de casos en el Asistente para la exportación a base de datos está limitada, bien a todos los casos o a los casos seleccionados a través de una condición de filtrado definida previamente. Si no hay ningún filtrado de casos activo, este panel no aparecerá y se exportarán todos los casos del conjunto de datos activo.

Figura 3-28

Asistente para la exportación a base de datos, selección de los casos para exportar



Si desea obtener información sobre la definición de una condición de filtrado para la selección de casos, consulte [Seleccionar casos en p. 194](#).

### **Emparejamiento de casos con registros**

Al añadir campos (columnas) a una tabla existente o reemplazar los valores de los campos existentes, es necesario asegurarse de que cada caso (fila) del conjunto de datos activo coincide correctamente con el correspondiente registro de la base de datos.

- En la base de datos, el campo o conjunto de campos que identifica de forma única cada registro suele estar designado como la **clave primaria**.

- Debe identificar las variables correspondientes a los campos de clave primaria u otros campos que identifican de manera única a cada registro.
- Los campos no tienen que ser la clave primaria de la base de datos, sin embargo, el valor de campo o la combinación de los valores de campo deben ser exclusivos para cada caso.

Para casar las variables con los campos de la base de datos que identifican cada registro de forma única:

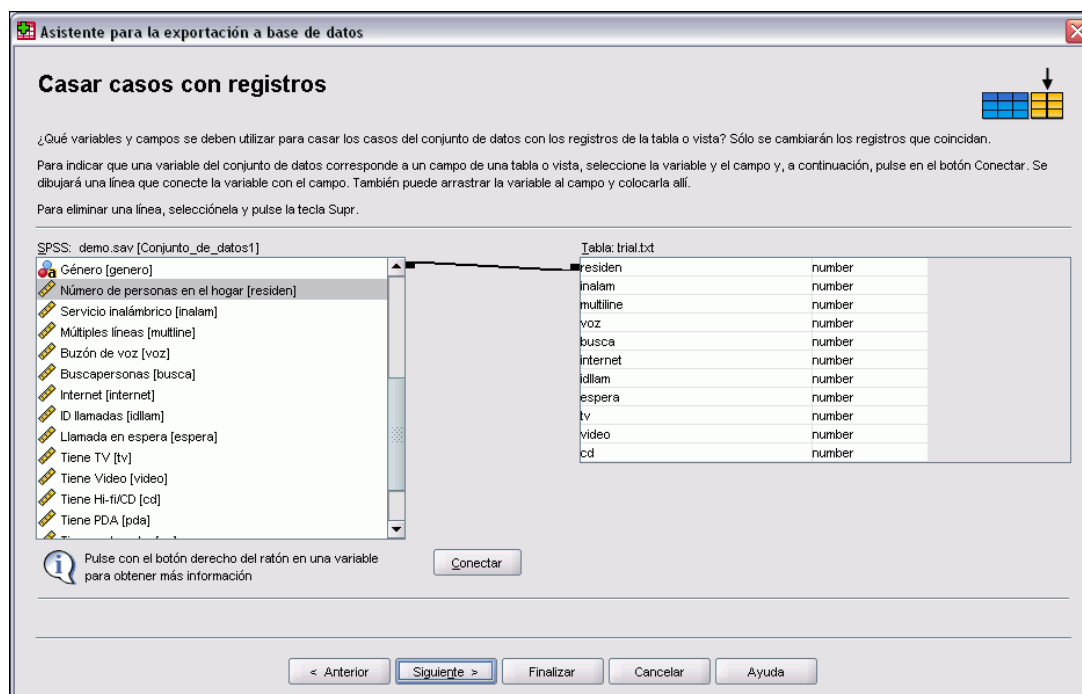
- ▶ Arrastre y coloque las variables en los campos correspondientes de la base de datos.
- o
- ▶ Seleccione una variable de la lista de variables, seleccione el campo correspondiente en la tabla de la base de datos y pulse en Conectar.

Para eliminar una línea de conexión:

- ▶ Seleccione la línea de conexión y pulse la tecla Supr.

Figura 3-29

Asistente para la exportación a base de datos, emparejamiento de casos con registros



**Nota:** los nombres de variable y los nombres de los campos de la base de datos es posible que no sean idénticos (ya que los nombres de la base de datos pueden contener caracteres que no admiten los nombres de variable de SPSS), pero si el conjunto de datos activo se creó a partir de la tabla de base de datos que está modificando, los nombres de variable o las etiquetas de variable normalmente serán como mínimo similares a los nombres de campo de la base de datos.

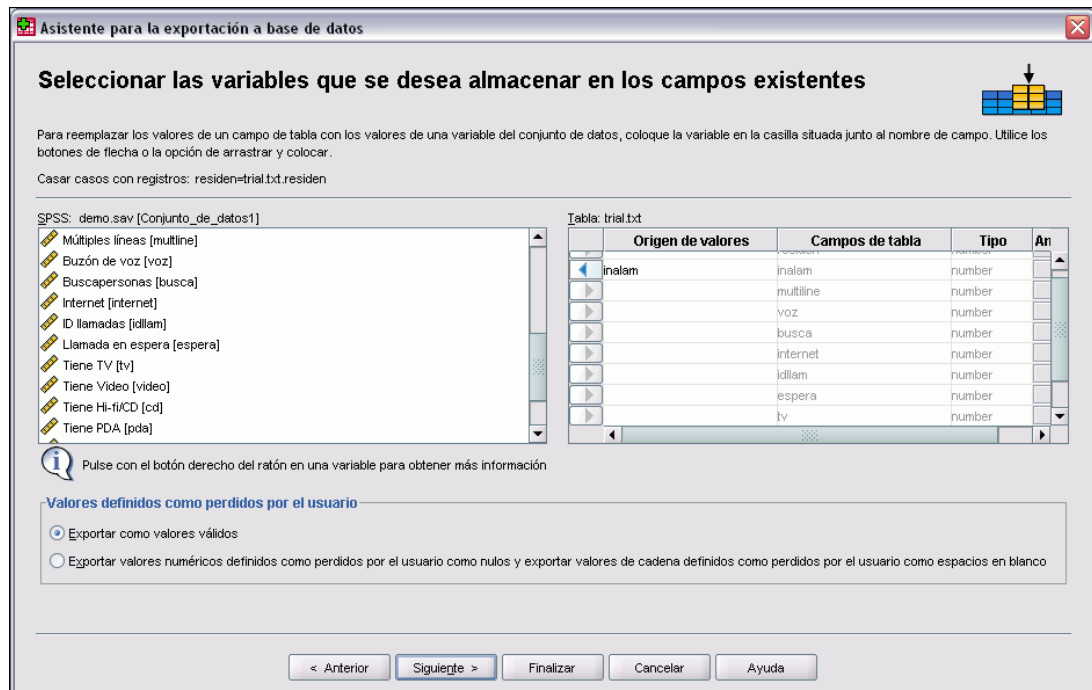
### Sustitución de los valores de los campos existentes

Para reemplazar los valores de los campos existentes en una base de datos:

- ▶ En el panel Seleccionar cómo exportar los datos del Asistente para la exportación a base de datos, seleccione Reemplazar los valores de los campos existentes.
- ▶ En el panel Seleccione una tabla o vista, seleccione la tabla de base de datos.
- ▶ En el panel Casar casos con registros, case las variables que identifican de forma única cada caso con los nombres de los campos de la base de datos correspondientes.
- ▶ Para cada campo del que desee reemplazar los valores, arrastre la variable que contiene los nuevos valores y colóquela en la columna Origen de valores, junto al nombre del campo de la base de datos correspondiente.

Figura 3-30

Asistente para la exportación a base de datos, sustitución de los valores de los campos existentes



- Como norma general, el tipo de datos básico (de cadena o numéricos) de la variable debe coincidir con el tipo de datos básico del campo de la base de datos. Si existe alguna discordancia de tipos de datos que la base de datos no pueda resolver, se producirá un error y no se exportará ningún dato a la base de datos. Por ejemplo, si exporta una variable de cadena a un campo de la base de datos con un tipo de datos numérico (por ejemplo, doble, real, entero), se producirá un error si algún valor de la variable de cadena contiene caracteres no numéricos. La letra *a* del icono situado junto a una variable denota una variable de cadena.
- No se puede modificar el nombre, el tipo ni la anchura del campo. Los atributos del campo de la base de datos originales se conservan, sólo se reemplazan los valores.

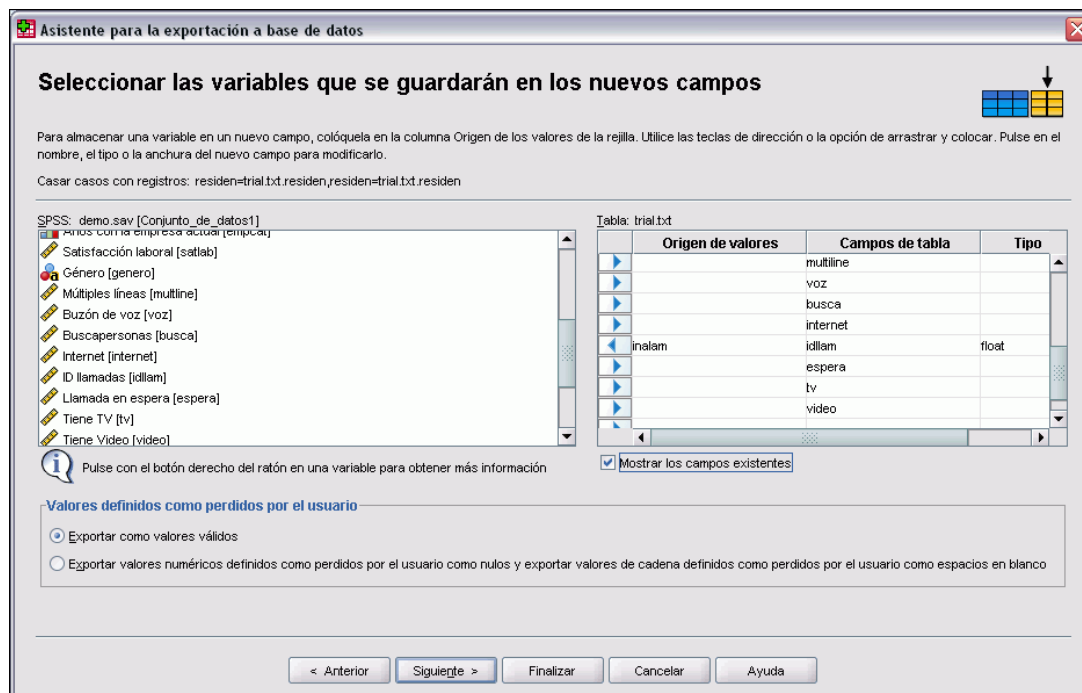
## Adición de nuevos campos

Para añadir nuevos campos a una tabla de base de datos existente:

- ▶ En el panel Seleccionar cómo exportar los datos del Asistente para la exportación a base de datos, seleccione Añadir nuevos campos a una tabla existente.
- ▶ En el panel Seleccione una tabla o vista, seleccione la tabla de base de datos.
- ▶ En el panel Casar casos con registros, case las variables que identifican de forma única cada caso con los nombres de los campos de la base de datos correspondientes.
- ▶ Arrastre las variables que desea añadir como campos nuevos y colóquelas en la columna Origen de valores.

Figura 3-31

Asistente para la exportación a base de datos, adición de nuevos campos a una tabla existente



Si desea obtener información sobre los nombres de campo y los tipos de datos, consulte la sección sobre la creación de campos de bases de datos a partir de las variables de SPSS en [Exportación a base de datos en p. 52](#).

**Mostrar los campos existentes.** Seleccione esta opción para mostrar una lista de campos existentes. No puede utilizar este panel en el Asistente para la exportación a base de datos para reemplazar campos existentes, pero puede resultar útil saber los campos que ya están presentes en la tabla. Si desea reemplazar los valores de los campos existentes, consulte [Sustitución de los valores de los campos existentes en p. 60](#).

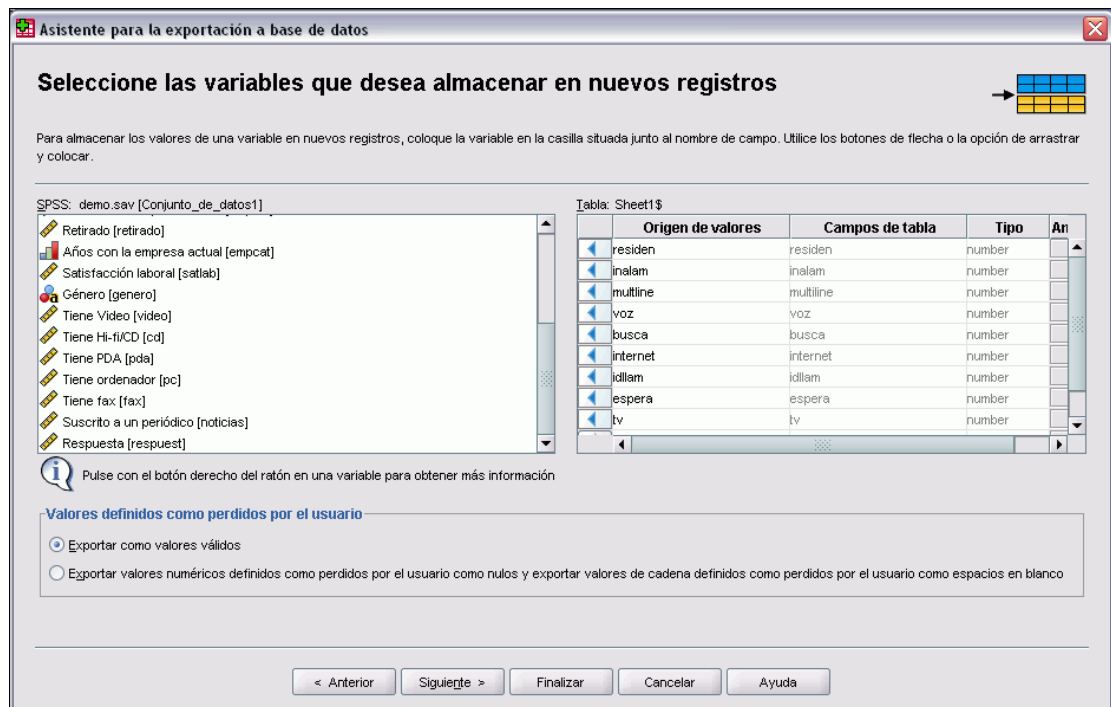
### Adición de nuevos registros (casos)

Para añadir nuevos registros (caso) a una tabla de base de datos:

- ▶ En el panel Seleccionar cómo exportar los datos del Asistente para la exportación a base de datos, seleccione Añadir nuevos registros a una tabla existente.
- ▶ En el panel Seleccione una tabla o vista, seleccione la tabla de base de datos.
- ▶ Haga coincidir las variables del conjunto de datos activo con los campos de la tabla arrastrando las variables y colocándolas en la columna Origen de valores.

Figura 3-32

Asistente para la exportación a base de datos, adición de nuevos registros (casos) a una tabla



El Asistente para la exportación a base de datos seleccionará automáticamente todas las variables que coincidan con los campos existentes utilizando la información sobre la tabla de base de datos original almacenada en el conjunto de datos activo (si está disponible) y/o los nombres de las variables que coinciden con los nombres de campo. Este emparejamiento inicial automático sólo pretende ser una guía y permite cambiar la forma en que se hacen coincidir variables con los campos de la base de datos.

Al añadir nuevos registros a una tabla existente, se aplican las siguientes reglas y limitaciones básicas:

- Todos los casos (o todos los casos seleccionados) en el conjunto de datos activo se añaden a la tabla. Si alguno de estos casos duplica los registros existentes en la base de datos, puede producirse un error si se encuentra un valor de clave duplicado. Si desea obtener

información sobre la exportación sólo de los casos seleccionados, consulte [Selección de casos para exportar en p. 58](#).

- Puede utilizar los valores de las variables nuevas creadas en la sesión como los valores de los campos existentes, pero no puede añadir campos nuevos ni cambiar los nombres de los existentes. Para añadir campos nuevos a una tabla de base de datos, consulte [Adición de nuevos campos en p. 61](#).
- Cualquier campo de la base de datos excluido que no coincida con una variable no tendrá ningún valor para los registros añadidos a la tabla de base de datos. (Si la casilla Origen de valores está vacía, no habrá ninguna variable que coincida con el campo.)

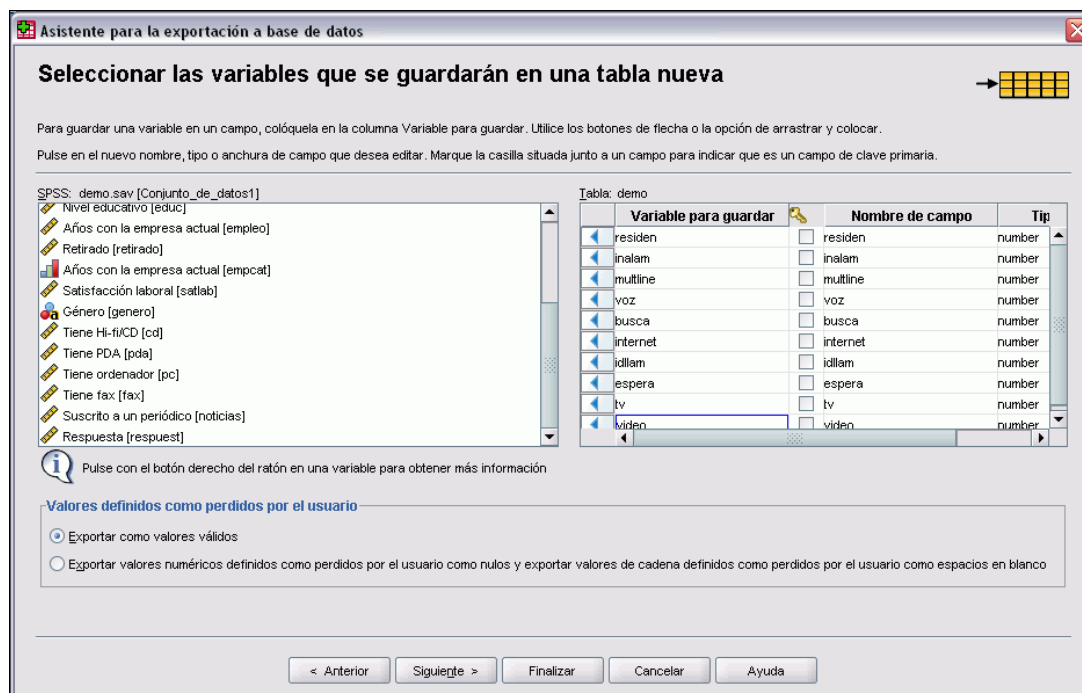
### Creación de una nueva tabla o sustitución de una tabla

Para crear una tabla de base de datos nueva o reemplazar una tabla de base de datos existente:

- ▶ En el panel Seleccionar cómo exportar los datos del asistente para la exportación, seleccione Eliminar una tabla existente y crear una tabla nueva con el mismo nombre o seleccione Crear una tabla nueva e introduzca un nombre para la nueva tabla.
- ▶ Si está reemplazando una tabla existente, en el panel Seleccione una tabla o vista, seleccione la tabla de base de datos.
- ▶ Arrastre las variables y colóquelas en la columna Variable para guardar.
- ▶ Si lo desea, puede designar variables o campos que definan la clave primaria, cambiar nombres de campos y cambiar el tipo de datos.

Figura 3-33

Asistente para la exportación a base de datos, selección de variables para una tabla nueva



**Clave primaria.** Para designar variables como la clave primaria de la tabla de base de datos, marque la casilla de la columna identificada con el icono de llave.

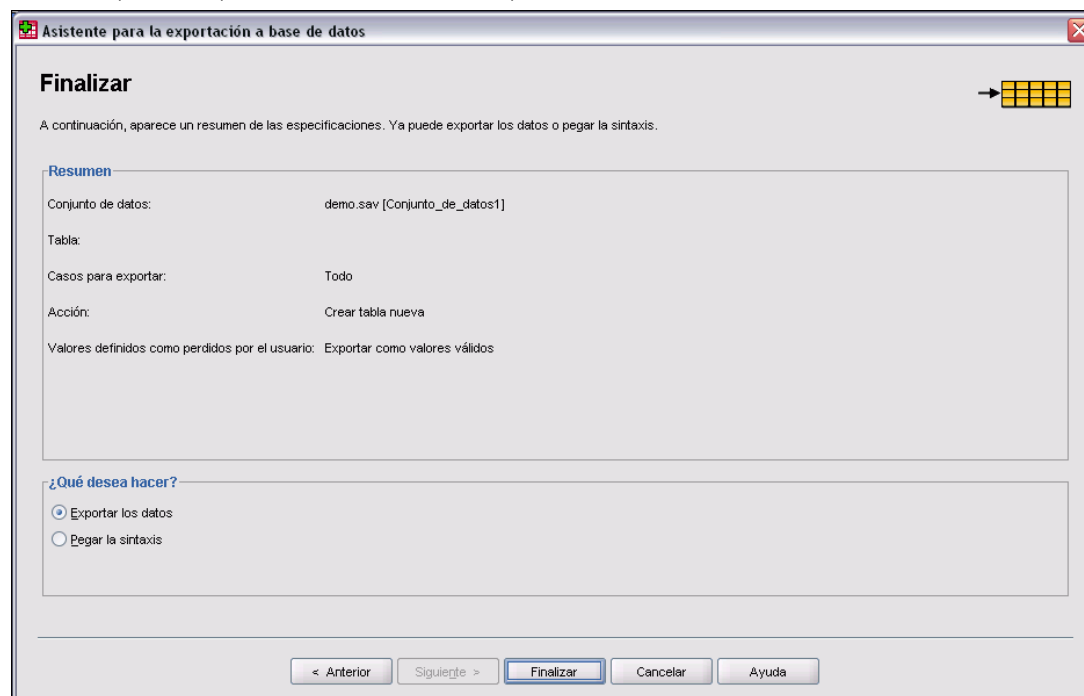
- Todos los valores de la clave primaria deben ser únicos, de lo contrario, se producirá un error.
- Si selecciona una única variable como la clave primaria, cada registro (caso) debe tener un valor único para esa variable.
- Si selecciona varias variables como clave primaria, esto define una clave primaria compuesta y la *combinación* de valores para las variables seleccionadas debe ser única para cada caso.

Si desea obtener información sobre los nombres de campo y los tipos de datos, consulte la sección sobre la creación de campos de bases de datos a partir de las variables de SPSS en [Exportación a base de datos en p. 52](#).

### **Finalización del Asistente para la exportación a base de datos**

El último panel del Asistente para la exportación a base de datos proporciona un resumen que indica los datos que se exportarán y cómo se exportarán. Además, ofrece la opción de exportar los datos o pegar la sintaxis de comandos subyacente en una ventana de sintaxis.

Figura 3-34  
Asistente para la exportación a base de datos, panel Finalizar



### **Información de resumen**

- **Conjunto de datos.** El nombre de sesión de SPSS para el conjunto de datos que se utilizará para exportar datos. Esta información es útil principalmente si existen varios orígenes de datos abiertos. Los orígenes de datos abiertos mediante la interfaz gráfica de usuario (por ejemplo, el Asistente para base de datos) tienen asignados automáticamente nombres



como *Conjunto\_de\_datos1*, *Conjunto\_de\_datos2*, etc. Un origen de datos abierto mediante sintaxis de comandos sólo tendrá un nombre de conjunto de datos si se ha asignado alguno explícitamente.

- **Tabla.** El nombre de la tabla que se va a modificar o crear.
- **Casos para exportar.** Se pueden exportar todos los casos o los seleccionados por un filtro previamente definido. Si desea obtener más información, consulte [Selección de casos para exportar](#) en p. 58.
- **Acción.** Indica cómo se modificará la base de datos (por ejemplo, crear una tabla nueva, añadir campos o registros a una tabla existente).
- **Valores definidos como perdidos por el usuario.** Los valores definidos como perdidos por el usuario se pueden exportar como valores válidos o se pueden tratar como valores perdidos del sistema para las variables numéricas y convertirlos en espacios en blanco para las variables de cadena. Este ajuste se controla en el panel en el que se seleccionan las variables que se van a exportar.

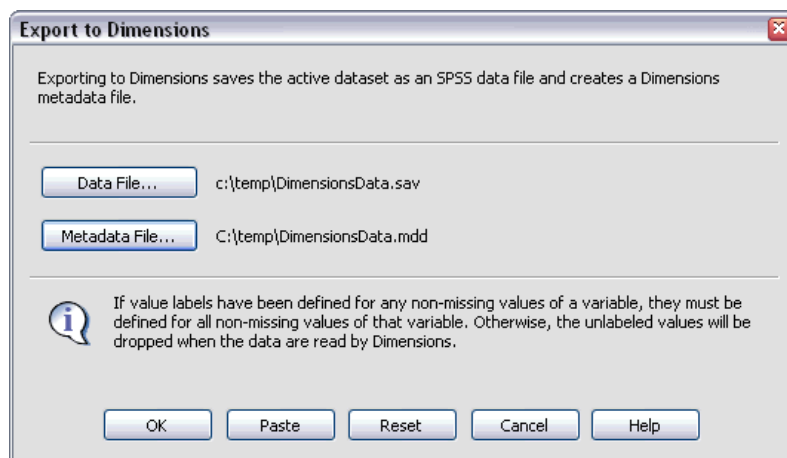
## ***Exportación a Dimensions***

El cuadro de diálogo Exportar a Dimensions crea un archivo de datos de SPSS y un archivo de metadatos de Dimensions que puede utilizar para leer los datos en las aplicaciones de Dimensions, como mrInterview y mrTables. Resulta particularmente útil cuando los datos van y vienen entre las aplicaciones de Dimensions y SPSS. Por ejemplo, puede leer un origen de datos de mrInterview en SPSS, calcular algunas nuevas variables y, a continuación, guardar los datos en un formato que mrTables pueda leer sin perder ninguno de los atributos de metadatos originales.

Para exportar datos que se van a utilizar en aplicaciones de Dimensions:

- ▶ En los menús de la ventana del Editor de datos correspondientes al conjunto de datos que contiene los datos que se desean exportar, seleccione:  
Archivo  
Exportar a Dimensions
- ▶ Pulse en Archivo de datos para especificar el nombre y la ubicación del archivo de datos de SPSS.
- ▶ Pulse en Archivo de metadatos para especificar el nombre y la ubicación del archivo de datos de Dimensions.

Figura 3-35  
Cuadro de diálogo *Exportar a Dimensions*



Para nuevas variables y conjuntos de datos no creados a partir de orígenes de datos de Dimensions, los atributos de variable de SPSS se asignan a atributos de metadatos de Dimensions en el archivo de metadatos siguiendo los métodos descritos en la documentación de SPSS SAV DSC en la biblioteca de desarrollo de Dimensions.

Si el conjunto de datos activo se ha creado a partir de un origen de datos de Dimensions:

- El nuevo archivo de metadatos se crea fusionando los atributos de metadatos originales con los atributos de metadatos de todas las nuevas variables, además de todos los cambios realizados a las variables originales que puedan afectar a sus atributos de metadatos (por ejemplo, adición o cambios de las etiquetas de variable).
- Para las variables originales leídas del origen de datos de Dimensions, todos los atributos de metadatos no reconocidos por SPSS se conservan en su estado original. Por ejemplo, SPSS convierte las variables de cuadrícula en variables de SPSS normales, pero los metadatos que definen dichas variables de cuadrículas se conservan al guardar el nuevo archivo de metadatos.
- Si los nombres de todas las variables de Dimensions se cambiaron automáticamente para que cumplieren las normas de denominación de variables de SPSS, el archivo de metadatos asigna los nombres convertidos a los nombres de variable originales de Dimensions.

La presencia o ausencia de etiquetas de valor puede afectar a los atributos de metadatos de las variables y, por tanto, a la manera en que dichas variables son leídas por las aplicaciones de Dimensions. Si se han definido etiquetas de valor para algún valor que no sea perdido de una variable, deberán definirse para todos los valores que no sean perdidos de dicha variable ya que, de no ser así, Dimensions eliminará los valores no etiquetados al leer el archivo de datos.

## ***Protección de datos originales***

Para evitar la modificación o eliminación accidental de los datos originales, puede marcar el archivo como un archivo de sólo lectura.

- En los menús del Editor de datos, elija:

Archivo  
 Marcar archivo como de sólo lectura

Si hace modificaciones posteriores de los datos y, a continuación, intenta guardar el archivo de datos, puede guardar los datos sólo con un nombre de archivo distinto; así, los datos originales no se verán afectados.

Puede restablecer los permisos de archivo a lectura/escritura seleccionando la opción Marcar archivo como de lectura/escritura en el menú Archivo.

## Archivo activo virtual

El archivo activo virtual permite trabajar con grandes archivos de datos sin que sea necesaria una cantidad igual de grande (o mayor) de espacio temporal en disco. Para la mayoría de los procedimientos de análisis y gráficos, el origen de datos original se vuelve a leer cada vez que se ejecuta un procedimiento diferente. Los procedimientos que modifican los datos necesitan una cierta cantidad de espacio temporal en disco para realizar un seguimiento de los cambios; además, algunas acciones necesitan disponer siempre de la cantidad suficiente de espacio en disco para, al menos, una copia completa del archivo de datos.

Figura 3-36  
 Requisitos de espacio temporal en disco

Acción	GET FILE = 'v1-5.sav'. FRECUENCIAS...	COMPUTE v6 = ... RECODE v4... REGRESSION... /SAVE ZPRED.	SORT CASES BY... o CACHE																																																																																																																																					
Archivo virtual activo	<table border="1"> <thead> <tr><th>v1</th><th>v2</th><th>v3</th><th>v4</th><th>v5</th></tr> </thead> <tbody> <tr><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td></tr> <tr><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td></tr> <tr><td>31</td><td>32</td><td>33</td><td>34</td><td>35</td></tr> <tr><td>41</td><td>42</td><td>43</td><td>44</td><td>45</td></tr> <tr><td>51</td><td>52</td><td>53</td><td>54</td><td>55</td></tr> <tr><td>61</td><td>62</td><td>63</td><td>64</td><td>65</td></tr> </tbody> </table>	v1	v2	v3	v4	v5	11	12	13	14	15	21	22	23	24	25	31	32	33	34	35	41	42	43	44	45	51	52	53	54	55	61	62	63	64	65	<table border="1"> <thead> <tr><th>v1</th><th>v2</th><th>v3</th><th>v4</th><th>v5</th><th>v6</th><th>zpre</th></tr> </thead> <tbody> <tr><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>1</td></tr> <tr><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>2</td></tr> <tr><td>31</td><td>32</td><td>33</td><td>34</td><td>35</td><td>36</td><td>3</td></tr> <tr><td>41</td><td>42</td><td>43</td><td>44</td><td>45</td><td>46</td><td>4</td></tr> <tr><td>51</td><td>52</td><td>53</td><td>54</td><td>55</td><td>56</td><td>5</td></tr> <tr><td>61</td><td>62</td><td>63</td><td>64</td><td>65</td><td>66</td><td>6</td></tr> </tbody> </table>	v1	v2	v3	v4	v5	v6	zpre	11	12	13	14	15	16	1	21	22	23	24	25	26	2	31	32	33	34	35	36	3	41	42	43	44	45	46	4	51	52	53	54	55	56	5	61	62	63	64	65	66	6	<table border="1"> <thead> <tr><th>v1</th><th>v2</th><th>v3</th><th>v4</th><th>v5</th><th>v6</th><th>zpre</th></tr> </thead> <tbody> <tr><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>1</td></tr> <tr><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>2</td></tr> <tr><td>31</td><td>32</td><td>33</td><td>34</td><td>35</td><td>36</td><td>3</td></tr> <tr><td>41</td><td>42</td><td>43</td><td>44</td><td>45</td><td>46</td><td>4</td></tr> <tr><td>51</td><td>52</td><td>53</td><td>54</td><td>55</td><td>56</td><td>5</td></tr> <tr><td>61</td><td>62</td><td>63</td><td>64</td><td>65</td><td>66</td><td>6</td></tr> </tbody> </table>	v1	v2	v3	v4	v5	v6	zpre	11	12	13	14	15	16	1	21	22	23	24	25	26	2	31	32	33	34	35	36	3	41	42	43	44	45	46	4	51	52	53	54	55	56	5	61	62	63	64	65	66	6
v1	v2	v3	v4	v5																																																																																																																																				
11	12	13	14	15																																																																																																																																				
21	22	23	24	25																																																																																																																																				
31	32	33	34	35																																																																																																																																				
41	42	43	44	45																																																																																																																																				
51	52	53	54	55																																																																																																																																				
61	62	63	64	65																																																																																																																																				
v1	v2	v3	v4	v5	v6	zpre																																																																																																																																		
11	12	13	14	15	16	1																																																																																																																																		
21	22	23	24	25	26	2																																																																																																																																		
31	32	33	34	35	36	3																																																																																																																																		
41	42	43	44	45	46	4																																																																																																																																		
51	52	53	54	55	56	5																																																																																																																																		
61	62	63	64	65	66	6																																																																																																																																		
v1	v2	v3	v4	v5	v6	zpre																																																																																																																																		
11	12	13	14	15	16	1																																																																																																																																		
21	22	23	24	25	26	2																																																																																																																																		
31	32	33	34	35	36	3																																																																																																																																		
41	42	43	44	45	46	4																																																																																																																																		
51	52	53	54	55	56	5																																																																																																																																		
61	62	63	64	65	66	6																																																																																																																																		
Datos almacenados en espacio temporal del disco	Ninguno	<table border="1"> <thead> <tr><th>v4</th><th>v6</th><th>zpre</th></tr> </thead> <tbody> <tr><td>14</td><td>16</td><td>1</td></tr> <tr><td>24</td><td>26</td><td>2</td></tr> <tr><td>34</td><td>36</td><td>3</td></tr> <tr><td>44</td><td>46</td><td>4</td></tr> <tr><td>54</td><td>56</td><td>5</td></tr> <tr><td>64</td><td>66</td><td>6</td></tr> </tbody> </table>	v4	v6	zpre	14	16	1	24	26	2	34	36	3	44	46	4	54	56	5	64	66	6	<table border="1"> <thead> <tr><th>v1</th><th>v2</th><th>v3</th><th>v4</th><th>v5</th><th>v6</th><th>zpre</th></tr> </thead> <tbody> <tr><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td><td>16</td><td>1</td></tr> <tr><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>2</td></tr> <tr><td>31</td><td>32</td><td>33</td><td>34</td><td>35</td><td>36</td><td>3</td></tr> <tr><td>41</td><td>42</td><td>43</td><td>44</td><td>45</td><td>46</td><td>4</td></tr> <tr><td>51</td><td>52</td><td>53</td><td>54</td><td>55</td><td>56</td><td>5</td></tr> <tr><td>61</td><td>62</td><td>63</td><td>64</td><td>65</td><td>66</td><td>6</td></tr> </tbody> </table>	v1	v2	v3	v4	v5	v6	zpre	11	12	13	14	15	16	1	21	22	23	24	25	26	2	31	32	33	34	35	36	3	41	42	43	44	45	46	4	51	52	53	54	55	56	5	61	62	63	64	65	66	6																																																															
v4	v6	zpre																																																																																																																																						
14	16	1																																																																																																																																						
24	26	2																																																																																																																																						
34	36	3																																																																																																																																						
44	46	4																																																																																																																																						
54	56	5																																																																																																																																						
64	66	6																																																																																																																																						
v1	v2	v3	v4	v5	v6	zpre																																																																																																																																		
11	12	13	14	15	16	1																																																																																																																																		
21	22	23	24	25	26	2																																																																																																																																		
31	32	33	34	35	36	3																																																																																																																																		
41	42	43	44	45	46	4																																																																																																																																		
51	52	53	54	55	56	5																																																																																																																																		
61	62	63	64	65	66	6																																																																																																																																		

### Las acciones que no necesitan ningún espacio temporal en disco son:

- La lectura de archivos de datos de SPSS
- La fusión de dos o más archivos de datos de SPSS
- La lectura de tablas de bases de datos con el Asistente para bases de datos
- La fusión de un archivo de datos de SPSS con una tabla de base de datos

- La ejecución de procedimientos que leen datos (por ejemplo, Frecuencias, Tablas de contingencia, Explorar)

***Las acciones que crean una o más columnas de datos en espacio temporal en disco son:***

- El cálculo de nuevas variables
- La recodificación de variables existentes
- La ejecución de procedimientos que crean o modifican variables (por ejemplo, almacenamiento de valores pronosticados en la regresión lineal)

***Las acciones que crean una copia completa del archivo de datos en espacio temporal en disco son:***

- La lectura de archivos de Excel
- La ejecución de procedimientos que ordenan los datos (por ejemplo, Ordenar casos, Segmentar archivo)
- La lectura de datos con los comandos `GET TRANSLATE` o `DATA LIST`
- La utilización de la unidad Datos de caché o el comando `CACHE`
- La activación de otras aplicaciones de SPSS que leen el archivo de datos (por ejemplo, AnswerTree, DecisionTime)

*Nota:* El comando `GET DATA` proporciona una funcionalidad comparable a `DATA LIST`, sin crear una copia completa del archivo de datos en el espacio temporal del disco. El comando `SPLIT FILE` de la sintaxis de comandos no ordena el archivo de datos y por lo tanto no crea una copia del archivo de datos. Este comando, sin embargo, necesita tener los datos ordenados para un funcionamiento apropiado y la interfaz del cuadro de diálogo para este procedimiento ordenará de forma automática el archivo de datos, con la consiguiente copia completa de dicho archivo. En la versión para estudiantes no está disponible la sintaxis de comandos.

***Acciones que crean una copia completa del archivo de datos por defecto:***

- Lectura de bases de datos con el Asistente para bases de datos
- La lectura de archivos de texto con el Asistente para la importación de texto

El Asistente para la importación de texto proporciona un ajuste opcional para crear de forma automática una caché de los datos. Por defecto, se selecciona esta opción. Para desactivar esta opción, simplemente desmarque la casilla de verificación Caché local de los datos. En el Asistente para bases de datos puede pegar la sintaxis de comando generada y eliminar el comando `CACHE`.

## ***Creación de una caché de datos***

Aunque el archivo actual virtual puede reducir de forma drástica la cantidad de espacio temporal en disco necesario, la falta de una copia temporal del archivo “activo” significa que el origen original de datos debe volver a leerse para cada procedimiento. Para archivos de datos grandes leídos desde un origen externo, la creación de una copia temporal de los datos puede mejorar el rendimiento. Por ejemplo, para tablas de datos leídas desde un origen de base de datos, la consulta SQL que lee la información de la base de datos debe volver a ejecutarse para cualquier comando o procedimiento que necesite leer los datos. Debido a que virtualmente todos los procedimientos de

análisis estadísticos y procedimientos gráficos necesitan leer los datos, la ejecución de la consulta SQL se repite para cada procedimiento, lo que puede significar un importante incremento en el tiempo de procesamiento si se ejecuta un gran número de procedimientos.

Si se dispone de suficiente espacio en disco en el ordenador que realiza el análisis (el ordenador local o el servidor remoto), se pueden eliminar varias consultas SQL y mejorar el tiempo de procesamiento mediante la creación de una caché de datos del archivo activo. La caché de datos es una copia temporal de todos los datos.

*Nota:* Por defecto, el Asistente para bases de datos crea de forma automática una caché de datos, pero si se utiliza el comando `GET DATA` en la sintaxis de comandos para leer una base de datos, no se creará una caché de datos de forma automática. En la versión para estudiantes no está disponible la sintaxis de comandos.

### **Para crear una caché de datos**

- ▶ Elija en los menús:
  - Archivo
  - Hacer caché de datos...
- ▶ Pulse en Aceptar o en Crear caché ahora.

Aceptar crea una caché de datos la siguiente vez que el programa lea los datos (por ejemplo, la próxima vez que se ejecute un procedimiento estadístico), que será lo que normalmente se quiera porque no necesita una lectura adicional de los datos. Crear caché ahora crea una caché de datos inmediatamente, lo cual no será necesario la mayoría de las veces. Crear caché ahora se utiliza principalmente por dos razones:

- Un origen de datos está “bloqueado” y no se puede actualizar por nadie hasta que finalice la sesión actual, abra un origen de datos diferente o haga una caché de los datos.
- Para grandes orígenes de datos, el desplazamiento por el contenido de la pestaña Vista de datos en el Editor de datos será mucho más rápido si se hace una caché de datos.

### **Para crear una caché de datos de forma automática**

Se puede utilizar el comando `SET` para crear de forma automática una caché de datos después de un número especificado de cambios en el archivo de datos activo. Por defecto, se crea una caché del archivo de datos de forma automática cada 20 cambios realizados sobre el archivo.

- ▶ Elija en los menús:
  - Archivo
  - Nuevo
  - Sintaxis
- ▶ En la ventana de sintaxis, escriba `SET CACHE n` (donde  $n$  representa el número de cambios realizados en el archivo de datos activo antes de crear una caché del archivo).
- ▶ En los menús de la ventana de sintaxis, elija:
  - Ejecutar
  - Todos

*Nota:* El ajuste de la caché no se almacena entre sesiones. Cada vez que se inicia una nueva sesión, se toma el valor por defecto de la opción que es 20.

## ***Análisis en modo distribuido***

El análisis en modo distribuido permite utilizar un ordenador que no es el local (o de escritorio) para realizar trabajos que requieren un gran consumo de memoria. Debido a que los servidores remotos utilizados para análisis distribuidos son normalmente más potentes y rápidos que los ordenadores locales, un análisis en modo distribuido puede reducir significativamente el tiempo de procesamiento del ordenador. El análisis distribuido con un servidor remoto puede ser útil si el trabajo trata:

- Archivos de datos, en particular lecturas de datos de fuentes de bases de datos.
- Tareas que requieren un gran consumo de memoria. Cualquier tarea que tarde bastante tiempo en el análisis en modo local será una buena candidata para el análisis distribuido,

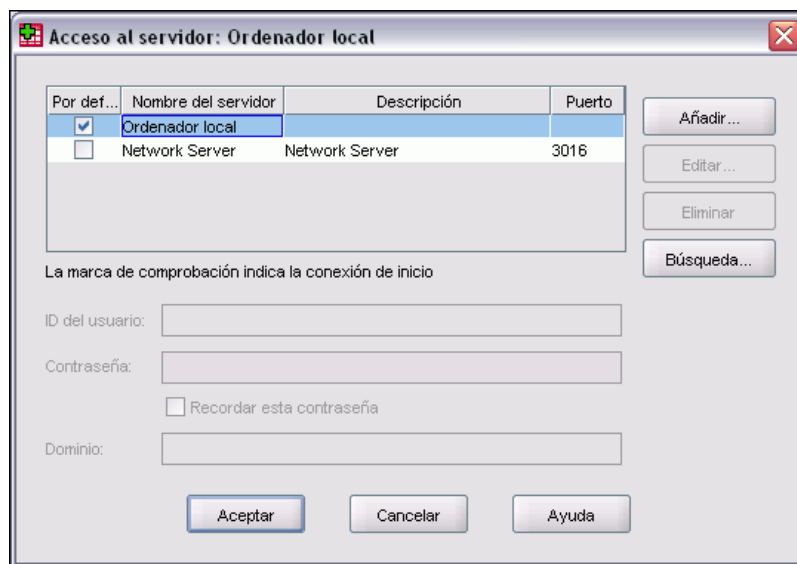
El análisis distribuido sólo afecta a las tareas relacionadas con los datos, como lectura de datos, transformación de datos, cálculo de nuevas variables y cálculo de estadísticos. El análisis distribuido no tiene ningún efecto sobre tareas relacionadas con la edición de resultados, como la manipulación de tablas pivote o la modificación de gráficos.

*Nota:* el análisis distribuido está disponible sólo si ambos equipos tienen una versión local y acceso a una versión de servidor con licencia del software instalado en un servidor remoto. Además, sólo está disponible para los ordenadores locales que ejecutan Windows. No está disponible para Mac y Linux.

### ***Acceso al servidor***

El cuadro de diálogo Acceso al servidor permite seleccionar el ordenador para procesar comandos y ejecutar procedimientos. Puede seleccionar el ordenador local o un servidor remoto.

Figura 4-1  
Cuadro de diálogo Acceso al servidor



Se pueden añadir, modificar o eliminar servidores remotos de la lista. Los servidores remotos requieren normalmente un ID de usuario y una contraseña; también puede ser necesario un nombre de dominio. Póngase en contacto con el administrador del sistema para obtener información acerca de servidores, ID de usuario y contraseñas, nombres de dominio disponibles y demás información necesaria para la conexión.

Puede seleccionar un servidor por defecto y guardar el ID de usuario, nombre de dominio y contraseña asociados a cualquier servidor. De esta manera, se conectará de forma automática al servidor por defecto en el momento de iniciar la sesión.

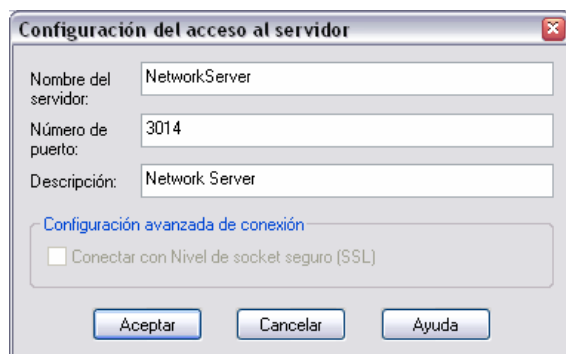
Si tiene licencia para utilizar SPSS Adaptor para Predictive Enterprise Services y tiene instalado SPSS Predictive Enterprise Services 3.5 o posterior, puede pulsar en Buscar... para ver una lista de los servidores disponibles en su red. Si no ha iniciado sesión en Predictive Enterprise Repository, se le solicitará que introduzca la información de conexión antes de poder ver la lista de servidores.

### **Adición y edición de la configuración de acceso al servidor**

Utilice el cuadro de diálogo Configuración del acceso al servidor para añadir o editar la información de conexión para servidores remotos para utilizar en los análisis en modo distribuido.



Figura 4-2  
Cuadro de diálogo Configuración del acceso al servidor



Para obtener una lista de servidores disponibles, los números de puerto para dichos servidores y toda la información adicional necesaria para la conexión, póngase en contacto con el administrador del sistema. No utilice el Nivel de socket seguro a menos que lo indique el administrador.

**Nombre del servidor.** Un “nombre” de servidor puede ser un nombre alfanumérico asignado a un ordenador (por ejemplo, ServidorRed) o una dirección IP única asignada a un ordenador (por ejemplo, 202.123.456.78).

**Número de puerto.** El número de puerto es el puerto que el software del servidor utiliza para las comunicaciones.

**Descripción.** Puede introducir una descripción opcional para que se visualice en la lista de servidores.

**Conectar con Nivel de socket seguro.** Las encriptaciones de Nivel de socket seguro (SSL) requieren el análisis distribuido cuando se envían al servidor remoto de SPSS. Antes de utilizar el SSL, consulte con el administrador. Para que esta opción se active, SSL debe estar configurado en su equipo de escritorio y en el servidor.

### ***Para seleccionar, cambiar o añadir servidores***

- ▶ Elija en los menús:  
Archivo  
Cambiar servidor...

#### **Para seleccionar un servidor por defecto:**

- ▶ En la lista de servidores, seleccione la casilla que se encuentra junto al servidor que desea utilizar.
- ▶ Introduzca el ID de usuario, nombre de dominio y contraseña suministrados por el administrador.

*Nota:* de esta manera, se conectará de forma automática al servidor por defecto en el momento de iniciar la sesión.

#### **Para cambiar a otro servidor:**

- ▶ Seleccione el servidor de la lista.
- ▶ Introduzca el ID de usuario, nombre de dominio y contraseña (si es necesario).

*Nota:* al cambiar de servidor durante una sesión, se cierran todas las ventanas abiertas. Se solicitará guardar los cambios antes de que se cierren las ventanas.

**Para añadir un servidor:**

- ▶ Solicite al administrador la información de conexión del servidor.
- ▶ Pulse en Añadir para abrir el cuadro de diálogo Configuración del acceso al servidor.
- ▶ Introduzca la información de conexión y la configuración opcional y pulse en Aceptar.

**Para editar un servidor:**

- ▶ Solicite al administrador la información de conexión revisada.
- ▶ Pulse en Editar para abrir el cuadro de diálogo Configuración del acceso al servidor.
- ▶ Introduzca los cambios y pulse en Aceptar.

**Para buscar servidores disponibles:**

*Nota:* la capacidad para buscar servidores disponibles sólo está disponible si tiene licencia para utilizar SPSS Adaptor para Enterprise Services y en su empresa se ejecuta SPSS Predictive Enterprise Services 3.5 o posterior.

- ▶ Pulse en Buscar... para abrir el cuadro de diálogo Buscar servidores. Si no ha iniciado sesión en Predictive Enterprise Repository, se le solicitará la información de conexión.
- ▶ Seleccione uno o varios servidores disponibles y pulse en Aceptar. Los servidores se mostrarán en el cuadro de diálogo Acceso al servidor.
- ▶ Para conectar con uno de los servidores siga las indicaciones para cambiar a otro servidor.

## ***Búsqueda de servidores disponibles***

Utilice el cuadro de diálogo Buscar servidores para elegir uno o varios servidores disponibles en la red. Este cuadro de diálogo aparece al pulsar en Buscar... en el cuadro de diálogo Acceso al servidor.

Figura 4-3  
Cuadro de diálogo Buscar servidores



Seleccione uno o más servidores y pulse en Aceptar para añadirlos al cuadro de diálogo Acceso al servidor. Aunque es posible añadir servidores manualmente al cuadro de diálogo Acceso al servidor, la opción de búsqueda de servidores disponibles permite conectar con los servidores sin necesidad de conocer el nombre correcto y número de puerto del servidor. Esta información se proporciona automáticamente. No obstante, deberá disponer de la información de inicio de sesión correcta, como nombre de usuario, dominio y contraseña.

## ***Apertura de archivos de datos desde un servidor remoto***

En el análisis en modo distribuido, el cuadro de diálogo Abrir archivo remoto sustituye al cuadro de diálogo estándar Abrir archivo.

- El contenido de la lista de archivos, carpetas y unidades muestra lo que hay disponible en o desde el servidor remoto. En la parte superior del cuadro de diálogo se indica el nombre del servidor actual.
- En el análisis en modo distribuido, sólo tendrá acceso a los archivos del equipo local si especifica la unidad como un dispositivo compartido y las carpetas que contienen los archivos de datos como carpetas compartidas. Consulte la documentación de su sistema operativo para obtener información sobre cómo “compartir” carpetas del equipo local con la red del servidor.
- Si el servidor está ejecutando un sistema operativo diferente (por ejemplo, usted dispone de Windows y el servidor se ejecuta bajo UNIX), probablemente no dispondrá de acceso a los archivos de datos locales en el análisis en modo distribuido, aunque los archivos estén en carpetas compartidas.

## ***Acceso a archivo en análisis en modo local y distribuido***

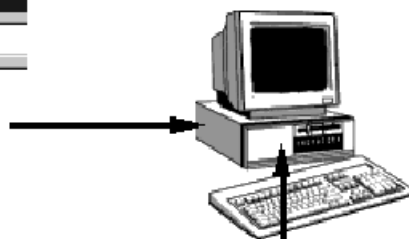
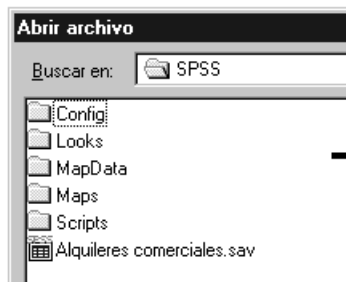
La presentación de los datos, las carpetas (directorios) y las unidades para el ordenador local y la red está en función del ordenador que está utilizando para procesar comandos y ejecutar procedimientos, que no es necesariamente el ordenador que tiene delante.

**Análisis en modo local.** Cuando utiliza el ordenador local como el “servidor”, la visualización de los archivos de datos, las carpetas y las unidades que ve en el cuadro de diálogo de acceso a los archivos (para la apertura de archivos de datos) es similar a lo que ve en otras aplicaciones o en el Explorador de Windows. Se pueden ver todos los archivos de datos y las carpetas en el ordenador y cualquier archivo y carpeta en las unidades de red.

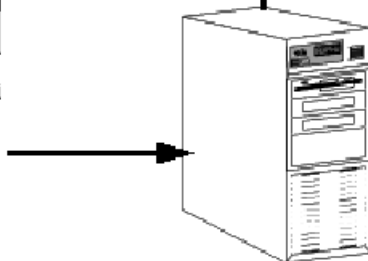
**Análisis en modo distribuido.** Cuando utiliza otro ordenador como “servidor remoto” para ejecutar comandos y procedimientos, la visualización de los archivos de datos y las unidades representa la vista desde el servidor remoto. Aunque vea nombres de carpetas que le son familiares (como *Archivos de programas* y unidades como *C*), estas *no* son las carpetas y unidades del ordenador local, sino las del servidor remoto.

Figura 4-4  
Vistas de los equipos local y remoto

### Vista del sistema local



### Vista del sistema remoto



En el análisis en modo distribuido, sólo tendrá acceso a los archivos de datos del equipo local si especifica la unidad como un dispositivo compartido y las carpetas que contienen los archivos de datos como carpetas compartidas. Si el servidor está ejecutando un sistema operativo diferente (por ejemplo, usted dispone de Windows y el servidor se ejecuta bajo UNIX), probablemente no dispondrá de acceso a los archivos de datos locales en el análisis en modo distribuido, aunque los archivos estén en carpetas compartidas.

El análisis en modo distribuido no es lo mismo que acceder a archivos de datos que se encuentran en otro ordenador de la red. Se puede acceder a archivos de datos en otros dispositivos de red tanto en análisis en modo local como en análisis en modo distribuido. En modo local, se accede a otros dispositivos desde el ordenador local. En el modo distribuido, se accede a otros dispositivos de red desde el servidor remoto.

Si no está seguro de si está utilizando el análisis en modo local o distribuido, mire la barra de título en el cuadro de diálogo para acceder a archivos de datos. Si el título del cuadro de diálogo contiene la palabra *remoto* (como en Abrir archivo remoto) o si el texto Servidor remoto: [nombre del servidor] aparece en la parte superior del cuadro de diálogo, significa que está en modo distribuido.

*Nota:* esta situación afecta sólo a los cuadros de diálogo para acceder a archivos de datos (por ejemplo, Abrir datos, Guardar datos, Abrir base de datos y Aplicar diccionario de datos). Para todos los demás tipos de archivos (por ejemplo, archivos del Visor, archivos de sintaxis y archivos de procesos) se utiliza la visualización local.

## **Disponibilidad de procedimientos en análisis en modo distribuido**

En el análisis en modo distribuido, estarán disponibles sólo aquellos procedimientos instalados en la versión local y en la versión del servidor remoto.

Si dispone de componentes opcionales instalados en el ordenador local que no están disponibles en el servidor remoto, y cambia del ordenador local a un servidor remoto, los procedimientos afectados se eliminarán de los menús y la sintaxis de comandos relacionada generará errores. Todos los procedimientos afectados se restaurarán al cambiar de nuevo al modo local.

## **Especificaciones de rutas absolutas frente a rutas relativas**

En el modo de análisis distribuido, las especificaciones de las rutas relativas para los archivos de datos y los archivos de sintaxis de comandos son relativas al servidor actual, no al equipo local. Una especificación de ruta de acceso como */misdocs/misdatos.sav* no indica un directorio y archivo en la unidad local, sino que indica un directorio y archivo en el disco duro del servidor remoto.

### **Especificaciones de ruta de acceso UNC para Windows**

Si utiliza la versión de servidor para Windows, puede usar las especificaciones de la UNC (convención de denominación universal) al acceder a los archivos de datos y sintaxis mediante la sintaxis de comandos. El formato general de una especificación UNC es:

```
\\nombre_servidor\nombre_compartido\ruta\nombre_archivo
```

- *Nombre\_servidor* es el nombre del ordenador que contiene el archivo de datos.
- *Nombre\_compartido* es la carpeta (directorio) en el ordenador que aparece designada como una carpeta compartida.
- *Ruta* es cualquier ruta de acceso de carpetas (subdirectorios) por debajo de la carpeta compartida.
- *Nombre\_archivo* es el nombre del archivo de datos.

A continuación se muestra un ejemplo:

```
GET FILE = '\\hqdev001\public\julio\ventas.sav'.
```

Si el ordenador no tiene un nombre asignado, puede utilizar su dirección IP, como en:

```
GET FILE = '\\204.125.125.53\public\julio\ventas.sav'.
```

Incluso con especificaciones de ruta de acceso UNC, sólo se puede acceder a archivos de datos y de sintaxis que estén en carpetas y dispositivos compartidos. Cuando se utiliza análisis en modo distribuido, esta situación incluye archivos de datos y de sintaxis del ordenador local.

### **Especificaciones de rutas absolutas para UNIX**

En las versiones de servidor para UNIX, no hay un equivalente a las rutas UNC y todas las rutas de acceso de los directorios deben ser rutas absolutas que comienzan en la raíz del servidor; las rutas relativas no están permitidas. Por ejemplo, si el archivo de datos está ubicado en */bin/spss/datos*

y el directorio actual también es */bin/spss/datos*, la sintaxis `GET FILE='ventas.sav'` no es válida; se debe especificar la ruta completa, como en:

```
GET FILE='/bin/spss/ventas.sav'. INSERT FILE='/bin/spss/trabajo_ventas.sps'.
```

## Editor de datos

El Editor de datos proporciona un método práctico (al estilo de las hojas de cálculo) para la creación y edición de archivos de datos. La ventana Editor de datos se abre automáticamente cuando se inicia una sesión.

El Editor de datos proporciona dos vistas de los datos.

- **Vista de datos.** Esta vista muestra los valores de datos reales o las etiquetas de valor definidas.
- **Vista de variables.** Esta vista muestra la información de definición de las variables, que incluye las etiquetas de la variable definida y de valor, tipo de dato (por ejemplo, cadena, fecha o numérico), nivel de medida (nominal, ordinal o de escala) y los valores perdidos definidos por el usuario.

En ambas vistas, se puede añadir, modificar y eliminar la información contenida en el archivo de datos.

### Vista de datos

Figura 5-1  
Vista de datos

	edad	marital	direcc	ingres	ingcat	coche
1	55	Casado	12	72,00	\$50 - \$74	36,
2	56	Sin casar	29	153,00	\$75+	76,
3	28	Casado	9	28,00	\$25 - \$49	13,
4	24	Casado	4	26,00	\$25 - \$49	12,
5	25	Sin casar	2	23,00	menos de	11,
6	45	Casado	9	76,00	\$75+	37,
7	42	Sin casar	19	40,00	\$25 - \$49	19,
8	35	Sin casar	15	57,00	\$50 - \$74	28,
9	46	Sin casar	26	24,00	menos de	12,
10	34	Casado	0	89,00	\$75+	46,
11	55	Casado	17	72,00	\$50 - \$74	35,

Muchas de las funciones de la Vista de datos son similares a las que se encuentran en aplicaciones de hojas de cálculo. Sin embargo, existen varias diferencias importantes:

- Las filas son casos. Cada fila representa un caso o una observación. Por ejemplo, cada individuo que responde a un cuestionario es un caso.
- Las columnas son variables. Cada columna representa una variable o una característica que se mide. Por ejemplo, cada elemento en un cuestionario es una variable.

- Las casillas contienen valores. Cada casilla contiene un valor único de una variable para cada caso. La casilla se encuentra en la intersección del caso y la variable. Las casillas sólo contienen valores de datos. A diferencia de los programas de hoja de cálculo, las casillas del Editor de datos no pueden contener fórmulas.
- El archivo de datos es rectangular. Las dimensiones del archivo de datos vienen determinadas por el número de casos y de variables. Se pueden introducir datos en cualquier casilla. Si introduce datos en una casilla fuera de los límites del archivo de datos definido, el rectángulo de datos se ampliará para incluir todas las filas y columnas situadas entre esa casilla y los límites del archivo. No hay casillas “vacías” en los límites del archivo de datos. Para variables numéricas, las casillas vacías se convierten en el valor perdido del sistema. Para variables de cadena, un espacio en blanco se considera un valor válido.

## Vista de variables

Figura 5-2  
Vista de variables

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores
1	edad	Numérico	4	0	Edad en años	Ninguna
2	marital	Numérico	4	0	Estado civil	{0, Sin casa...
3	direcc	Numérico	4	0	Años en la dire...	Ninguna
4	ingres	Numérico	8	2	Ingresos del ho...	Ninguna
5	ingcat	Numérico	8	2	Categoría de in...	{1.00, Meno...
6	coche	Numérico	8	2	Precio del coch...	Ninguna
7	cochecat	Numérico	8	2	Categoría del pr...	{1.00, Econ...
8	educ	Numérico	4	0	Nivel educativo	{1, No comp...
9	empleo	Numérico	4	0	Años con la em...	Ninguna
10	retirado	Numérico	4	0	Retirado	{0, No}...
11	empcat	Numérico	4	0	Años con la em...	{1, Menos d...
12	satlab	Numérico	4	0	Satisfacción la...	{1, Muy ins...
13	genero	Cadena	2	0	Género	{h, Hombre}...
14	residen	Numérico	4	0	Número de per...	Ninguna
15	inalam	Numérico	4	0	Servicio inalám...	{0, Sí}...
16	multline	Numérico	4	0	Múltiples líneas	{0, Sí}...

La Vista de variables contiene descripciones de los atributos de cada variable del archivo de datos. En la Vista de variables:

- Las filas son variables.
- Las columnas son atributos de las variables.



Se pueden añadir o eliminar variables, y modificar los atributos de las variables, incluidos los siguientes:

- Nombre de variable
- Tipo de datos
- Número de dígitos o caracteres
- Número de decimales
- Las etiquetas descriptivas de variable y de valor
- Valores perdidos definidos por el usuario
- Ancho de columna.
- Nivel de medida

Todos estos atributos se guardan al guardar el archivo de datos.

Además de la definición de propiedades de variables en la Vista de variables, hay dos otros métodos para definir las propiedades de variables:

- El Asistente para la copia de propiedades de datos ofrece la posibilidad de utilizar un archivo de datos de SPSS externo u otro conjunto de datos que esté disponible en la sesión actual como plantilla para definir las propiedades del archivo y las variables del conjunto de datos activo. También puede utilizar variables del conjunto de datos activo como plantillas para otras variables del conjunto de datos activo. La opción Copiar propiedades de datos está disponible en el menú Datos en la ventana Editor de datos.
- La opción Definir propiedades de variables (también disponible en el menú Datos de la ventana Editor de datos) explora los datos y muestra una lista con todos los valores de datos únicos para las variables seleccionadas, indica los valores sin etiquetas y ofrece una función de etiquetas automáticas. Este método es especialmente útil para las variables categóricas que utilizan códigos numéricos para representar las categorías (por ejemplo, 0 = *hombre*, 1 = *mujer*).

### ***Para visualizar o definir los atributos de las variables***

- ▶ Haga que el editor de datos sea la ventana activa.
- ▶ Pulse dos veces en un nombre de variable en la parte superior de la columna en la Vista de datos o bien pulse en la pestaña Vista de variables.
- ▶ Para definir variables nuevas, introduzca un nombre de variable en cualquier fila vacía.
- ▶ Seleccione los atributos que desea definir o modificar.

## **Nombres de variable**

Para los nombres de variable se aplican las siguientes normas:

- Cada nombre de variable debe ser único; no se permiten duplicados.
- Los nombres de variable pueden tener una longitud de hasta 64 bytes y el primer carácter debe ser una letra o uno de estos caracteres: @, # o \$. Los caracteres posteriores puede ser cualquier combinación de letras, números, caracteres que no sean signos de puntuación y un punto (.). En el modo de página de código, sesenta y cuatro bytes suelen equivaler a 64 caracteres en idiomas de un solo byte (por ejemplo, inglés, francés, alemán, español, italiano, hebreo, ruso, griego, árabe y tailandés) y 32 caracteres en los idiomas de dos bytes (por ejemplo, japonés, chino y coreano). Muchos caracteres de una cadena ocuparán un solo byte en el modo de página de código y dos o más bytes en el modo Unicode. Por ejemplo, ocupa un byte en el formato de página de código pero dos bytes en el formato Unicode; por lo que *rsum* ocupa seis bytes en un archivo de página de código y ocho bytes en modo Unicode.

*Nota:* las letras incluyen todos los caracteres que no son signos de puntuación y se utilizan al escribir palabras normales en los idiomas admitidos en el juego de caracteres de la plataforma.

- Las variables no pueden contener espacios.
- Un carácter # en la primera posición de un nombre de variable define una variable transitorio. Sólo puede crear variables transitorios mediante la sintaxis de comandos. No puede especificar un # como primer carácter de una variable en los cuadros de diálogo que permiten crear nuevas variables.
- Un signo \$ en la primera posición indica que la variable es una variable del sistema. El signo \$ no se admite como carácter inicial de una variable definida por el usuario.
- El punto, el subrayado y los caracteres \$, # y @ se pueden utilizar dentro de los nombres de variable. Por ejemplo, *A\_#\$@#I* es un nombre de variable válido.
- Se deben evitar los nombres de variable que terminan con un punto, ya que el punto puede interpretarse como un terminador del comando. Sólo puede crear variables que finalicen con un punto en la sintaxis de comandos. No puede crear variables que terminen con un punto en los cuadros de diálogo que permiten crear nuevas variables.
- Se deben evitar los nombres de variable que terminan con un carácter de subrayado, ya que tales nombres puede entrar en conflicto con los nombres de variable creados automáticamente por comandos y procedimientos.
- Las palabras reservadas no se pueden utilizar como nombres de variable. Las palabras reservadas son ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO y WITH.
- Los nombres de variable se pueden definir combinando de cualquier manera caracteres en mayúsculas y en minúsculas, esta distinción entre mayúsculas y minúsculas se conserva en lo que se refiere a la presentación.
- Cuando es necesario dividir los nombres largos de variable en varias líneas en los resultados, se dividen las líneas aprovechando los subrayados, los puntos y los cambios de minúsculas a mayúsculas.

## Nivel de medida de variable

Puede especificar el nivel de medida como Escala (datos numéricos de una escala de intervalo o de razón), Ordinal o Nominal. Los datos nominales y ordinales pueden ser de cadena (alfanuméricos) o numéricos.

- **Nominal.** Una variable puede ser tratada como nominal cuando sus valores representan categorías que no obedecen a una ordenación intrínseca. Por ejemplo, el departamento de la compañía en el que trabaja un empleado. Son ejemplos de variables nominales: la región, el código postal y la confesión religiosa.
- **Ordinal.** Una variable puede ser tratada como ordinal cuando sus valores representan categorías con alguna ordenación intrínseca. Por ejemplo, los niveles de satisfacción con un servicio, que abarquen desde muy insatisfecho hasta muy satisfecho. Son ejemplos de variables ordinales: las puntuaciones de actitud que representan el nivel de satisfacción o confianza y las puntuaciones de evaluación de la preferencia.
- **Escala.** Una variable puede ser tratada como de escala cuando sus valores representan categorías ordenadas con una métrica con significado, por lo que son adecuadas las comparaciones de distancia entre valores. Son ejemplos de variables de escala: la edad en años y los ingresos en dólares.

*Nota:* para variables de cadena ordinales, se supone que el orden alfabético de los valores de cadena indica el orden correcto de las categorías. Por ejemplo, en una variable de cadena cuyos valores sean *bajo*, *medio*, *alto*, se interpreta el orden de las categorías como *alto*, *bajo*, *medio* (orden que no es el correcto). Por norma general, se puede indicar que es más fiable utilizar códigos numéricos para representar datos ordinales.

A las nuevas variables numéricas creadas durante una sesión se les asigna un nivel de medida de escala. Para los datos leídos de los formatos de archivo externos y los archivos de datos SPSS creados antes de la versión 8.0, la asignación del nivel de medida por defecto se basa en las siguientes reglas:

- Las variables numéricas con menos de 24 valores únicos y las variables de cadena se establecen en nominales.
- Las variables numéricas con 24 o más valores únicos se establecen en variables de escala.

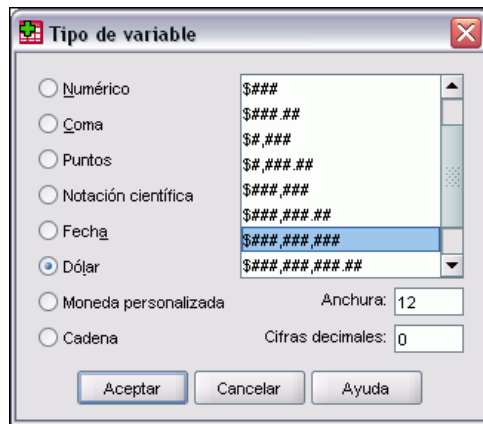
Se puede cambiar el valor de corte de escala/nominal para las variables numéricas en el cuadro de diálogo Opciones. Si desea obtener más información, consulte [Datos: Opciones](#) en Capítulo 45 en p. 482.

El cuadro de diálogo Definir propiedades de variables, disponible en el menú Datos, puede ayudarle a asignar el nivel de medida correcto. Si desea obtener más información, consulte [Asignación del nivel de medida](#) en Capítulo 7 en p. 111.

## Tipo de variable

Tipo de variable especifica los tipos de datos de cada variable. Por defecto se asume que todas las variables nuevas son numéricas. Se puede utilizar Tipo de variable para cambiar el tipo de datos. El contenido del cuadro de diálogo Tipo de variable depende del tipo de datos seleccionado. Para algunos tipos de datos, hay cuadros de texto para la anchura y el número de decimales; para otros tipos de datos, simplemente puede seleccionar un formato de una lista desplegable de ejemplos.

Figura 5-3  
Cuadro de diálogo Tipo de variable



Los tipos de datos disponibles son los siguientes:

**Numérico.** Una variable cuyos valores son números. Los valores se muestran en formato numérico estándar. El Editor de datos acepta valores numéricos en formato estándar o en notación científica.

**Coma.** Una variable numérica cuyos valores se muestran con comas que delimitan cada tres posiciones y con el punto como delimitador decimal. El Editor de datos acepta valores numéricos para este tipo de variables con o sin comas, o bien en notación científica. Los valores no pueden contener comas a la derecha del indicador decimal.

**Punto.** Una variable numérica cuyos valores se muestran con puntos que delimitan cada tres posiciones y con la coma como delimitador decimal. El Editor de datos acepta valores numéricos para este tipo de variables con o sin puntos, o bien en notación científica. Los valores no pueden contener puntos a la derecha del indicador decimal.

**Notación científica.** Una variable numérica cuyos valores se muestran con una E intercalada y un exponente con signo que representa una potencia de base 10. El Editor de datos acepta para estas variables valores numéricos con o sin el exponente. El exponente puede aparecer precedido por una E o una D con un signo opcional, o bien sólo por el signo (por ejemplo, 123, 1,23E2, 1,23D2, 1,23E+2 y 1,23+2).

**Fecha.** Una variable numérica cuyos valores se muestran en uno de los diferentes formatos de fecha-calendario u hora-reloj. Seleccione un formato de la lista. Puede introducir las fechas utilizando como delimitadores: barras, guiones, puntos, comas o espacios. El rango de siglo para los valores de año de dos dígitos está determinado por la configuración de las opciones (en el menú Edición, seleccione Opciones y, a continuación, pulse en la pestaña Datos).

**Dólar.** Una variable numérica que se muestra con un signo dólar inicial (\$), comas que delimitan cada tres posiciones y un punto como delimitador decimal. Se pueden introducir valores de datos con o sin el signo dólar inicial.

**Moneda personalizada.** Una variable numérica cuyos valores se muestran en uno de los formatos de moneda personalizados que se hayan definido previamente en la pestaña Moneda del cuadro de diálogo Opciones. Los caracteres definidos en la moneda personalizada no se pueden emplear en la introducción de datos pero sí se mostrarán en el Editor de datos.

**Cadena.** Una variable cuyos valores no son numéricos y, por lo tanto, no se utilizan en los cálculos. Los valores pueden contener cualquier carácter siempre que no se exceda la longitud definida. Las mayúsculas y las minúsculas se consideran diferentes. Este tipo también se conoce como variable alfanumérica.

### **Para definir el tipo de variable**

- ▶ Pulse en el botón de la casilla *Tipo* de la variable que se quiere definir.
- ▶ Seleccione el tipo de datos en el cuadro de diálogo Tipo de variable.
- ▶ Pulse en Aceptar.

### **Formatos de entrada frente a formatos de presentación**

Dependiendo del formato, la presentación de valores en la Vista de datos puede ser diferente del valor real que se ha introducido y almacenado internamente. A continuación, se proporcionan algunas normas generales:

- Para formatos numéricos, de coma y de punto, se pueden introducir valores con cualquier número de dígitos decimales (hasta 16) y el valor completo se almacena internamente. La Vista de datos muestra sólo el número definido de dígitos decimales y redondea los valores con más decimales. Sin embargo, el valor completo se utiliza en todos los cálculos.
- Para las variables de cadena, todos los valores se rellenan por la derecha hasta el ancho máximo. Para una variable de cadena con un ancho de tres, un valor de *No* se almacena internamente como 'No ' y no es equivalente a ' No '.
- Para formatos de fecha, se pueden utilizar guiones, barras, espacios, comas o puntos como separadores entre valores de día, mes y año; se pueden introducir números, abreviaciones de tres letras o nombres completos para el valor de mes. Las fechas del formato general *dd-mmm-aa* aparecen separadas por guiones y con abreviaciones de tres letras para el mes. Las fechas del formato general *dd/mm/aa* y *mm/dd/aa* se muestran con barras como separadores y números para el mes. Internamente, las fechas se almacenan como el número de segundos transcurridos desde el 14 de octubre de 1582. El rango de siglo para los valores de año de dos dígitos está determinado por la configuración de las Opciones (en el menú Edición, seleccione Opciones y, a continuación, pulse en la pestaña Datos).
- Para formatos de hora, se pueden utilizar dos puntos, puntos o espacios como separadores entre horas, minutos y segundos. Las horas se muestran separadas por dos puntos. Internamente, las horas se almacenan como el número de segundos que representa un intervalo de tiempo. Por ejemplo, 10:00:00 se almacena internamente como 36000, que es 60 (segundos por minuto) x 60 (minutos por hora) x 10 (horas).

### **Etiquetas de variable**

Puede asignar etiquetas de variable descriptivas de hasta 256 caracteres de longitud (128 caracteres en los idiomas de doble byte). Las etiquetas de variable pueden contener espacios y caracteres reservados que no se admiten en los nombres de variable.

### **Para especificar etiquetas de variable**

- ▶ Haga que el editor de datos sea la ventana activa.
- ▶ Pulse dos veces en un nombre de variable en la parte superior de la columna en la Vista de datos o bien pulse en la pestaña Vista de variables.
- ▶ Escriba la etiqueta de variable descriptiva en la casilla *Etiqueta* de la variable.

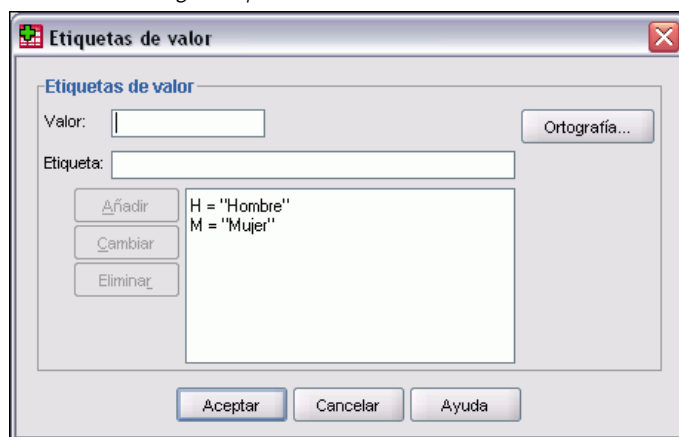
### **Etiquetas de valor**

Puede asignar etiquetas de valor descriptivas a cada valor de una variable. Este proceso es especialmente útil si el archivo de datos utiliza códigos numéricos para representar categorías que no son numéricas (por ejemplo, códigos 1 y 2 para *hombre* y *mujer*).

- Las etiquetas de valor pueden ocupar hasta 120 bytes.

Figura 5-4

Cuadro de diálogo *Etiquetas de valor*



### **Para especificar etiquetas de valor**

- ▶ Pulse en el botón de la casilla *Valores* de la variable que se quiere definir.
- ▶ Para cada valor, escriba el valor y una etiqueta.
- ▶ Pulse en Añadir para introducir la etiqueta de valor.
- ▶ Pulse en Aceptar.

### **Inserción de saltos de línea en etiquetas**

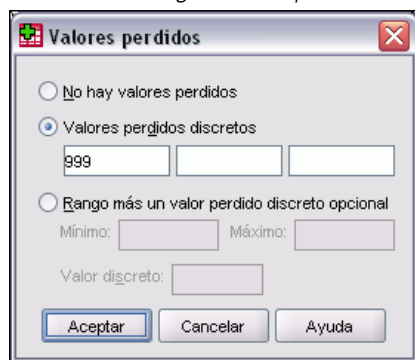
Las etiquetas de valor y las de variable se dividen automáticamente en varias líneas en los gráficos y en las tablas pivote si el ancho de casilla o el área no es suficiente para mostrar la etiqueta entera en una línea. Se pueden editar los resultados para insertar saltos de línea manuales si se quiere dividir la etiqueta en un punto diferente. También puede crear etiquetas de variable y de valor que *siempre* se dividan en puntos especificados y se muestren en varias líneas.

- ▶ Para etiquetas de variable, seleccione la casilla *Etiqueta* de la variable en la Vista de variables del Editor de datos.
- ▶ Para etiquetas de valor, seleccione la casilla *Valores* correspondiente a la variable en la Vista de variables del Editor de datos, pulse en el botón de la casilla y, a continuación, seleccione la etiqueta que desea modificar en el cuadro de diálogo Etiquetas de valor.
- ▶ En el punto de la etiqueta en el que desea dividir la etiqueta, escriba \n.  
El \n no aparece en las tablas pivote ni en los gráficos; se interpreta como un carácter de salto de línea.

## Valores perdidos

Valores perdidos define los valores de los datos definidos como **perdidos por el usuario**. Por ejemplo, es posible que quiera distinguir los datos perdidos porque un encuestado se niegue a responder de los datos perdidos porque la pregunta no afecta a dicho encuestado. Los valores de datos que se especifican como perdidos por el usuario aparecen marcados para un tratamiento especial y se excluyen de la mayoría de los cálculos.

Figura 5-5  
Cuadro de diálogo Valores perdidos



- Se pueden introducir hasta tres valores perdidos (individuales) de tipo discreto, un rango de valores perdidos o un rango más un valor de tipo discreto.
- Sólo pueden especificarse rangos para las variables numéricas.
- Se considera que son válidos todos los valores de cadena, incluidos los valores vacíos o nulos, a no ser que se definan explícitamente como perdidos.
- Los valores perdidos de las variables de cadena no pueden tener más de ocho bytes. (No hay ningún límite respecto al ancho definido de la variable de cadena, pero los valores definidos como perdidos no pueden tener más de ocho bytes.)
- Para definir como perdidos los valores nulos o vacíos de una variable de cadena, escriba un espacio en blanco en uno de los campos debajo de la selección Valores perdidos discretos.

### Para definir los valores perdidos

- ▶ Pulse en el botón de la casilla *Perdido* de la variable que se quiere definir.

- ▶ Introduzca los valores o el rango de valores que representen los datos perdidos.

### ***Ancho de columna***

Se puede especificar un número de caracteres para el ancho de la columna. Los anchos de columna también se pueden cambiar en la Vista de datos pulsando y arrastrando los bordes de las columnas.

Los formatos de columna afectan sólo a la presentación de valores en el Editor de datos. Al cambiar el ancho de columna no se cambia el ancho definido de una variable.

### ***Alineación de la variable***

La alineación controla la presentación de los valores de los datos y/o de las etiquetas de valor en la Vista de datos. La alineación por defecto es a la derecha para las variables numéricas y a la izquierda para las variables de cadena. Este ajuste sólo afecta a la presentación en la Vista de datos.

### ***Aplicación de atributos de definición de variables a varias variables***

Tras definir los atributos de definición de variables correspondientes a una variable, puede copiar uno o más atributos y aplicarlos a una o más variables.

Se utilizan las operaciones básicas de copiar y pegar para aplicar atributos de definición de variables. Tiene la posibilidad de:

- Copiar un único atributo (por ejemplo, etiquetas de valor) y pegarlo en la misma casilla de atributo para una o más variables.
- Copiar todos los atributos de una variable y pegarlos en una o más variables.
- Crear varias variables nuevas con todos los atributos de una variable copiada.

### ***Aplicación de atributos de definición de variables a otras variables***

#### ***Para aplicar atributos individuales de una variable definida***

- ▶ En Vista de variables, seleccione la casilla de atributos que quiere aplicar a otras variables.
- ▶ Elija en los menús:
  - Edición
  - Copiar
- ▶ Seleccione la casilla de atributos a la que quiere aplicar el atributo. (Puede seleccionar varias variables de destino.)



- ▶ Elija en los menús:

- Edición
  - Pegar

Si pega el atributo en filas vacías, se crean nuevas variables con atributos por defecto para todos los atributos excepto para el seleccionado.

#### ***Para aplicar todos los atributos de una variable definida***

- ▶ En Vista de variables, seleccione el número de fila para la variable con los atributos que quiere utilizar. (Se resaltará la fila entera.)
- ▶ Elija en los menús:
  - Edición
  - Copiar
- ▶ Seleccione los números de fila de las variables a la que desea aplicar los atributos. (Puede seleccionar varias variables de destino.)
- ▶ Elija en los menús:
  - Edición
  - Pegar

#### ***Generación de varias variables nuevas con los mismos atributos***

- ▶ En la Vista de variables, pulse en el número de fila de la variable que tiene los atributos que quiere utilizar para la nueva variable. (Se resaltará la fila entera.)
- ▶ Elija en los menús:
  - Edición
  - Copiar
- ▶ Pulse en el número de la fila vacía situada bajo la última variable definida en el archivo de datos.
- ▶ Elija en los menús:
  - Edición
  - Pegar variables...
- ▶ En el cuadro de diálogo Pegar variables, escriba el número de variables que desea crear.
- ▶ Introduzca un prefijo y un número inicial para las nuevas variables.
- ▶ Pulse en Aceptar.

Los nombres de las nuevas variables se compondrán del prefijo especificado, más un número secuencial que comienza por el número indicado.

#### ***Atributos de variable personalizados***

Además de los atributos de variable estándar (como las etiquetas de valores, los valores perdidos y el nivel de medida), puede cerrar sus propios atributos de variable personalizados. Al igual que los atributos de variable estándar, estos atributos personalizados se guardan en los archivos de datos con formato SPSS. De esta forma, puede crear un atributo de variable que identifique el tipo

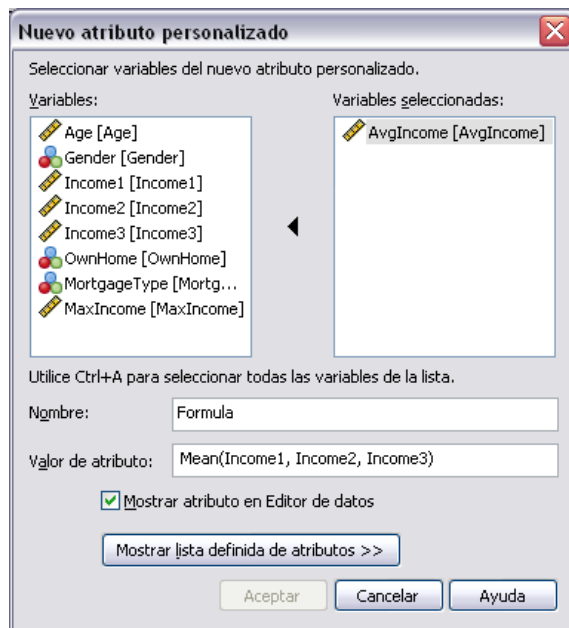
de respuesta para las preguntas de encuesta (por ejemplo, selección única, selección múltiple, rellenar) o las fórmulas empleadas para el cálculo de variables.

### **Creación de atributos de variable personalizados**

Para crear nuevos atributos personalizados:

- ▶ En la Vista de variables, elija en los menús:  
   Datos  
   Nuevo atributo personalizado...
- ▶ Arrastre las variables a las que desea asignar el nuevo atributo a la lista y colóquelas en la lista Variables seleccionadas.
- ▶ Escriba el nombre del atributo. Los nombres de atributo deben cumplir las mismas reglas que los nombres de variable. Si desea obtener más información, consulte [Nombres de variable](#) en p. 82.
- ▶ Introduzca un valor opcional para el atributo. Si selecciona varias variables, el valor se asignará a todas las variables seleccionadas. Puede dejar este campo en blanco y especificar valores para cada variable en la Vista de variables.

Figura 5-6  
 Cuadro de diálogo Nuevo atributo personalizado



**Mostrar atributo en Editor de datos.** Muestra el atributo en la Vista de variables del Editor de datos. Si desea obtener información acerca del control de la presentación de los atributos personalizados, consulte [Presentación y edición de los atributos de variable personalizados](#) a continuación.

**Mostrar lista definida de atributos.** Muestra una lista de atributos personalizados ya definidos para el conjunto de datos. Los nombres de atributo que comienzan con un signo de dólar (\$) son atributos reservados que no es posible modificar.

### Presentación y edición de los atributos de variable personalizados

Puede mostrar y editar los atributos de variable personalizados en la Vista de variables del Editor de datos.

Figura 5-7

Atributos de variables personalizados mostrados en la Vista de variables

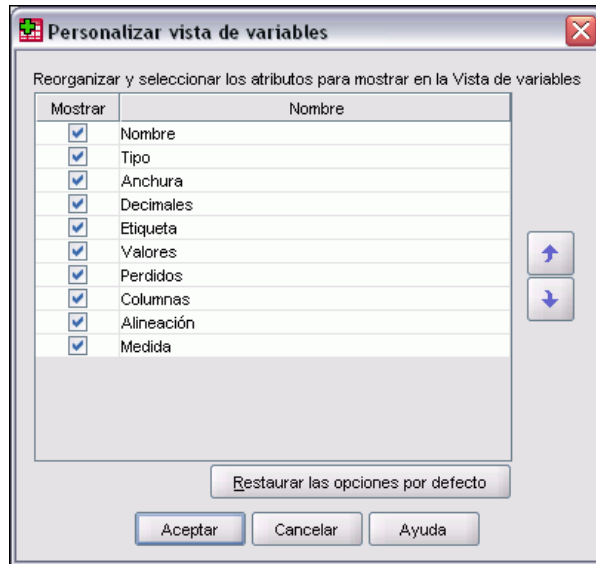
	[AnswerFormat]	[DerivedFrom]	[Formula]	[Notes]
1	Fill-in			Empty
2	Single select			Empty
3	Fill-in			Empty
4	Fill-in			Empty
5	Fill-in			Empty
6	Single select			Empty
7	Single select			Empty
8	Computed value	Array...	mean(Income1, Income2, Income3)	Empty
9	Computed value	Array...	max(Income1, Income2, Income3)	Empty
10				
11				
12				

- Los nombres de atributos de variable personalizados aparecen entre corchetes.
- Los nombres de atributo que comienzan con un signo de dólar son reservados y no se pueden modificar.
- Una casilla en blanco indica que el atributo no existe para la variable; el texto *Vacío* mostrado en una casilla indica que el atributo existe para dicha variable pero no se ha asignado ningún valor al atributo de dicha variable. Una vez que se escribe texto en la casilla, existe el atributo para dicha variable con el valor que ha introducido.
- Si aparece el texto *Matriz...*, en una casilla, indica que se trata de una matriz de atributos, un atributo que contiene varios valores. Pulse en el botón de la casilla para mostrar la lista de valores.

#### Para mostrar y editar atributos de variable personalizados

- ▶ En la Vista de variables, elija en los menús:
  - Ver
  - Personalizar Vista de variables...
- ▶ Seleccione (marque) los atributos de variable personalizados que desea mostrar. (Los atributos de variable personalizados son los que aparecen entre corchetes.)

Figura 5-8  
Personalizar Vista de variables

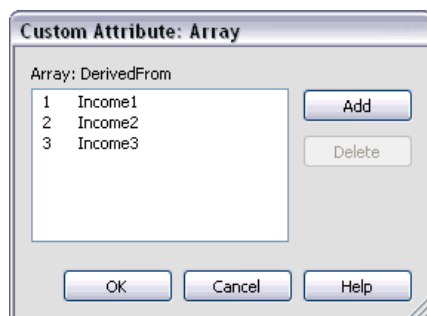


Una vez que los atributos aparecen en la Vista de variables, puede modificarlos directamente en el Editor de datos.

### **Matrices de atributos de variable**

El texto *Matriz...* que se muestra en una casilla para un atributo de variable personalizado en Vista de variables o en el cuadro de diálogo Atributos personalizados de variables en Definir propiedades de variables indica que es una **matriz de atributos**, un atributo que contiene varios valores. Por ejemplo, podría tener una matriz de atributos que identificara todas las variables origen para calcular una variable derivada. Pulse en el botón de la casilla para mostrar y editar la lista de valores.

Figura 5-9  
Cuadro de diálogo Atributo personalizado: Matriz



## Personalización de la Vista de variables

Puede utilizar la opción Personalizar Vista de variables para controlar qué atributos se muestran en la Vista de variables (por ejemplo, nombre, tipo, etiqueta) y el orden en el que aparecen.

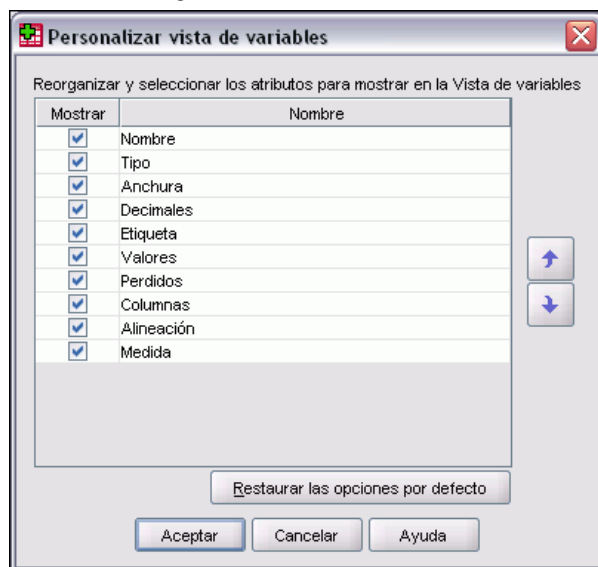
- Todos los atributos de variable personalizados asociados al conjunto de datos aparecen entre corchetes. Si desea obtener más información, consulte [Creación de atributos de variable personalizados](#) en p. 90.
- Los ajustes de presentación personalizados se guardan en los archivos de datos con formato SPSS.
- También puede controlar la presentación por defecto y el orden de los atributos en la Vista de variables. Si desea obtener más información, consulte [Cambio de la Vista de variables por defecto](#) en Capítulo 45 en p. 484.

### Para personalizar la Vista de variables

- ▶ En la Vista de variables, elija en los menús:  
Ver  
Personalizar Vista de variables...
- ▶ Seleccione (marque) los atributos de variable que desea mostrar.
- ▶ Utilice los botones de dirección hacia arriba y hacia abajo para cambiar el orden de la presentación de los atributos.

Figura 5-10

Cuadro de diálogo Personalizar vista de variables



**Restaurar las opciones por defecto.** Aplica los ajustes por defecto de presentación y orden.

## **Revisión ortográfica de las etiquetas de los valores y las variables**

Para revisar la ortografía de las etiquetas de los valores y las variables:

- ▶ En la Vista de variables, pulse con el botón derecho del ratón en la columna *Etiquetas* o *Valores* y elija en el menú contextual:  
Ortografía

*o*

- ▶ En la Vista de variables, elija en los menús:  
Utilidades  
Ortografía

*o*

- ▶ En el cuadro de diálogo *Etiquetas de valor*, pulse en *Ortografía*. (Con esto, la revisión ortográfica se limitará a las etiquetas de valor de una determinada variable.)

La revisión ortográfica se limita a las etiquetas de los valores y las variables de la Vista de variables del Editor de datos.

## **Introducción de datos**

En la Vista de datos, puede introducir datos directamente en el Editor de datos. Se puede introducir datos en cualquier orden. Asimismo, se pueden introducir datos por caso o por variable, para áreas seleccionadas o para casillas individuales.

- Se resaltará la casilla activa.
- El nombre de la variable y el número de fila de la casilla activa aparecen en la esquina superior izquierda del Editor de datos.
- Cuando seleccione una casilla e introduzca un valor de datos, el valor se muestra en el editor de casillas situado en la parte superior del Editor de datos.
- Los valores de datos no se registran hasta que se pulsa *Intro* o se selecciona otra casilla.
- Para introducir datos distintos de los numéricos, en primer lugar, se debe definir el tipo de variable.

Si introduce un valor en una columna vacía, el Editor de datos creará automáticamente una nueva variable y asignará un nombre de variable.

### **Para introducir datos numéricos**

- ▶ Seleccione una casilla en la Vista de datos.
- ▶ Introduzca el valor de los datos. (El valor se muestra en el editor de casillas situado en la parte superior del Editor de datos.)
- ▶ Para registrar el valor, pulse *Intro* o seleccione otra casilla.

### **Para introducir datos no numéricos**

- ▶ Pulse dos veces en un nombre de variable en la parte superior de la columna en la Vista de datos o bien pulse en la pestaña Vista de variables.
- ▶ Pulse en el botón de la casilla *Tipo* de la variable.
- ▶ Seleccione el tipo de datos en el cuadro de diálogo Tipo de variable.
- ▶ Pulse en Aceptar.
- ▶ Pulse dos veces en el número de fila o pulse en la pestaña Vista de datos.
- ▶ Introduzca en la columna los datos de la variable que se va a definir.

### **Para utilizar etiquetas de valor en la introducción de datos**

- ▶ Si las etiquetas de valor no aparecen en la Vista de datos, elija en los menús:  
Ver  
Etiquetas de valor
- ▶ Pulse la casilla en la que quiere introducir el valor.
- ▶ Elija una etiqueta de valor en la lista desplegable.  
De este modo se introducirá el valor y la etiqueta de valor se mostrará en la casilla.  
*Nota:* este proceso sólo funciona si ha definido etiquetas de valor para la variable.

### **Restricciones de los valores de datos en el Editor de datos**

El ancho y el tipo de variable definidos determinan el tipo de valor que se puede introducir en la casilla en la Vista de datos.

- Si escribe un carácter no permitido por el tipo de variable definido, no se introducirá dicho carácter.
- Para variables de cadena, no se permiten los caracteres que sobrepasen el ancho definido.
- Para variables numéricas, se pueden introducir valores enteros que excedan el ancho definido, pero el Editor de datos mostrará la notación científica o una parte del valor seguido por puntos suspensivos (...) para indicar que el valor es más ancho que el ancho definido. Para mostrar el valor de la casilla, cambie el ancho definido de la variable.

*Nota:* cambiar el ancho de la columna no afecta al ancho de la variable.

### **Edición de datos**

Con el Editor de datos es posible modificar un archivo de datos en Vista de datos de muchas maneras. Tiene la posibilidad de:

- Cambiar los valores de datos
- Cortar, copiar y pegar valores de datos
- Añadir y eliminar casos

- Añadir y eliminar variables
- Cambiar el orden de las variables

### ***Para reemplazar o modificar un valor de datos***

#### ***Para eliminar el valor anterior e introducir un valor nuevo***

- ▶ En la Vista de datos, pulse dos veces en la casilla. (Su valor aparecerá en el editor de casillas.)
- ▶ Edite el valor directamente en la casilla o en el editor de casillas.
- ▶ Pulse Intro o seleccione otra casilla para registrar el nuevo valor.

### ***Cortar, copiar y pegar valores de datos***

Puede cortar, copiar y pegar valores de casillas individuales o grupos de valores en el Editor de datos. Tiene la posibilidad de:

- Mover o copiar un único valor de casilla a otra casilla
- Mover o copiar un único valor de casilla a un grupo de casillas
- Mover o copiar los valores de un único caso (fila) a varios casos
- Mover o copiar los valores de una única variable (columna) a varias variables
- Mover o copiar un grupo de valores de casillas a otro grupo de casillas

### ***Conversión de datos para valores pegados en el Editor de datos***

Si los tipos de variable definidos de las casillas de origen y de destino no son iguales, el Editor de datos intentará convertir el valor. Si no es posible realizar la conversión, el valor perdido del sistema se insertará en la casilla de destino.

**Conversión de numérico o fecha a cadena.** Los formatos numéricos (por ejemplo, numérico, dólar, de punto o de coma) y de fechas se convierten en cadenas si se pegan en una casilla de variable de cadena. El valor de cadena es el valor numérico tal como se muestra en la casilla. Por ejemplo, para la variable con formato de dólar, el signo dólar que se muestra se convierte en parte del valor de cadena. Los valores que sobrepasan el ancho de la variable de cadena definida quedan cortados.

**Conversión de cadena a numérico o fecha.** Los valores de cadena que contienen caracteres admisibles por el formato numérico o de fecha de la casilla de destino se convierten al valor numérico o de fecha equivalente. Por ejemplo, un valor de cadena de 25/12/91 se convierte a una fecha válida si el tipo de formato de la casilla de destino es uno de los formatos día-mes-año, pero se convierte en perdido por el sistema si el tipo de formato de la casilla de destino es uno de los formatos mes-día-año.

**Conversión de fecha a numérico.** Los valores de fecha y hora se convierten a un número de segundos si la casilla de destino es uno de los formatos numéricos (por ejemplo, numérico, dólar, de punto o de coma). Al almacenarse internamente las fechas como el número de segundos transcurridos desde el 14 de octubre de 1582, la conversión de fechas a valores numéricos puede



generar números extremadamente grandes. Por ejemplo, la fecha 10/29/91 se convierte al valor numérico 12.908.073.600.

**Conversión de numérico a fecha u hora.** Los valores numéricos se convierten a fechas u horas si el valor representa un número de segundos que puede producir una fecha u hora válidos. Para las fechas, los valores numéricos menores que 86.400 se convierten al valor perdido del sistema.

### ***Inserción de nuevos casos***

Al introducir datos en una casilla de una fila vacía, se crea automáticamente un nuevo caso. El Editor de datos inserta el valor perdido del sistema para el resto de las variables de dicho caso. Si hay alguna fila vacía entre el nuevo caso y los casos existentes, las filas en blanco también se convierten en casos nuevos con el valor perdido del sistema para todas las variables. También puede insertar nuevos casos entre casos existentes.

#### ***Para insertar nuevos casos entre los casos existentes***

- ▶ En la Vista de datos, seleccione cualquier casilla del caso (fila) debajo de la posición donde desea insertar el nuevo caso.
- ▶ Elija en los menús:
  - Edición
  - Insertar casos

Se inserta una fila nueva para el caso y todas las variables reciben el valor perdido del sistema.

### ***Inserción de nuevas variables***

La introducción de datos en una columna vacía en la Vista de datos o en una fila vacía en la Vista de variables crea de forma automática una variable nueva con un nombre de variable por defecto (el prefijo *var* y un número secuencial) y un tipo de formato de datos por defecto (numérico). El Editor de datos inserta el valor perdido del sistema en todos los casos de la nueva variable. Si hay columnas vacías en la Vista de datos o filas vacías en Vista de variables entre la nueva variable y las variables existentes, estas filas o columnas también se convierten en nuevas variables con el valor perdido del sistema para todos los casos. También se pueden insertar variables nuevas entre las variables existentes.

#### ***Para insertar nuevas variables entre variables existentes***

- ▶ Seleccione cualquier casilla de la variable a la derecha (Vista de datos) o debajo (Vista de variables) de la posición donde desea insertar la nueva variable.
- ▶ Elija en los menús:
  - Edición
  - Insertar variable

Se insertará una nueva variable con el valor perdido del sistema para todos los casos.

**Para mover variables**

- ▶ Para seleccionar la variable, pulse en el nombre de variable de la Vista de datos o en el número de fila para la variable de la Vista de variables.
- ▶ Arrastre y suelte la variable en la nueva ubicación.
- ▶ Si desea colocar la variable entre dos ya existentes: En la Vista de datos, arrastre la variable sobre la columna de variables a la derecha del lugar donde desea colocar la variable, o en la Vista de variables, arrastre la variable a la fila de variables debajo de donde desee colocarla.

**Para cambiar el tipo de datos**

Puede cambiar el tipo de datos de una variable en cualquier momento mediante el cuadro de diálogo Tipo de variable de la Vista de variables. El Editor de datos intentará convertir los valores existentes en el nuevo tipo. Si no se puede realizar esta conversión, se asignará el valor perdido del sistema. Las reglas de conversión son las mismas que las del pegado de valores de datos en una variable con distinto tipo de formato. Si el cambio del formato de los datos puede generar la pérdida de las especificaciones de valores perdidos o de las etiquetas de valor, el Editor de datos mostrará un cuadro de alerta solicitando confirmación para proseguir o cancelar la operación.

**Búsqueda de casos o variables**

A continuación, el cuadro de diálogo Ir a caso/Variable busca el número (fila) del caso o el nombre de la variable en el Editor de datos.

- ▶ Elija en los menús:
  - Edición
  - Ir a caso

*o*

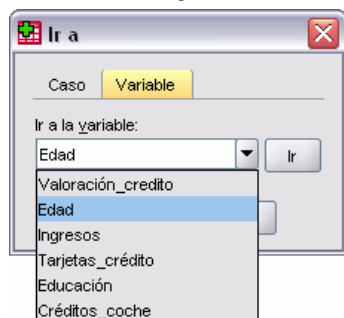
- ▶ Elija en los menús:
  - Edición
  - Ir a la variable
- ▶ Para los casos, escriba un valor entero que represente el número de fila actual en la Vista de datos.

*Nota:* el número de fila actual de un determinado caso puede cambiar debido al orden o a otras acciones.

*o*

- ▶ Para las variables, escriba el nombre de la variable o seleccione la variable en la lista desplegable.

Figura 5-11  
Cuadro de diálogo Ir a



## Búsqueda y sustitución de datos y valores de atributo

Para buscar o sustituir valores de datos en la Vista de datos o valores de atributos en la Vista de variables:

- ▶ Pulse en una casilla de la columna en la que desea buscar. (La búsqueda y sustitución de valores se limita a una única columna.)
  - ▶ Elija en los menús:
    - Edición
    - Buscar
- o
- Edición
  - Reemplazar

### Vista de datos

- No se puede buscar en la Vista de datos. La dirección de búsqueda es siempre hacia abajo.
- Para fechas y horas, se buscan los valores con formato, es decir, tal como aparecen en la Vista de datos. Por ejemplo, si se busca la fecha 10-28-2007 no se encontrará una fecha que aparezca como 10/28/2007.
- Para las demás variables numéricas, Contiene, Comienza por y Termina por buscan valores con formato. Por ejemplo, con la opción Comienza por, un valor de búsqueda de \$123 para una variable con formato dólar encontrará tanto \$123,00 como \$123,40 pero no \$1,234. Con la opción Casilla entera, el valor de búsqueda puede tener formato o no (formato numérico F simple), pero sólo se buscarán valores numéricos exactos (con la precisión mostrada en el Editor de datos).
- Si se muestran las etiquetas de valor para la columna de variable seleccionada, se buscará el texto de la etiqueta y no el valor de datos subyacente. Además, no podrá sustituir el texto de la etiqueta.

### Vista de variables

- La búsqueda sólo está disponible para *Nombre*, *Etiqueta*, *Valores*, *Perdidos* y las columnas de atributos de variable personalizados.

- La sustitución sólo está disponible para *Etiqueta*, *Valores* y columnas de atributos personalizados.
- En la columna *Valores* (etiquetas de valor), la cadena de búsqueda puede buscar el valor de datos o una etiqueta de valor.

*Nota:* la sustitución del valor de datos eliminará cualquier etiqueta de valor anteriormente asociada a dicho valor.

## Estado de selección de casos en el Editor de datos

Si ha seleccionado un subconjunto de casos pero no ha descartado los casos no seleccionados, éstos se marcarán en el Editor de datos con una línea diagonal (barra diagonal) atravesando el número de fila.

Figura 5-12  
Casos filtrados en el Editor de datos

	edad	marital	direcc	ingres	ingcat	coche	cochecat
1	55	Casado	12	72,00	\$50 - \$74	36,20	Lujo
2	56	Sin casar	29	153,00	\$75+	76,90	Lujo
3	28	Casado	9	28,00	\$25 - \$49	13,70	Económico
4	24	Casado	4	26,00	\$25 - \$49	12,50	Económico
5	25	Sin casar	2	23,00	Menos de ...	11,30	Económico
6	45	Casado	9	76,00	\$75+	37,20	Lujo
7	42	Sin casar	19	40,00	\$25 - \$49	19,80	Estándar
8	35	Sin casar	15	57,00	\$50 - \$74	28,20	Estándar
9	46	Sin casar	26	24,00	Menos de ...	12,20	Económico
10	34	Casado	0	89,00	\$75+	46,10	Lujo
11	55	Casado	17	72,00	\$50 - \$74	35,50	Lujo
12	28	Sin casar	3	24,00	Menos de ...	11,80	Económico
13	31	Casado	9	40,00	\$25 - \$49	21,30	Estándar
14	42	Sin casar	8	137,00	\$75+	68,90	Lujo

## Editor de datos: Opciones de presentación

El menú Ver proporciona varias opciones de presentación para el Editor de datos:

**Fuentes.** Esta opción controla las características de fuentes de la presentación de datos.

**Líneas de cuadrícula.** Esta opción activa y desactiva la presentación de las líneas de cuadrícula.

**Etiquetas de valor.** Esta opción activa y desactiva la presentación de los valores reales de los datos y las etiquetas de valor descriptivas definidas por el usuario. Esta opción sólo está disponible en la Vista de datos.

### **Uso de varias vistas**

En la Vista de datos, puede crear varias vistas (paneles) mediante los divisores situados debajo de la barra de desplazamiento horizontal y a la derecha de la barra de desplazamiento vertical.

También puede utilizar el menú Ventana para insertar y eliminar divisores de paneles. Para insertar divisores:

- ▶ En la Vista de datos, elija en los menús:

Ventana  
Dividir

Los divisores se insertan sobre y a la izquierda de la casilla seleccionada.

- Si se ha seleccionado la casilla superior izquierda, los divisores se insertan para dividir la vista actual aproximadamente por la mitad horizontal y verticalmente.
- Si se selecciona una casilla distinta de la casilla superior de la primera columna, se inserta un divisor de paneles horizontales sobre la casilla seleccionada.
- Si se selecciona una casilla distinta de la primera casilla de fila superior, se inserta un divisor de paneles verticales a la izquierda de la casilla seleccionada.

## **Impresión en el Editor de datos**

Los archivos de datos se imprimen tal y como aparece en la pantalla.

- Se imprime la información que está en la vista actualmente mostrada. En la Vista de datos, se imprimen los datos. En la Vista de variables, se imprime la información de definición de los datos.
- Las líneas de cuadrícula se imprimen si aparecen actualmente en la vista seleccionada.
- Las etiquetas de valor se imprimen si aparecen actualmente en la Vista de datos. En caso contrario, se imprimirán los valores de datos reales.

Utilice el menú Ver en la ventana Editor de datos para mostrar u ocultar las líneas de cuadrícula y para que se muestren o no los valores de los datos y las etiquetas de valor.

### **Para imprimir el contenido del Editor de datos**

- ▶ Haga que el editor de datos sea la ventana activa.
- ▶ Puse la pestaña de la vista que desea imprimir.
- ▶ Elija en los menús:  
Archivo  
Imprimir...

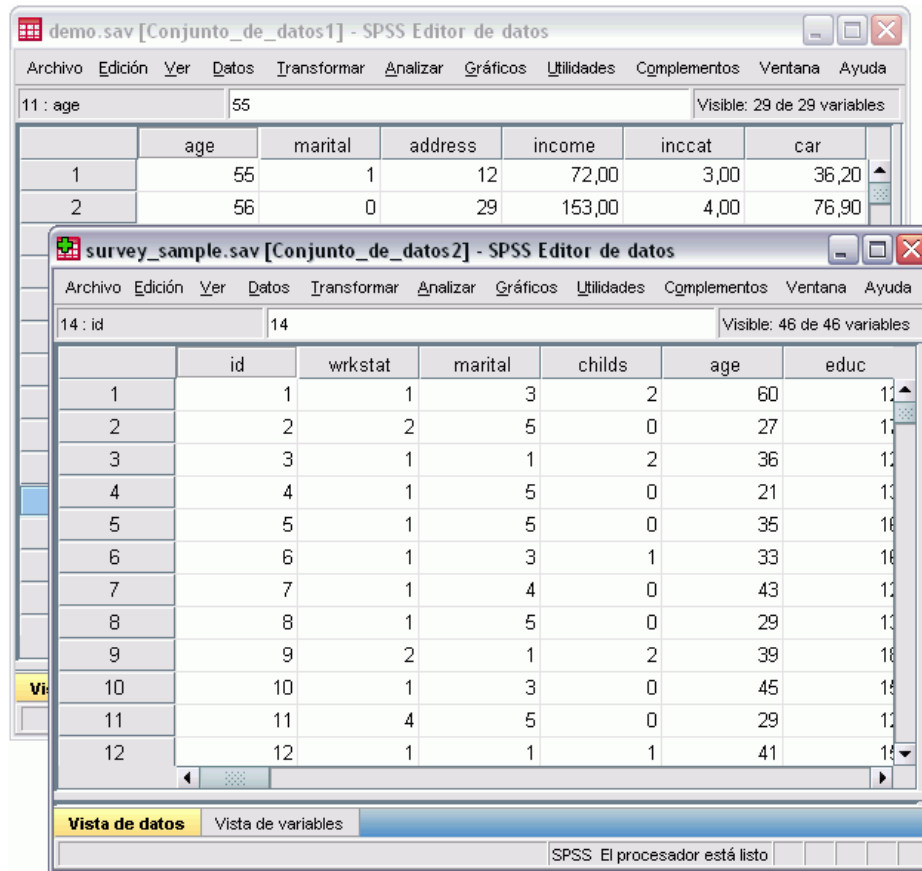
## ***Trabajo con varios orígenes de datos***

A partir de SPSS 14.0, se pueden tener varios orígenes de datos abiertos al mismo tiempo, lo que facilita:

- Cambiar de un origen de datos a otro.
- Comparar el contenido de diferentes orígenes de datos.
- Copiar y pegar datos entre orígenes de datos.
- Crear varios subconjuntos de casos y/o variables para su análisis.
- Fundir varios orígenes de datos con diferentes formatos de datos (por ejemplo, hojas de cálculo, bases de datos, datos en texto) sin tener que guardar antes cada origen de datos.

## Tratamiento básico de varios orígenes de datos

Figura 6-1  
Dos orígenes de datos abiertos al mismo tiempo



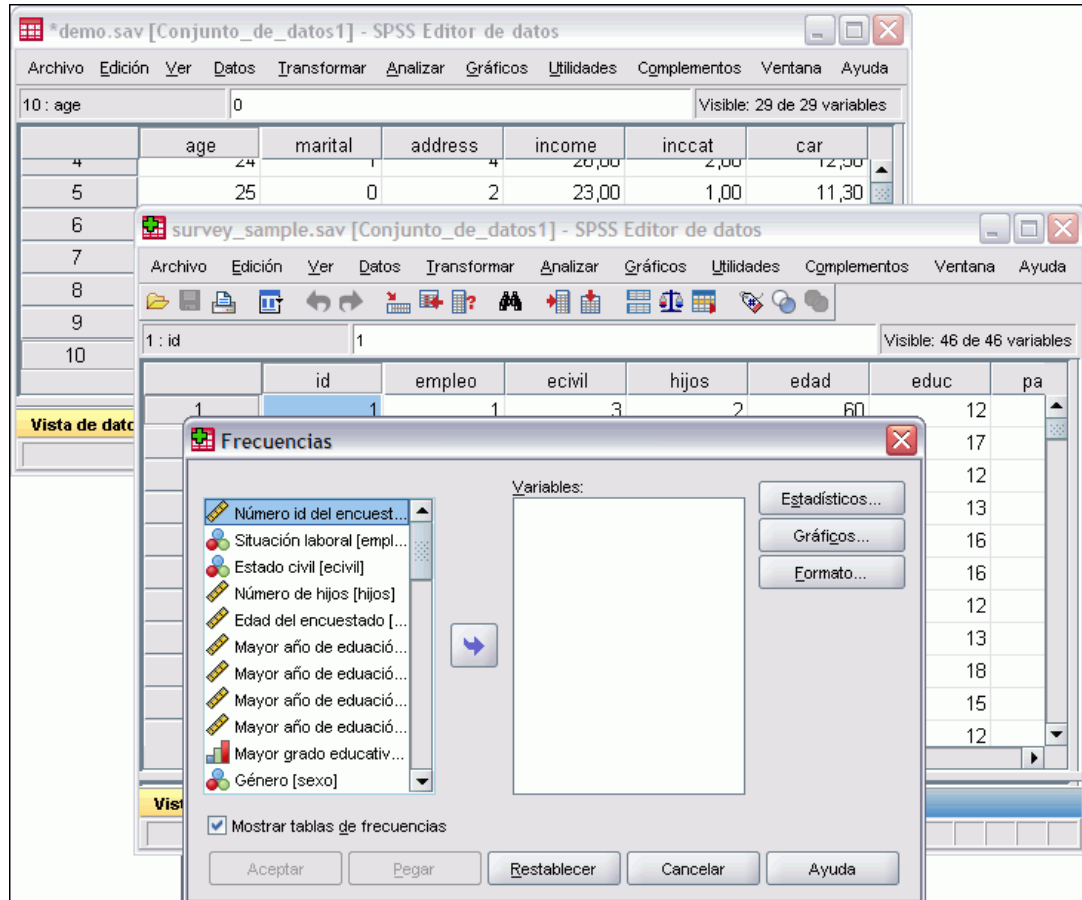
Por defecto, cada origen de datos que se abra aparecerá en una nueva ventana del Editor de datos. (Consulte [General: Opciones](#) para obtener información sobre cómo cambiar el comportamiento por defecto para que sólo muestre un conjunto de datos al mismo tiempo, en una sola ventana Editor de datos.)

- Todos los orígenes de datos que haya abierto anteriormente permanecerán abiertos y estarán disponibles para su uso.
- Al abrir por primera vez un origen de datos, se convierte automáticamente en el **conjunto de datos activo**.
- Para cambiar el conjunto de datos activo basta con pulsar en cualquier parte de la ventana del Editor de datos del origen de datos que desee utilizar o bien seleccionar la ventana del Editor de datos correspondiente a dicho origen de datos en el menú Ventana.

- Sólo será posible analizar las variables del conjunto de datos activo.

Figura 6-2

Lista de variables con las variables del conjunto de datos activo



- No se puede cambiar el conjunto de datos activo mientras esté abierto cualquier cuadro de diálogo que acceda a los datos (incluidos todos los cuadros de diálogo que muestran las listas de variables).
- Al menos una ventana del Editor de datos debe estar abierta durante una sesión. Al cerrar la última ventana abierta del Editor de datos, se cierra automáticamente SPSS, preguntándole antes si desea guardar los cambios.

## ***Trabajo con varios conjuntos de datos en la sintaxis de comandos***

Si utiliza la sintaxis de comandos para abrir orígenes de datos (por ejemplo, GET FILE, GET DATA), deberá utilizar el comando DATASET NAME para indicar explícitamente el nombre de cada conjunto de datos y así poder tener más de un origen de datos abierto al mismo tiempo.

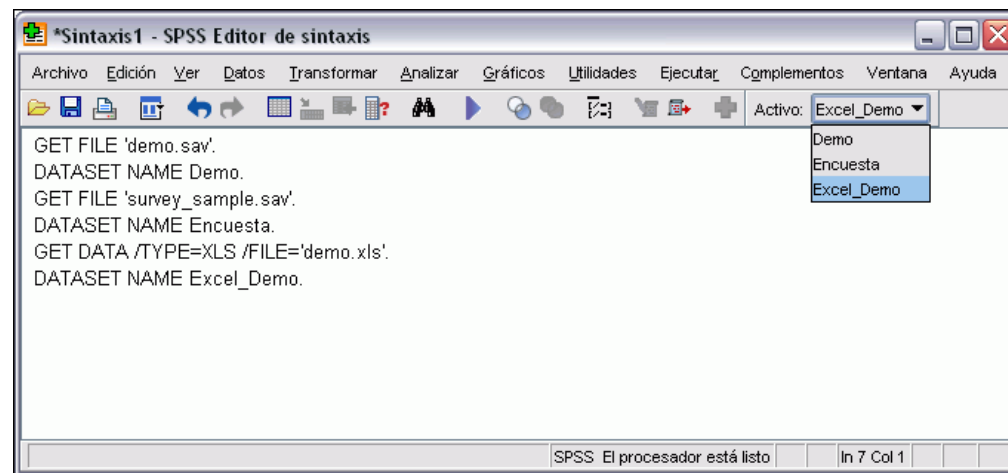


Al trabajar con la sintaxis de comandos, aparece el nombre del conjunto de datos activo en la barra de herramientas de la ventana de sintaxis. Todas las acciones siguientes pueden cambiar el conjunto de datos activo:

- Usar el comando `DATASET ACTIVATE`.
- Pulsar en cualquier punto de la ventana Editor de datos de un conjunto de datos.
- Seleccione el nombre de un conjunto de datos en la barra de herramientas de la ventana de sintaxis.

Figura 6-3

Conjuntos de datos abiertos mostrados en la barra de herramientas de la ventana de sintaxis



## Copia y pegado de información entre conjuntos de datos

Puede copiar tanto datos como atributos de definición de variables de un conjunto de datos a otro, básicamente de la misma manera que copia y pega información en un archivo de datos único.

- Al copiar y pegar determinadas casillas de datos en la Vista de datos se pegan únicamente los valores de los datos, sin los atributos de definición de variables.
- Si se copia y pega una variable entera en la Vista de datos seleccionando el nombre de dicha variable que aparece en la parte superior de la columna, se pegarán todos los datos y todos los atributos de definición de variables correspondientes a dicha variable.
- Al copiar y pegar los atributos de definición de variables o las variables enteras en la Vista de variables, se pegarán los atributos seleccionados (o toda la definición de la variable) pero no se pegará ningún valor de los datos.

## Cambio del nombre de los conjuntos de datos

Al abrir un origen de datos utilizando los menús y los cuadros de diálogo, se le asignará automáticamente a cada origen de datos un nombre de conjunto de datos *Conjunto\_de\_datosn*, donde *n* es un número entero secuencial. Al abrir un origen de datos utilizando la sintaxis de comandos, no se asignará ningún nombre de conjunto de datos a menos que se especifique

uno explícitamente utilizando `DATASET NAME`. Para especificar nombres de conjuntos de datos más descriptivos:

- ▶ En los menús de la ventana del Editor de datos correspondientes al conjunto de datos cuyo nombre desea cambiar, seleccione:
  - Archivo
  - Cambiar nombre de conjunto de datos...
- ▶ Escriba un nuevo nombre de conjunto de datos que cumpla las reglas de denominación de variables. Si desea obtener más información, consulte [Nombres de variable](#) en Capítulo 5 en p. 82.

## ***Supresión de varios conjuntos de datos***

Si prefiere tener un único conjunto de datos disponible al mismo tiempo y desea suprimir la función de varios conjuntos de datos:

- ▶ Elija en los menús:
  - Edición
  - Opciones...
- ▶ Pulse en la pestaña General.

Seleccione (active) Abrir sólo un conjunto de datos cada vez.

Si desea obtener más información, consulte [General: Opciones](#) en Capítulo 45 en p. 479.

# Preparación de datos

Cuando abra un archivo de datos o introduzca datos en el Editor de datos, podrá empezar a crear informes, gráficos y análisis sin ningún trabajo preliminar adicional. Sin embargo, hay algunas funciones de preparación adicional de los datos que pueden resultarle útiles, entre las que se incluyen:

- Asignar propiedades de las variables que describan los datos y determinen cómo se deben tratar ciertos valores.
- Identificar los casos que pueden contener información duplicada y excluir dichos casos de los análisis o eliminarlos del archivo de datos.
- Crear nuevas variables con algunas categorías distintas que representen rangos de valores de variables que tengan un mayor número de valores posibles.

## Propiedades de variables

Los datos introducidos en la Vista de datos del Editor de datos o leídos desde un formato de archivos externo (como una hoja de cálculo de Excel o un archivo de datos de texto) carecen de ciertas propiedades de variables que pueden resultar muy útiles, como:

- Definición de etiquetas de valor descriptivas para códigos numéricos (por ejemplo, 0 = *Hombre* y 1 = *Mujer*).
- Identificación de códigos de valores perdidos (por ejemplo, 99 = *No procede*).
- Asignación del nivel de medida (nominal, ordinal o de escala).

Todas estas propiedades de variables (y otras) se pueden asignar en la Vista de variables del Editor de datos. También hay algunas utilidades que le pueden ofrecer asistencia en este proceso:

- **Definir propiedades de variables** puede ayudarle a definir etiquetas de valor descriptivas y valores perdidos. Esto es especialmente útil para datos categóricos con códigos numéricos utilizados para valores de categorías.
- **Copiar propiedades de datos** ofrece la posibilidad de utilizar un archivo de datos con formato SPSS como plantilla para definir las propiedades de variables y archivos en el archivo de datos actual. Esto es particularmente útil si utiliza frecuentemente archivos de datos con un formato externo que tenga un contenido similar, como puedan ser informes mensuales en formato Excel.

## Definición de propiedades de variables

Definir propiedades de variables está diseñado para ayudarle en el proceso de creación de etiquetas de valor descriptivas para variables categóricas (nominales u ordinales). Definir propiedades de variables:

- Explora los datos reales y enumera todos valores de datos únicos para cada variable seleccionada.
- Identifica valores sin etiquetas y ofrece una función de “etiquetas automáticas”.
- Permite copiar etiquetas de valor definidas de otra variable en la variable seleccionada o de la variable seleccionada a varias variables adicionales.

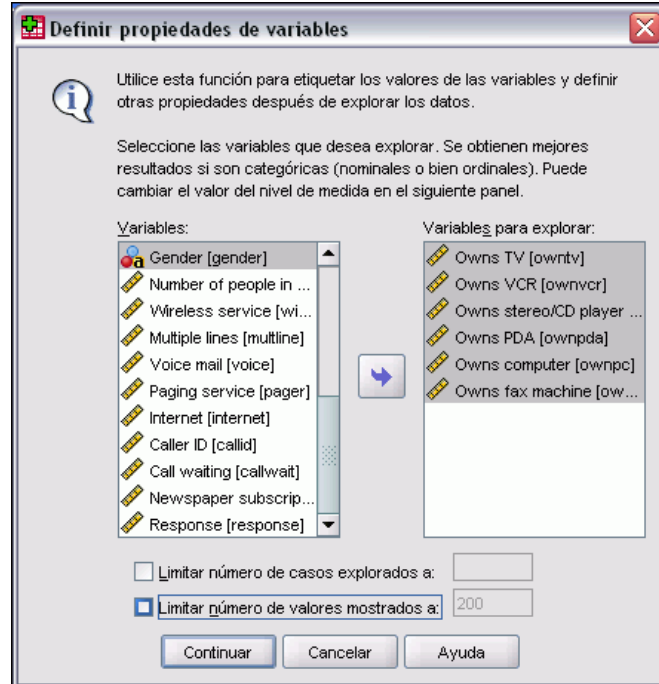
*Nota:* Para utilizar Definir propiedades de variables sin explorar primero los casos, introduzca 0 para el número de casos que se van a explorar.

### Para definir propiedades de variables

- Elija en los menús:  
Datos  
Definir propiedades de variables...

Figura 7-1

Cuadro de diálogo inicial para seleccionar las variables que se van a definir



- Seleccione las variables numéricas o de cadena para las que desea crear etiquetas de valor o definir o cambiar otras propiedades de las variables, como los valores perdidos o las etiquetas de variable descriptivas.

- ▶ Especificar el número de casos que se van a explorar para generar la lista de valores únicos. resulta especialmente útil para los archivos de datos con mayor número de casos, para el cual una exploración del archivo de datos completo podría tardar una gran cantidad de tiempo.
- ▶ Especifique un límite superior para el número de valores únicos que se va a visualizar. Esto es especialmente útil para evitar que se generen listas de cientos, miles o incluso millones de valores para las variables de escala (intervalo continuo, razón).
- ▶ Pulse en Continuar para abrir el cuadro de diálogo Definir propiedades de variables principal.
- ▶ Seleccione una variable para la que desee crear etiquetas de valor o definir o cambiar otras propiedades de las variables.
- ▶ Introduzca el texto de etiqueta para los valores sin etiquetas que se visualicen en Rejilla etiqueta valores.
- ▶ Si hay valores para los que desea crear etiquetas de valor, pero no se visualizan dichos valores, puede introducirlos en la columna *Valores* por debajo del último valor explorado.
- ▶ Repita este proceso para cada variable de la lista para la que desee crear etiquetas de valor.
- ▶ Pulse en Aceptar para aplicar las etiquetas de valor y otras propiedades de las variables.

## Definición de etiquetas de valor y otras propiedades de las variables

Figura 7-2  
Cuadro de diálogo principal Definir propiedades de las variables

Definir propiedades de las variables

Lista de variables exploradas

S.	M	Variable
<input type="checkbox"/>	<input type="checkbox"/>	pc
<input type="checkbox"/>	<input type="checkbox"/>	fax
<input type="checkbox"/>	<input type="checkbox"/>	pda
<input type="checkbox"/>	<input type="checkbox"/>	cd
<input type="checkbox"/>	<input type="checkbox"/>	tv
<input type="checkbox"/>	<input type="checkbox"/>	video

Variable actual: tienefax Etiqueta: Tiene fax

Nivel de medida: Ordinal  Tipo: Numéricc

Valores sin etiqueta: 0  Ancho: 4 Decimales: 0

Rejilla eti. valores:  Añada etiquetas a la rejilla o edítelas. Puede añadir valores abajo.

	Cambiado	Perdido	Recuento	Valor	Etiqueta
1	<input type="checkbox"/>	<input type="checkbox"/>	5198	0	No
2	<input type="checkbox"/>	<input type="checkbox"/>	1202	1	Sí
3	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	0	99	No contesta
4	<input type="checkbox"/>	<input type="checkbox"/>			

Copiar propiedades:

Valores sin etiquetas:

Casos explorados: 6400  
Límite lista valores: 200

El cuadro de diálogo principal Definir propiedades de variables proporciona la siguiente información para las variables exploradas:

**Lista de variables exploradas.** Para cada variable explorada, aparecerá una marca de verificación en la columna *Sin etiqueta (S/E)* indicando que la variable contiene valores sin etiquetas de valor asignadas.

Para ordenar la lista de variables para que aparezcan todas las variables con valores sin etiquetas en la parte superior de la lista:

- Pulse en el encabezado de columna *Sin etiqueta* debajo de la Lista de variables exploradas.

También puede ordenarla por nombre de variable o nivel de medida pulsando en el encabezado de columna correspondiente debajo de la Lista de variables exploradas.

### **Rejilla etiqueta valores**

- **Etiqueta.** Muestra las etiquetas de valor que ya se han definido. Puede añadir o cambiar las etiquetas de esta columna.
- **Valor.** Valores únicos para cada variable seleccionada. Esta lista de valores únicos se basa en el número de casos explorados. Por ejemplo, si sólo ha explorado los primeros 100 casos del archivo de datos, la lista reflejará sólo los valores únicos presentes en esos casos. Si el archivo de datos ya se ha ordenado por la variable para la que desea asignar etiquetas de valor, la lista puede mostrar muchos menos valores únicos de los que hay realmente presentes en los datos.
- **Recuento.** Número de veces que aparece cada valor en los casos explorados.
- **Perdidos.** Valores definidos para representar valores perdidos. Puede cambiar la designación de la categoría de los valores perdidos pulsando en la casilla de verificación. Una marca indica que la categoría se ha definido como categoría perdida por el usuario. Si una variable ya tiene un rango de valores definidos como perdidos por el usuario (por ejemplo 90 - 99), no podrá añadir ni eliminar categorías de valores perdidos para esa variable con Definir propiedades de variables. Puede utilizar la Vista de variables del Editor de datos para modificar las categorías de valores perdidos para la variable con rangos de valores perdidos. Si desea obtener más información, consulte [Valores perdidos](#) en Capítulo 5 en p. 87.
- **Cambiado.** Indica que ha añadido o cambiado una etiqueta de valor.

*Nota:* Si ha especificado 0 para el número de casos que se van a explorar en el cuadro de diálogo inicial, la rejilla etiqueta valores estará en blanco al principio, a excepción de algunas etiquetas de valor ya existentes y/o categorías de valores perdidos definidas para la variable seleccionada. Además, se desactivará el botón Sugerir para el nivel de medida.

**Nivel de medida.** Las etiquetas de valor son especialmente útiles para las variables categóricas (nominales u ordinales), y algunos procedimientos tratan a las variables categóricas y de escala de manera diferente, por lo que a veces es importante asignar el nivel de medida correcto. Sin embargo, por defecto, todas las nuevas variables numéricas se asignan al nivel de medida de escala. Por tanto, puede que muchas variables que son de hecho categóricas, aparezcan inicialmente como variables de escala.

Si no está seguro de qué nivel de medida debe asignar a una variable, pulse en Sugerir.

**Copiar propiedades.** Puede copiar las etiquetas de valor y otras propiedades de las variables de otra variable a la variable seleccionada en ese momento o desde la variable seleccionada en ese momento a una o varias otras variables.

**Valores sin etiquetas.** Para crear automáticamente etiquetas para valores sin etiquetas, pulse en Etiquetas automáticas.

### ***Etiqueta de variable y formato de presentación***

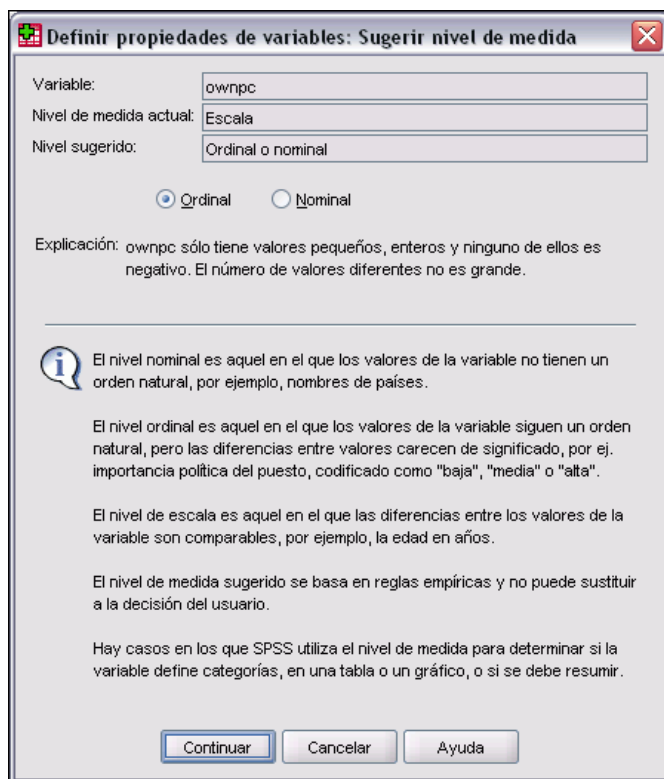
Puede cambiar de la etiqueta de variable descriptiva y el formato de presentación.

- No puede cambiar el tipo fundamental de la variable (numérica o de cadena).
- Para las variables de cadena, sólo puede cambiar la etiqueta de variable, no el formato de presentación.
- Para las variables numéricas, puede cambiar el tipo numérico (como numérico, fecha, dólar o moneda personalizada), el ancho (número máximo de dígitos, incluyendo los indicadores decimales y/o de agrupación) y el número de posiciones decimales.
- Para el formato de fecha numérica, puede seleccionar un formato de fecha específico (como dd-mm-aaaa, mm/dd/aa, aaaadddd)
- Para formato numérico personalizado, puede seleccionar uno de los cinco formatos de moneda personalizados (de CCA a CCE). Si desea obtener más información, consulte [Moneda: Opciones](#) en Capítulo 45 en p. 484.
- Aparece un asterisco en la columna *Valor* si el ancho especificado es inferior al ancho de los valores explorados o los valores mostrados para etiquetas de valor definidas ya existentes o categorías de valores perdidos.
- Aparece un período (.) si los valores explorados o los valores mostrados para etiquetas de valor definidas ya existentes o categorías de valores perdidos no son válidos para el tipo de formato de presentación seleccionado. Por ejemplo, un valor numérico interno inferior a 86.400 no es válido para una variable de formato de fecha.

### ***Asignación del nivel de medida***

Cuando pulse en Sugerir para seleccionar un nivel de medida en el cuadro de diálogo principal Definir propiedades de variables, la variable actual se evalúa en función de los casos explorados y las etiquetas de valor definidas y se sugiere un nivel de medida en el cuadro de diálogo Sugerir nivel de medida que se abre. El área Explicación ofrece una breve descripción de los criterios utilizados para proporcionar el nivel de medida sugerido.

Figura 7-3  
Cuadro de diálogo Sugerir nivel de medida



*Nota:* Los valores definidos para representar valores perdidos no se incluyen en la evaluación para el nivel de medida. Por ejemplo, la explicación del nivel de medida sugerido puede indicar que la sugerencia se basa, en parte, en el hecho de que la variable no contiene valores negativos, mientras que, de hecho, puede contener valores negativos, pero dichos valores ya se han definido como valores perdidos.

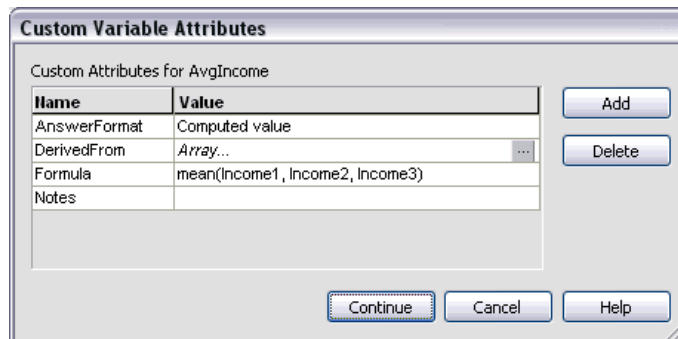
- Pulse en Continuar para aceptar el nivel de medida sugerido o en Cancelar para mantener el mismo.

### ***Atributos de variable personalizados***

El botón Atributos del cuadro de diálogo Definir propiedades de variables abre el cuadro de diálogo Atributos personalizados de variables. Además de los atributos de variable estándar, como las etiquetas de valores, los valores perdidos y el nivel de medida, puede crear sus propios atributos de variable personalizados. Al igual que los atributos de variable estándar, estos atributos personalizados se guardan en los archivos de datos con formato SPSS.



Figura 7-4  
Atributos de variable personalizados



**Nombre.** Los nombres de atributo deben cumplir las mismas reglas que los nombres de variable. Si desea obtener más información, consulte [Nombres de variable](#) en Capítulo 5 en p. 82.

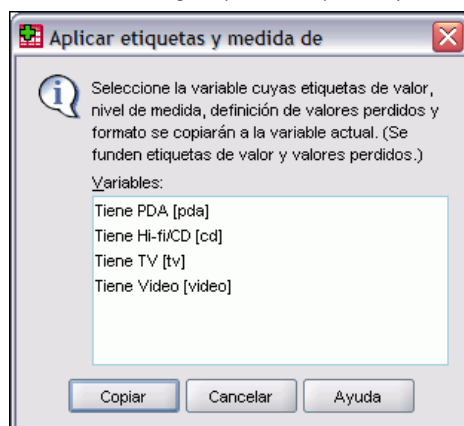
**Valor.** Valor asignado al atributo de la variable seleccionada.

- Los nombres de atributo que comienzan con un signo de dólar son reservados y no se pueden modificar.
- Si aparece el texto *Matriz...*, en una casilla de valor, indica que se trata de una **matriz de atributos**, un atributo que contiene varios valores. Pulse en el botón de la casilla para mostrar la lista de valores.

## Copia de propiedades de variables

El cuadro de diálogo Aplicar etiquetas y nivel a aparece al pulsar en De otra variable o A otras variables en el cuadro de diálogo principal Definir propiedades de variables. Muestra todas las variables exploradas que coinciden con el tipo de variable actual (de cadena o numérico). Para las variables de cadena, también debe coincidir la anchura definida.

Figura 7-5  
Cuadro de diálogo Aplicar etiquetas y medida a



- ▶ Seleccione una única variable desde la que va a copiar las etiquetas de valor y otras propiedades de las variables (excepto la etiqueta de la variable).  
*o*
- ▶ Seleccione una o más variables a las que va a copiar las etiquetas de valor y otras propiedades de las variables.
- ▶ Pulse en Copiar para copiar las etiquetas de valor y el nivel de medida.
  - Las etiquetas de valor existentes y categorías de valores perdidos para las variables de destino no se sustituyen.
  - Las etiquetas de valor y las categorías de valores perdidos para los valores que no se han definido aún para las variables de destino se añaden al conjunto de etiquetas de valor y categorías de valores perdidos para las variables de destino.
  - El nivel de medida para las variables de destino siempre se sustituye.
  - Si la variable de origen o de destino tiene un rango definido de valores perdidos, no se copian las definiciones de los valores perdidos.

## ***Conjuntos de respuestas múltiples***

Las Tablas personalizadas admiten un tipo especial de “variable” al que se denomina **conjunto de respuestas múltiples**. En realidad, los conjuntos de respuestas múltiples no son, en sentido estricto, “variables”. No aparecen en el Editor de datos y los demás procedimientos no los reconocen. Los conjuntos de respuestas múltiples utilizan varias variables para registrar respuestas a preguntas en las que el encuestado puede ofrecer más de una respuesta. Los conjuntos de respuestas múltiples se consideran variables categóricas y la mayor parte de las acciones que se pueden realizar con las variables categóricas se pueden realizar también con conjuntos de respuestas múltiples.

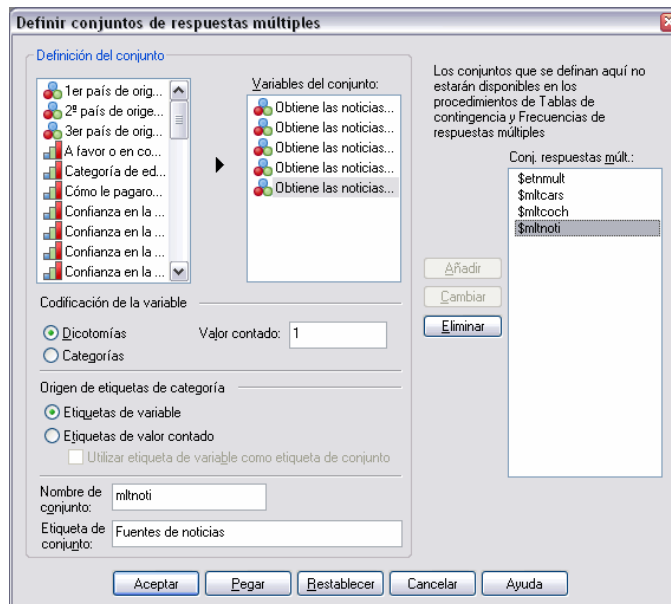
Los conjuntos de respuestas múltiples se crean a partir de múltiples variables del archivo de datos. Un conjunto de respuestas múltiples es un constructo especial perteneciente a un archivo de datos con formato SPSS. Se pueden definir y guardar varios conjuntos en un archivo de datos con formato SPSS, pero no se pueden importar o exportar conjuntos de respuestas múltiples desde o a otros formatos de archivo. Se pueden copiar conjuntos de respuestas múltiples de otros archivos de datos de SPSS mediante Copiar propiedades de datos en el menú Datos en la ventana Editor de datos. Si desea obtener más información, consulte [Copia de propiedades de datos](#) en p. 117.

### ***Para definir conjuntos de respuestas múltiples***

Para definir conjuntos de respuestas múltiples:

- ▶ Elija en los menús:
  - Datos
  - Definir conjuntos de respuestas múltiples...

Figura 7-6  
Cuadro de diálogo Definir conjuntos de respuestas múltiples



- ▶ Seleccione dos o más variables. Si las variables están codificadas como dicotomías, indique qué valor desea contar.
- ▶ Escriba un nombre distintivo para cada conjunto de respuestas múltiples. El nombre puede tener hasta 63 bytes. De forma automática, se añade un signo dólar al comienzo del nombre del conjunto.
- ▶ Escriba una etiqueta descriptiva para el conjunto. (Esto es opcional.)
- ▶ Pulse Añadir para añadir el conjunto de respuestas múltiples a la lista de conjuntos definidos.

### Dicotomías

Un conjunto de dicotomías múltiples se compone normalmente de múltiples variables dicotómicas: variables con sólo dos valores posibles del tipo sí/no, presente/ausente, seleccionado/no seleccionado. Si bien las variables pueden no ser estrictamente dicotómicas, todas las variables del conjunto se codifican de la misma manera, y el valor contado representa la condición correspondiente a afirmativo/ presente/seleccionado.

Por ejemplo, una encuesta formula la pregunta, “¿En cuáles de las siguientes fuentes confía para obtener noticias?” y proporciona cinco posibles respuestas. El encuestado puede señalar varias opciones marcando un cuadro situado junto a cada opción. Las cinco respuestas se convierten en cinco variables en el archivo de datos, con las codificaciones 0 para *No* (no seleccionado) y 1 para *Sí* (seleccionado). En el conjunto de dicotomías múltiples, el valor contado es 1.

El archivo de datos de muestra *survey\_sample.sav* ya contiene tres conjuntos de respuestas múltiples definidos. *\$mltnews* es un conjunto de dicotomías múltiples.

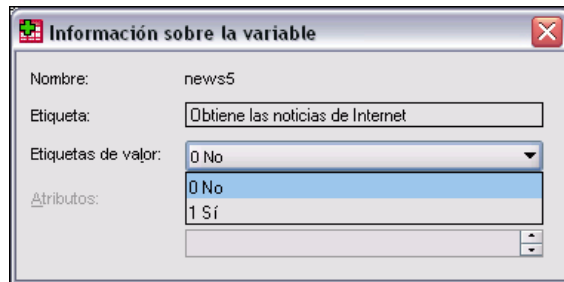
- ▶ Seleccione (pulse en) *\$mltnews* en la lista Conj. respuestas múlt..

Con ello se muestran las variables y las opciones utilizadas para definir este conjunto de respuestas múltiples.

- La lista Variables del conjunto, muestra las cinco variables utilizadas para construir el conjunto de respuestas múltiples.
  - El grupo Codificación de la variable indica que las variables son dicotómicas.
  - El valor contado es 1.
- ▶ Seleccione (pulse en) una de las variables de la lista Variables del conjunto.
  - ▶ Pulse con el botón derecho del ratón en la variable y seleccione Información sobre la variable en el menú contextual emergente.
  - ▶ En la ventana Información sobre la variable, pulse en la flecha de la lista desplegable Etiquetas de valor para mostrar toda la lista de etiquetas de valor definidas.

Figura 7-7

Información de las variables para una variable origen de dicotomías múltiples



Las etiquetas de valor indican que la variable es una dicotomía con valores de 0 y 1, que representan *No* y *Sí*, respectivamente. Las cinco variables de la lista están codificadas de la misma manera y el valor de 1 (el código para *Sí*) es el valor contado para el conjunto de dicotomías múltiples.

### **Categorías**

Un conjunto de categorías múltiples se compone de varias variables, todas ellas codificadas de la misma manera, a menudo con muchas posibles categorías de respuestas. Por ejemplo, un elemento de la encuesta pregunta, “Nombre hasta tres nacionalidades que mejor describan su herencia étnica”. Puede haber cientos de respuestas posibles, pero por cuestiones de codificación se ha limitado la lista a las 40 nacionalidades más comunes, con cualquier otra opción relegada a la categoría “otras”. En el archivo de datos, las tres opciones se convierten en tres variables, cada una con 41 categorías (40 nacionalidades codificadas más la categoría “otras”).

En el archivo de datos de muestra, *\$ethmult* y *\$mltcars* son conjuntos de categorías múltiples.

### **Origen de etiquetas de categoría**

Para dicotomías múltiples, puede controlar cómo se etiquetan los conjuntos.

- **Etiquetas de variable.** Utiliza las etiquetas de variable definidas (o los nombres de variable para las variables que no tienen etiquetas de variable definidas) como las etiquetas de categoría de conjunto. Por ejemplo, si todas las variables del conjunto tienen la misma etiqueta de valor (o no tienen etiquetas de valor definidas) para el valor contado (por ejemplo, *Sí*), debe utilizar las etiquetas de variable como las etiquetas de categoría de conjunto.

- **Etiquetas de valores contados.** Utiliza las etiquetas de valor definidas de los valores contados como etiquetas de categoría de conjunto. Seleccione esta opción sólo si todas las variables tienen una etiqueta de valor definida para el valor contado y la etiqueta de valor para el valor contado es distinta para cada variable.
- **Utilizar etiqueta de variable como etiqueta de conjunto.** Si selecciona Etiquetas de valores contados, también puede utilizar la etiqueta de variable para la primera variable del conjunto con una etiqueta de variable definida como la etiqueta de conjunto. Si ninguna de las variables del conjunto tiene etiquetas de variable definidas, el nombre de la primera variable del conjunto se utiliza como la etiqueta de conjunto.

## ***Copia de propiedades de datos***

El Asistente para la copia de propiedades de datos ofrece la posibilidad de utilizar un archivo de datos de SPSS externo como plantilla para definir las propiedades del archivo y las variables del conjunto de datos activo. También puede utilizar variables del conjunto de datos activo como plantillas para otras variables del conjunto de datos activo. Tiene la posibilidad de:

- Copiar las propiedades de archivo seleccionadas de un archivo de datos externo o de un conjunto de datos abierto en el conjunto de datos activo. Las propiedades de archivo incluyen documentos, etiquetas de archivos, conjuntos de respuestas múltiples, conjuntos de variables y ponderación.
- Copiar las propiedades de archivo seleccionadas de un archivo de datos externo o de un conjunto de datos abierto en las variables coincidentes del conjunto de datos activo. Las propiedades de variable incluyen etiquetas de valor, valores perdidos, nivel de medida, etiquetas de variable, formatos de impresión y escritura, alineación y ancho de columna (en el Editor de datos).
- Copiar las propiedades de variable seleccionadas de una variable, ya sea del archivo de datos externo, de un conjunto de datos abierto o del conjunto de datos activo, en diversas variables del conjunto de datos activo.
- Crear nuevas variables en el conjunto de datos activo basándose en las variables seleccionadas del archivo de datos externo o un conjunto de datos abierto.

Al copiar las propiedades de datos, se aplicarán las reglas siguientes:

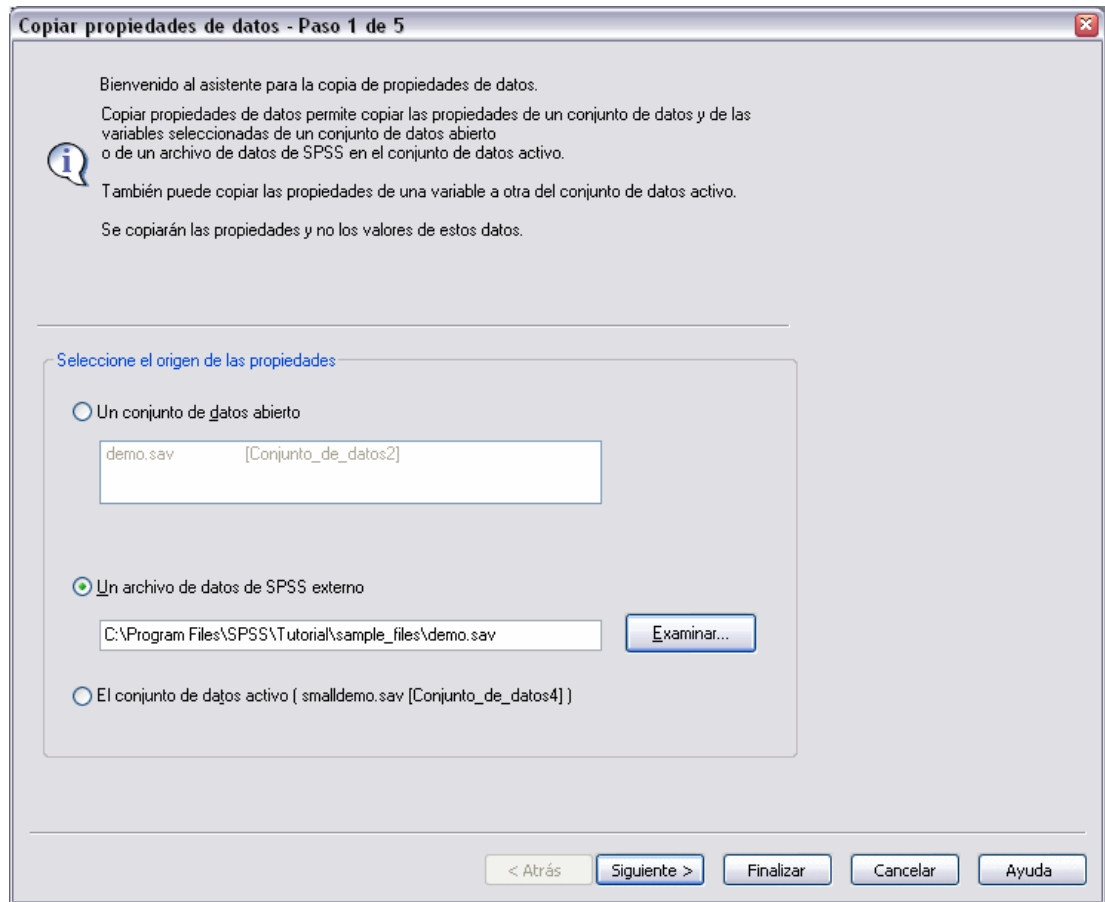
- Para copiar un archivo de datos externo como archivo de datos de origen, deberá tratarse de un archivo de datos con formato SPSS.
- Para utilizar el conjunto de datos activo como archivo de datos de origen, deberá contener al menos una variable. No podrá utilizar un conjunto de datos activo que esté completamente en blanco como archivo de datos de origen.
- Las propiedades no definidas (vacías) del conjunto de datos de origen no sobrescriben las propiedades definidas en el conjunto de datos activo.
- Las propiedades de variable se copian desde la variable de origen únicamente a las variables de destino de un tipo coincidente: de cadena (alfanuméricas) o numérico (incluidas numéricas, fecha y moneda).

*Nota:* En el menú Archivo, Copiar propiedades de datos sustituirá a Aplicar diccionario de datos, disponible anteriormente.

## Para copiar propiedades de datos

- ▶ Seleccione en los menús de la ventana Editor de datos:  
Datos  
Copiar propiedades de datos...

Figura 7-8  
Asistente para la copia de propiedades de datos: Paso 1

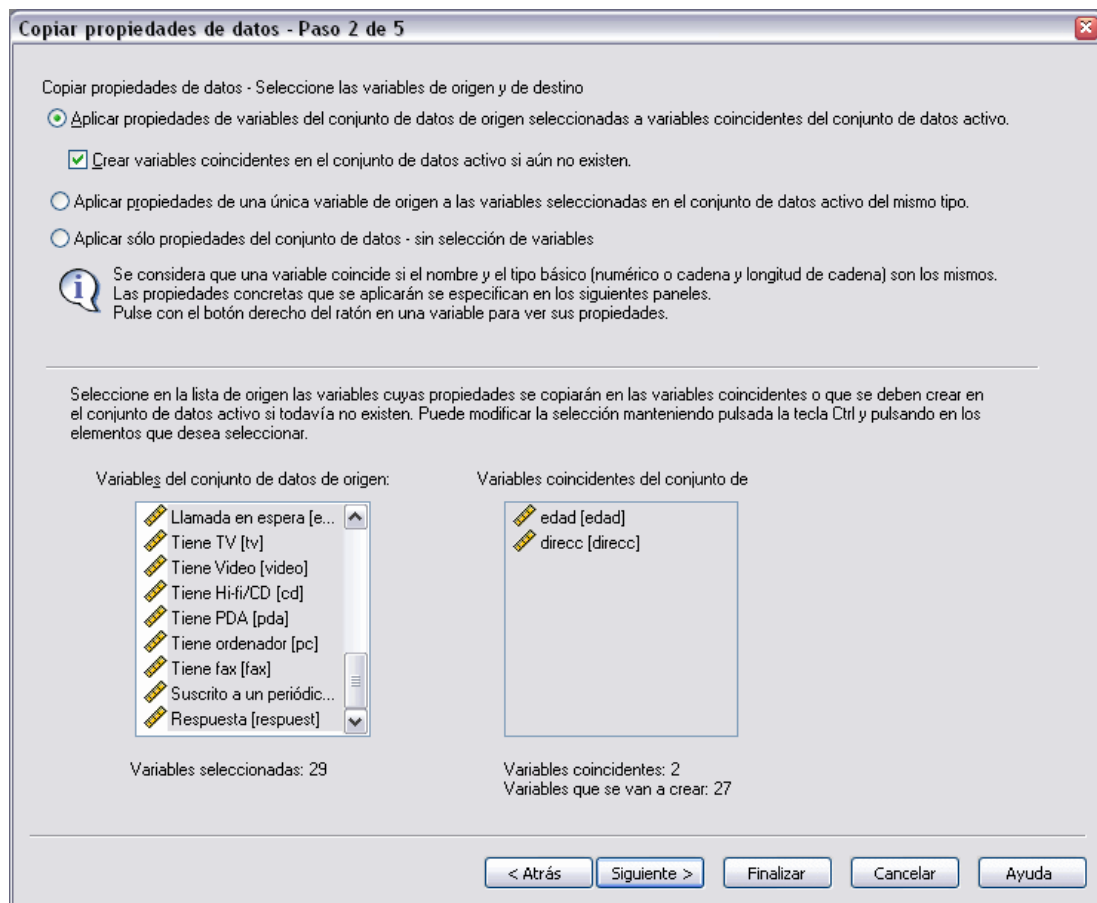


- ▶ Seleccione el archivo de datos que contenga las propiedades de archivo y/o variable que desea copiar. Puede ser un conjunto de datos abierto actualmente, un archivo de datos con formato SPSS externo o el conjunto de datos activo.
- ▶ Siga las instrucciones detalladas del Asistente para la copia de propiedades de datos.

## Selección de las variables de origen y de destino

En este paso, puede especificar tanto las variables de origen que contienen las propiedades de variable que desea copiar como las variables de destino en las que se copiarán estas propiedades de variable.

Figura 7-9  
Asistente para la copia de propiedades de datos: Paso 2



**Aplicar propiedades de variables del conjunto de datos de origen seleccionadas a variables coincidentes del conjunto de datos activo.** Las propiedades de variable se copian desde una o más variables de origen seleccionadas en las variables coincidentes del conjunto de datos activo. Las variables “coinciden” si el nombre y el tipo de variable (de cadena o numérico) son los mismos. En el caso de las variables de cadena, la longitud también debe ser la misma. Por defecto, sólo se muestran en las dos listas de variables las variables coincidentes.

- **Crear variables coincidentes en el conjunto de datos activo si aún no existen.** Actualiza la lista de origen para que muestre todas las variables del archivo de datos de origen. Si se seleccionan variables de origen que no existen en el conjunto de datos activo (basándose en el nombre de variable), se crearán nuevas variables en el conjunto de datos activo con los nombres y las propiedades de variable del archivo de datos de origen.

Si el conjunto de datos activo no contiene variables (un nuevo conjunto de datos en blanco), se mostrarán todas las variables del archivo de datos de origen y se crearán automáticamente en el conjunto de datos activo nuevas variables basadas en las variables de origen seleccionadas.

**Aplicar propiedades de una única variable de origen a variables seleccionadas del mismo tipo del conjunto de datos activo.** Las propiedades de variable de una única variable seleccionada en la lista de origen se pueden aplicar a una o más variables seleccionadas de la lista del conjunto de

datos activo. En esta lista sólo se mostrarán las variables que sean del mismo tipo (numérico o de cadena) que la variable seleccionada en la lista de origen. Si se trata de variables de cadena, sólo se mostrarán las cadenas con la misma longitud definida que la variable de origen. Esta opción no está disponible si el conjunto de datos activo no contiene variables.

*Nota:* No se pueden crear nuevas variables en el conjunto de datos activo con esta opción.

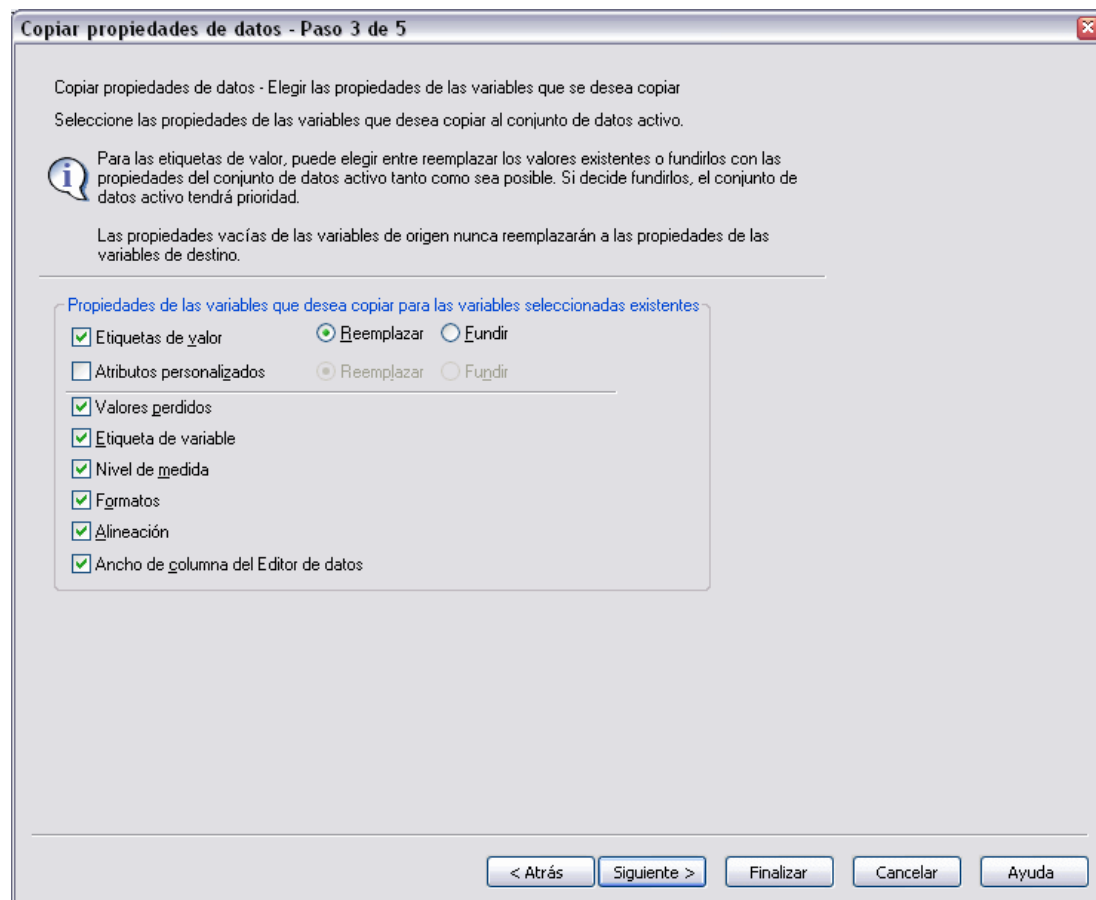
**Aplicar sólo propiedades de conjunto de datos (sin selección de variables).** Sólo se pueden aplicar al conjunto de datos activo las propiedades de archivo (por ejemplo, documentos, etiquetas de archivo, grosor). No se podrá aplicar ninguna propiedad de variable. Esta opción no está disponible si el conjunto de datos activo es también el archivo de datos de origen.

### ***Selección de propiedades de variable para copiar***

Desde las variables de origen, las propiedades de variable seleccionadas no se pueden copiar en las variables de destino. Las propiedades no definidas (vacías) de las variables de origen no sobrescriben las propiedades definidas en las variables de destino.

Figura 7-10

*Asistente para la copia de propiedades de datos: Paso 3*





**Etiquetas de valor.** Las etiquetas de valor son etiquetas descriptivas asociadas a valores de datos. Se suelen utilizar cuando se seleccionan valores de datos numéricos para representar categorías no numéricas (por ejemplo, códigos 1 y 2 para *Hombre* y *Mujer*). Puede reemplazar o fundir las etiquetas de valor en las variables de destino.

- Reemplazar elimina todas las etiquetas de valor definidas para la variable de destino y las reemplaza por las etiquetas de valor definidas en la variable de origen.
- Fundir funde las etiquetas de valor definidas en la variable de origen con cualquier etiqueta de valor definida existente en la variable de destino. Si existe una etiqueta de valor definida con el mismo valor tanto en la variable de origen como en la de destino, la etiqueta de valor de la variable de destino permanecerá inalterada.

**Atributos personalizados.** Atributos de variable personalizados definidos por el usuario. Si desea obtener más información, consulte [Atributos de variable personalizados](#) en Capítulo 5 en p. 89.

- Reemplazar elimina todos los atributos personalizados para la variable de destino y los reemplaza por los atributos definidos en la variable de origen.
- Fundir funde los atributos definidos de la variable de origen con cualquier atributo definido existente en la variable de destino.

**Valores perdidos.** Los valores perdidos son valores identificados como representantes de datos perdidos (por ejemplo, 98 para *No se conoce* y 99 para *No procede*). Por lo general, estos valores tienen también etiquetas de valor definidas que describen el significado de códigos de valores perdidos. Todos los valores perdidos existentes definidos para la variable de destino se eliminarán y se reemplazarán por los valores perdidos de la variable de origen.

**Etiqueta de variable.** Las etiquetas de variable descriptivas pueden contener espacios y caracteres reservados que no se permiten en los nombres de las variables. Si desea utilizar esta opción para copiar propiedades de variable desde una variable de origen en varias variables de destino, reflexione antes de hacerlo.

**Nivel de medida.** El nivel de medida puede ser nominal, ordinal o de escala. En los procedimientos que diferencian entre los distintos niveles de medida, tanto el nivel nominal como el ordinal se consideran **categoricos**.

**Formatos.** Controla el tipo numérico (como numérico, fecha o moneda), el ancho (número total de caracteres que se muestran, incluidos los caracteres iniciales y finales y el indicador decimal) y el número de decimales que se van a mostrar para las variables numéricas. Esta opción no se tendrá en cuenta para las variables de cadena.

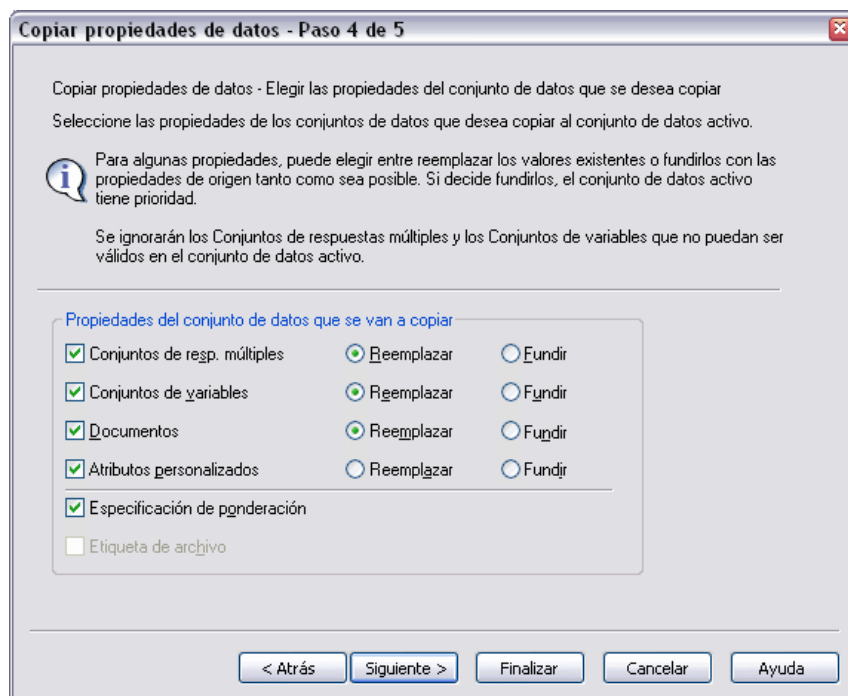
**Alineación.** Afecta únicamente a la alineación (izquierda, derecha, central) del Editor de datos de la Vista de datos.

**Ancho de columna del Editor de datos.** Afecta únicamente al ancho de columna de la Vista de datos del Editor de datos.

## ***Copia de propiedades (de archivo) de conjunto de datos***

Las propiedades de conjunto de datos globales seleccionadas del archivo de datos de origen se pueden aplicar al conjunto de datos activo. (Esta opción no está disponible si el conjunto de datos activo es el archivo de datos de origen.)

Figura 7-11  
Asistente para la copia de propiedades de datos: Paso 4



**Conjuntos resp. múltiples.** Aplica varias definiciones del conjunto de respuestas múltiples del archivo de datos de origen al conjunto de datos activo. (*Nota:* en este momento, únicamente el generador de gráficos y el componente adicional Tablas utilizan los conjuntos de respuestas múltiples.)

- Se ignorarán los conjuntos de respuestas múltiples del archivo de datos de origen que contengan variables no existentes en el conjunto de datos activo, a menos que se creen estas variables basándose en las especificaciones del paso 2 (Selección de las variables de origen y de destino) del Asistente para la copia de propiedades de datos.
- Reemplazar elimina todos los conjuntos de respuestas múltiples del conjunto de datos activo y los reemplaza por los incluidos en el archivo de datos de origen.
- Fundir añade los conjuntos de respuestas múltiples del archivo de datos de origen a la colección de este tipo de conjuntos incluida en el conjunto de datos activo. En caso de que exista un conjunto con el mismo nombre en ambos archivos, el conjunto existente del conjunto de datos activo permanecerá inalterado.

**Conjuntos de variables.** Los conjuntos de variables se utilizan para controlar la lista de variables que se muestra en los cuadros de diálogo. Para definir conjuntos de variables, seleccione Definir conjuntos en el menú Utilidades.

- Se ignorarán los conjuntos del archivo de datos de origen que contengan variables no existentes en el conjunto de datos activo, a menos que se creen estas variables basándose en las especificaciones del paso 2 (Selección de las variables de origen y de destino) del Asistente para la copia de propiedades de datos.

- Reemplazar elimina todos los conjuntos de variables existentes en el conjunto de datos activo y los reemplaza por los incluidos en el archivo de datos de origen.
- Fundir añade los conjuntos de variables del archivo de datos de origen a la colección de este tipo de conjuntos incluida en el conjunto de datos activo. En caso de que exista un conjunto con el mismo nombre en ambos archivos, el conjunto existente del conjunto de datos activo permanecerá inalterado.

**Documentos.** Notas añadidas al archivo de datos a través del comando `DOCUMENT`.

- Reemplazar elimina todos los documentos existentes en el conjunto de datos activo y los reemplaza por los incluidos en el archivo de datos de origen.
- Fundir combina los documentos incluidos en los conjuntos de datos de origen y de trabajo. Los documentos exclusivos del archivo de origen que no existan en el conjunto de datos activo se añadirán al conjunto de datos activo. A continuación, todos los documentos se ordenarán por fecha.

**Atributos personalizados.** Atributos del archivo de datos personalizados, creados normalmente por el comando `DATAFILE ATTRIBUTE` en la sintaxis de comandos.

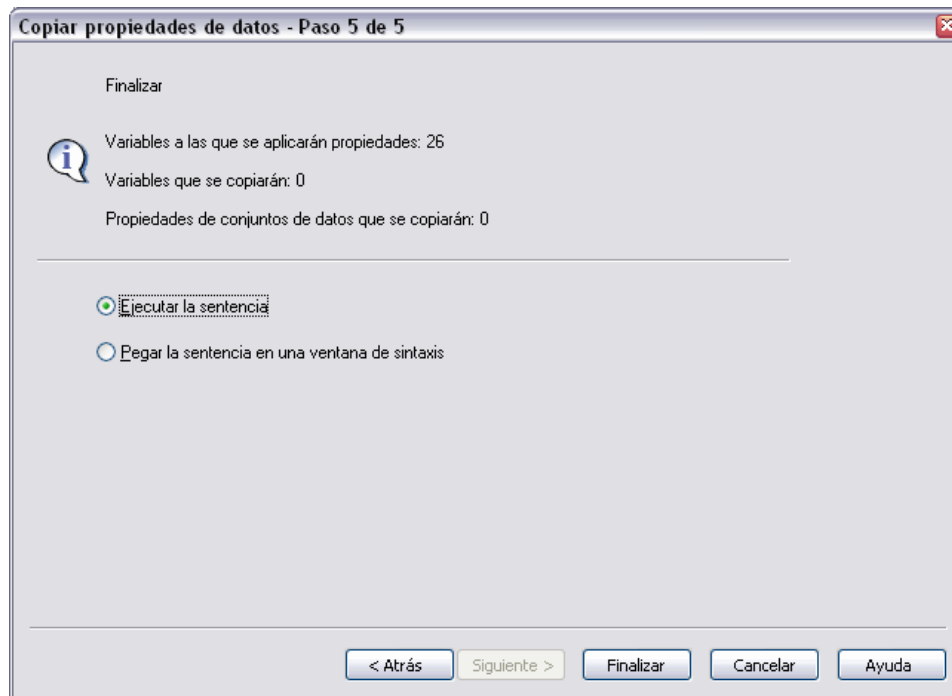
- Reemplazar elimina todos los atributos del archivo de datos personalizados existentes en el conjunto de datos activo y los reemplaza por los incluidos en el archivo de datos de origen.
- Fundir combina los del archivo de datos de los conjuntos de datos de origen y activo. Los nombres de atributos exclusivos del archivo de origen que no existan en el conjunto de datos activo se añadirán al conjunto de datos activo. En caso de que exista un atributo con el mismo nombre en ambos archivos de datos, el atributo con nombre existente en el conjunto de datos activo permanecerá inalterado.

**Especificación de ponderación.** Pondera los casos por la variable de ponderación actual del archivo de datos de origen, siempre que exista una variable coincidente en el conjunto de datos activo. Sobrescribe cualquier ponderación activada actualmente en el conjunto de datos activo.

**Etiqueta de archivo.** Etiqueta descriptiva que se aplica a un archivo de datos mediante el comando `FILE LABEL`.

## Resultados

Figura 7-12  
Asistente para la copia de propiedades de datos: Paso 5



El último paso del Asistente para la copia de propiedades de datos proporciona información sobre el número de variables para las que se van a copiar las propiedades de variable del archivo de datos de origen, el número de nuevas variables que se van a crear y el número de propiedades (de archivo) de conjunto de datos que se van a copiar.

También puede pegar la sintaxis de comandos generada en una ventana de sintaxis y guardarla para su posterior uso.

## Identificación de casos duplicados

Puede haber distintos motivos por los que haya casos “duplicados” en los datos, entre ellos:

- Errores en la entrada de datos si por accidente se introduce el mismo caso más de una vez.
- Casos múltiples que comparten un valor de identificador primario común pero tienen valores diferentes de un identificador secundario, como los miembros de una familia que viven en el mismo domicilio.
- Casos múltiples que representan el mismo caso pero con valores diferentes para variables que no sean las que identifican el caso, como en el caso de varias compras realizadas por la misma persona o empresa de diferentes productos o en diferentes momentos.

La identificación de los casos duplicados le permite definir prácticamente como quiera lo que se considera *duplicado* y le proporciona cierto control sobre la determinación automática de los casos primarios frente a los duplicados.

### Para identificar y marcar los casos duplicados

- ▶ Elija en los menús:  
Datos  
Identificar casos duplicados...
- ▶ Seleccione una o varias variables que identifiquen los casos coincidentes.
- ▶ Seleccione una o varias de las opciones del grupo Crear variables.  
  
Si lo desea, puede:
  - ▶ Seleccionar una o varias variables para ordenar los casos dentro de los bloques definidos por las variables seleccionadas de casos coincidentes. El orden definido por estas variables determina el “primer” y el “último” caso de cada bloque. En caso contrario, se utilizará el orden del archivo original.
  - ▶ Filtrar automáticamente los casos duplicados de manera que no se incluyan en los informes, los gráficos o los cálculos de estadísticos.

Figura 7-13  
Cuadro de diálogo Identificar casos duplicados

**Definir casos coincidentes por.** Los casos se consideran duplicados si sus valores coinciden para *todas* las variables seleccionadas. Si desea identificar únicamente aquellos casos que coincidan al 100% en todos los aspectos, seleccione todas las variables.

**Ordenar dentro de los bloques coincidentes por.** Los casos se ordenan automáticamente por las variables que definen los casos coincidentes. Puede seleccionar otras variables de ordenación que determinarán el orden secuencial de los casos en cada bloque de coincidencia.

- Para cada variable de ordenación, el orden puede ser ascendente o descendente.
- Si selecciona más de una variable de ordenación, los casos se ordenarán por cada variable dentro de las categorías de la variable anterior de la lista. Por ejemplo, si selecciona *fecha* como la primera variable de ordenación y *cantidad* como la segunda, los casos se ordenarán por cantidad dentro de cada fecha.
- Utilice los botones de flecha hacia arriba y hacia abajo que hay a la derecha de la lista para cambiar el orden de las variables.
- El orden determina el “primer” y el “último” caso de cada bloque de coincidencia, que determina el valor de la variable indicador del caso primario opcional. Por ejemplo, si desea descartar todos los casos salvo el más reciente de cada bloque de coincidencia, puede ordenar los casos del bloque en orden ascendente por una variable de fecha, lo cual haría que la fecha más reciente fuese la última fecha del bloque.

**Variable indicador de casos primarios.** Crea una variable con un valor de 1 para todos los casos únicos y para el caso identificado como caso primario en cada bloque de casos coincidentes y un valor de 0 para los duplicados no primarios de cada bloque.

- El caso primario puede ser el primer o el último caso de cada bloque de coincidencia, según determine el orden del bloque de coincidencia. Si no especifica ninguna variable de ordenación, el orden del archivo original determina el orden de los casos dentro de cada bloque.
- Puede utilizar la variable indicador como una **variable de filtro** para excluir los duplicados que no sean primarios de los informes y los análisis sin eliminar dichos casos del archivo de datos.

**Recuento secuencial de casos coincidentes en cada bloque.** Crea una variable con un valor secuencial de 1 a  $n$  para los casos de cada bloque de coincidencia. La secuencia se basa en el orden actual de los casos de cada bloque, que puede ser el orden del archivo original o el orden determinado por las variables de ordenación especificadas.

**Mover los casos coincidentes a la parte superior del archivo.** Ordena el archivo de datos de manera que todos los bloques de casos coincidentes estén en la parte superior del archivo de datos, facilitando la inspección visual de los casos coincidentes en el Editor de datos.

**Mostrar tabla de frecuencias de las variables creadas.** Las tablas de frecuencias contienen los recuentos de cada valor de las variables creadas. Por ejemplo, para la variable de indicador de caso primario, la tabla mostraría tanto el número de casos con un valor de 0 en esa variable, que indica el número de duplicados, como el número de casos con un valor de 1 para esa variable, que indica el número de casos únicos y primarios.

**Valores perdidos.** En el caso de variables numéricas, los valores perdidos del sistema se tratan como cualquier otro valor: los casos que tengan el valor perdido del sistema para una variable de identificación se tratarán como si tuviesen valores coincidentes para dicha variable. En el caso de variables de cadena, los casos que no tengan ningún valor para una variable de identificación se tratarán como si tuviesen valores coincidentes para dicha variable.

## Agrupación visual

La agrupación visual está concebida para ayudarle en el proceso de creación de nuevas variables basadas en la agrupación de los valores contiguos de las variables existentes para dar lugar a un número limitado de categorías diferentes. Puede utilizar la agrupación visual para:

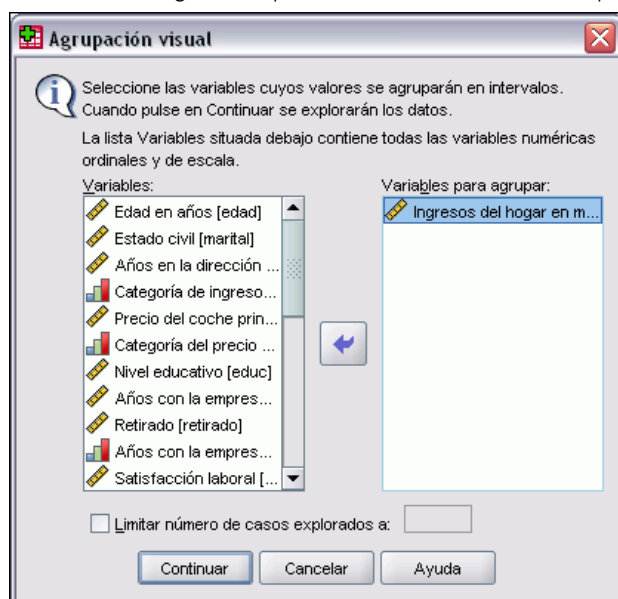
- Crear variables categóricas a partir de variables de escala continuas. Por ejemplo, puede utilizar una variable de escala con los ingresos para crear una variable categórica nueva que contenga intervalos de ingresos.
- Colapsar un número elevado de categorías ordinales en un conjunto menor de categorías. Por ejemplo, es posible colapsar una escala de evaluación de nueve categorías en tres categorías que representen: bajo, medio y alto.

En el primer paso, puede:

- Seleccione las variables numéricas de escala u ordinales para las que desee crear nuevas variables categóricas (en agrupaciones).

Figura 7-14

Cuadro de diálogo inicial para seleccionar las variables que se van a agrupar



Como alternativa, puede limitar la cantidad de casos que se van a explorar. Con los archivos de datos que contengan un elevado número de casos, la limitación del número de casos que se va a explorar puede ahorrar tiempo, pero debe evitarse este procedimiento en lo posible, ya que afectará a la distribución de los valores que se utilizarán en los cálculos posteriores en la agrupación visual.

*Nota:* Las variables de cadena y las variables numéricas nominales no se muestran en la lista de variables origen. La agrupación visual requiere que las variables sean numéricas, medidas bien a nivel ordinal o de escala, puesto que supone que los valores de los datos representan algún tipo de orden lógico que se puede utilizar para agrupar los valores con sentido. Puede cambiar el nivel de medida de una variable en la Vista de variables del Editor de datos. Si desea obtener más información, consulte [Nivel de medida de variable](#) en Capítulo 5 en p. 83.

## Para agrupar variables

- ▶ Seleccione en los menús de la ventana Editor de datos:  
Transformar  
Agrupación visual...
- ▶ Seleccione las variables numéricas de escala u ordinales para las que desee crear nuevas variables categóricas (en agrupaciones).
- ▶ Seleccione una variable de la Lista de variables exploradas.
- ▶ Escriba el nombre de la nueva variable agrupada. Los nombres de variable no pueden repetirse, y deben seguir las normas de denominación de variables. Si desea obtener más información, consulte [Nombres de variable](#) en Capítulo 5 en p. 82.
- ▶ Defina los criterios de agrupación para la nueva variable. Si desea obtener más información, consulte [Agrupación de variables](#) en p. 128.
- ▶ Pulse en Aceptar.

## Agrupación de variables

Figura 7-15  
Agrupación visual, cuadro de diálogo principal

**Agrupación visual**

Lista de variables exploradas:

M	Variable
	Ingresos del hogar en miles

Variable actual: ingres      Etiqueta: Ingresos del hogar en miles

Variable agrupada: ingcat2      Etiqueta: Ingresos del hogar en miles (agrupado)

Mínimo: 9,00      Valores no perdidos      Máximo: 1116,00

Rejilla:

	Valor	Etiqueta
1	25,00	<=25,00
2	50,00	25,00 - 49,00
3	75,00	50,00 - 74,00
4	SUPERIOR	75,00+
5		

Límites superiores:  
 Incluidos (<=)  
 Excluidos (<)

Crear puntos de corte...  
 Crear etiquetas  
 Invertir escala

Casos explorados: 6400  
 Valores perdidos: 0

Copiar intervalos:

Aceptar    Pegar    Restablecer    Cancelar    Ayuda



El cuadro de diálogo principal de la agrupación visual proporciona la siguiente información sobre las variables exploradas:

**Lista de variables exploradas.** Muestra las variables que fueron seleccionadas en el cuadro de diálogo inicial. Puede ordenar la lista por el nivel de medida (de escala u ordinal) o por la etiqueta o el nombre de variable, pulsando en los encabezados de las columnas.

**Casos explorados.** Indica el número de casos explorados. Todos los casos explorados sin valores perdidos definidos por el usuario o del sistema para la variable seleccionada, se usan en la generación de la distribución de valores que se emplea en los cálculos de la agrupación visual, incluyendo el histograma que se visualiza en el cuadro de diálogo principal y los puntos de corte basados en percentiles o unidades de desviación típica.

**Valores perdidos.** Indica el número de casos explorados con valores perdidos definidos por el usuario y perdidos del sistema. Los valores perdidos no se incluyen en ninguno de las categorías agrupadas. Si desea obtener más información, consulte [Valores perdidos definidos por el usuario en la agrupación visual](#) en p. 135.

**Variable actual.** El nombre y etiqueta de variable (si existe) de la variable actualmente seleccionada y que se usará como base para la nueva variable agrupada.

**Variable agrupada.** Nombre y etiqueta de variable alternativa para la nueva variable agrupada.

- **Nombre.** Debe introducir un nombre para la nueva variable. Los nombres de variable no pueden repetirse, y deben seguir las normas de denominación de variables. Si desea obtener más información, consulte [Nombres de variable](#) en Capítulo 5 en p. 82.
- **Etiqueta.** Puede introducir una etiqueta de variable descriptiva con una longitud de hasta 255 caracteres. La etiqueta de variable por defecto será la etiqueta de variable (si existe) o el nombre de variable de la variable origen con el texto (*Agrupada*) añadido al final de la etiqueta.

**Mínimo y Máximo.** Valores mínimo y máximo para la variable seleccionada actualmente, basados en los casos explorados y excluyendo los valores definidos como perdidos por el usuario.

**Valores no perdidos.** El histograma muestra la distribución de valores no perdidos correspondiente a la variable seleccionada actualmente, basándose en los casos explorados.

- Después de haber definido los intervalos para la nueva variable, se mostrarán líneas verticales en el histograma para indicar los puntos de corte que definen los intervalos.
- Puede pulsar y arrastrar las líneas de los puntos de corte a distintos puntos del histograma, modificando así la amplitud de los intervalos.
- Puede eliminar intervalos arrastrando las líneas de los puntos de corte fuera del histograma.

*Nota:* El histograma (que muestra valores no perdidos), el mínimo y el máximo se basan en los casos explorados. Si no incluye todos los casos en la exploración, es posible que no se refleje con precisión la distribución real, sobre todo si el archivo de datos se ordenó según la variable seleccionada. Si no explora ningún caso, no encontrará disponible información sobre la distribución de valores.

**Cuadrícula.** Muestra los valores que definen los puntos de corte superiores de cada intervalo, así como las etiquetas de valor opcionales para cada intervalo.

- **Valor.** Valores que definen los puntos de corte superiores en cada intervalo. Puede introducir los valores o utilizar **Crear puntos de corte** para crear automáticamente los intervalos basándose en los criterios seleccionados. Por defecto, se incluye automáticamente un punto de corte con el valor *SUPERIOR*. Este intervalo contendrá cualesquiera valores no perdidos por encima de los restantes puntos de corte. El intervalo definido por el punto de corte inferior incluirá todos los valores no perdidos que sean menores o iguales que dicho valor (o, sencillamente, inferiores a ese valor, dependiendo de la forma en que haya definido los puntos de corte superiores).
- **Etiqueta.** Etiquetas opcionales y descriptivas de los valores de la nueva variable agrupada. Puesto que los valores de la nueva variable sólo serán números enteros en secuencia, del 1 a *n*, las etiquetas que describan lo que representan los valores pueden resultar muy útiles. Puede introducir las etiquetas o usar **Crear etiquetas** para crear las etiquetas de valor de forma automática.

#### ***Para eliminar un intervalo de la rejilla***

- ▶ Pulse con el botón derecho en las casillas *Valor* o *Etiqueta* del intervalo.
- ▶ En el menú contextual emergente, seleccione **Eliminar fila**.

*Nota:* Si elimina el intervalo *SUPERIOR*, los casos con valores superiores al valor del último punto de corte especificado recibirán el valor perdido del sistema en la nueva variable.

#### ***Para eliminar todas las etiquetas o todos los intervalos definidos***

- ▶ Pulse en cualquier parte de la rejilla con el botón derecho del ratón.
- ▶ En el menú contextual emergente, seleccione **Eliminar todas las etiquetas** o **Eliminar todos los puntos de corte**.

**Límites superiores.** Controla el tratamiento de los valores de los límites superiores introducidos en la columna *Valor* de la rejilla.

- **Incluidos (<=).** Los casos con el valor especificado en la casilla *Valor* se incluyen en la categoría agrupada. Por ejemplo, si especifica los valores 25, 50 y 75, los casos con el valor exacto 25 se incluirán en el primer intervalo, ya que se incluirán todos los casos con valor menor o igual que 25.
- **Excluido (<).** Los casos con el valor especificado en la casilla *Valor* no se incluyen en la categoría agrupada. Por el contrario, se incluyen en el siguiente intervalo. Por ejemplo, si especifica los valores 25, 50 y 75, los casos con el valor exacto 25 se incluirán en el segundo intervalo en vez de en el primero, puesto que el primero sólo contendrá casos con valores inferiores a 25.

**Crear puntos de corte.** Genera categorías agrupadas automáticamente para crear intervalos de igual amplitud, intervalos con el mismo número de casos o intervalos basados en un número de desviaciones típicas. Esta posibilidad no está disponible si no se ha explorado ningún caso. Si desea obtener más información, consulte [Generación automática de categorías agrupadas](#) en p. 131.

**Crear etiquetas.** Genera etiquetas descriptivas para los valores enteros consecutivos contenidos en la nueva variable agrupada, en función de los valores de la rejilla y el tratamiento especificado para los límites superiores (incluidos o excluidos).

**Invertir la escala.** Por defecto, los valores de la nueva variable agrupada serán números enteros consecutivos, del 1 a  $n$ . La inversión de la escala convierte los valores en números enteros consecutivos, de  $n$  a 1.

**Copiar intervalos.** Puede copiar las especificaciones de agrupación de otra variable a la variable seleccionada en ese momento, o desde la variable seleccionada en ese momento a otras varias variables. Si desea obtener más información, consulte [Copia de categorías agrupadas](#) en p. 133.

### ***Generación automática de categorías agrupadas***

El cuadro de diálogo Crear puntos de corte permite la creación automática de categorías agrupadas en función de los criterios seleccionados.

#### ***Para utilizar el cuadro de diálogo Crear puntos de corte***

- ▶ Seleccione (pulse) una variable de la Lista de variables exploradas.
- ▶ Pulse Crear puntos de corte.
- ▶ Seleccione los criterios de generación de los puntos de corte que definirán las categorías agrupadas.
- ▶ Pulse en Aplicar.

Figura 7-16  
Cuadro de diálogo *Crear puntos de corte*

*Nota:* El cuadro de diálogo *Crear puntos de corte* no está disponible si no se ha explorado ningún caso.

**Intervalos de igual amplitud.** Genera categorías agrupadas de igual amplitud (por ejemplo, 1–10, 11–20 y 21–30), basándose en dos (cualquiera) de los tres criterios siguientes:

- **Posición del primer punto de corte.** Valor que define el límite superior de la categoría agrupada inferior (por ejemplo, el valor 10 indica un intervalo que incluya todos los valores hasta 10).
- **Número de puntos de corte.** El número de categorías agrupadas es el número de puntos de corte *más uno*. Por ejemplo, 9 puntos de corte generan 10 categorías agrupadas.
- **Amplitud.** La amplitud de cada intervalo. Por ejemplo, el valor 10 agrupará la variable *Edad en años* en intervalos de 10 años.

**Percentiles iguales basados en los casos explorados.** Genera categorías agrupadas con un número igual de casos en cada intervalo (utilizando el **algoritmo “aempirical”** para el cálculo de percentiles), según uno de los criterios siguientes:

- **Número de puntos de corte.** El número de categorías agrupadas es el número de puntos de corte *más uno*. Por ejemplo, tres puntos de corte generan cuatro intervalos percentiles (cuartiles), conteniendo cada uno el 25% de los casos.
- **% de casos.** Amplitud de cada intervalo, expresado en forma de porcentaje sobre el número total de casos. Por ejemplo, el valor 33,3 generaría tres categorías agrupadas (dos puntos de corte), conteniendo cada una el 33,3% de los casos.

Si la variable origen contiene un número relativamente pequeño de valores distintos o un gran número de casos con el mismo valor, es posible que obtenga menos intervalos que las solicitadas. En caso de haber varios valores idénticos en un punto de corte, todos se incluyen en el mismo intervalo; por consiguiente, los porcentajes reales pueden no ser siempre iguales.

**Puntos de corte en media y desviaciones típicas seleccionadas, basadas en casos explorados.**

Genera categorías agrupadas basándose en los valores de la media y la desviación típica de la distribución de la variable.

- Si no selecciona ninguno de los intervalos de desviación típica, se crearán dos categorías agrupadas, siendo la media el punto de corte que divide los intervalos.
- Puede seleccionar cualquier combinación de los intervalos de desviación típica, basándose en una, dos o tres desviaciones típicas. Por ejemplo, al seleccionar las tres opciones se obtendrán ocho categorías agrupadas: seis intervalos distanciados en una desviación típica de amplitud y dos intervalos para los casos que se encuentren a más de tres desviaciones típicas por encima y por debajo de la media.

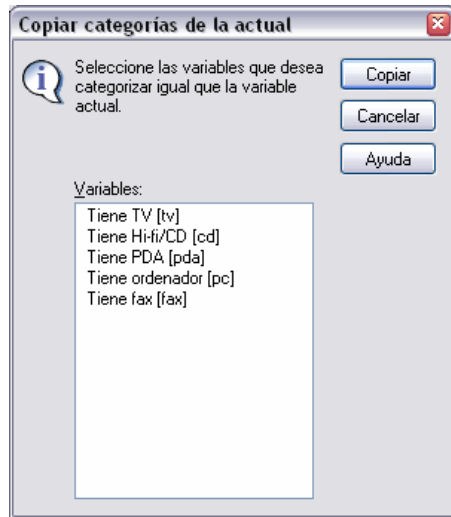
En una distribución normal, el 68% de los casos se encuentra dentro de una distancia de una desviación típica respecto a la media, el 95% entre dos desviaciones típicas y el 99% dentro de tres desviaciones típicas. La creación de categorías agrupadas basadas en desviaciones típicas puede ocasionar que algunos intervalos queden definidos fuera del rango real de los datos, e incluso fuera del rango de valores posibles de los datos (por ejemplo, un rango de salarios negativos).

*Nota:* Los cálculos de los percentiles y las desviaciones típicas se basan en los casos explorados. Si limita el número de casos explorados, puede que los intervalos resultantes no incluyan la proporción de casos deseada en dichos intervalos, sobre todo si el archivo de datos se ordenó según la variable origen. Por ejemplo, si limita la exploración a los primeros 100 casos de un archivo de datos con 1000 casos y el archivo de datos está ordenado en orden descendente por edad del encuestado, en lugar de cuatro intervalos percentiles de la edad, cada uno con el 25% de los casos, podría encontrarse con que los tres primeros intervalos contuvieran cada uno sólo en torno al 3,3% de los casos, mientras que el último intervalo albergaría el 90% de los casos.

### ***Copia de categorías agrupadas***

Al crear categorías agrupadas para una o más variables, puede copiar las especificaciones de agrupación de otra variable a la seleccionada en ese momento o desde la variable seleccionada en ese momento a varias otras variables.

**Figura 7-17**  
*Copia de intervalos en la variable actual o desde ella*



***Para copiar especificaciones de intervalos***

- ▶ Defina las categorías agrupadas para una variable como mínimo; pero *no* pulse en Aceptar ni en Pegar.
  - ▶ Seleccione (pulse) una variable de la Lista de variables exploradas para la cual haya definido categorías agrupadas.
  - ▶ Pulse A otras variables.
  - ▶ Seleccione las variables para las que desea crear nuevas variables con las mismas categorías agrupadas.
  - ▶ Pulse Copiar.
- o*
- ▶ Seleccione (pulse) una variable de la Lista de variables exploradas sobre la cual desea copiar categorías agrupadas ya definidas.
  - ▶ Pulse De otra variable.
  - ▶ Seleccione la variable que contiene las categorías agrupadas definidas que desea copiar.
  - ▶ Pulse Copiar.

También se copiarán las etiquetas de valor si se especificaron en la variable cuyas especificaciones de agrupación se van a copiar.

*Nota:* Una vez que haya pulsado en Aceptar en el cuadro de diálogo principal de la agrupación visual, para crear nuevas variables agrupadas (o cerrado el cuadro de diálogo de alguna otra forma), no podrá usar de nuevo la agrupación visual para copiar dichas categorías agrupadas en otras variables.

### **Valores perdidos definidos por el usuario en la agrupación visual**

Los valores perdidos definidos por el usuario (valores identificados como los códigos para los datos perdidos) para la variable origen no se incluyen en las categorías agrupadas de la nueva variable. Los valores perdidos definidos por el usuario de las variables se copian como valores perdidos definidos por el usuario en la nueva variable, copiándose también cualquier otra etiqueta de valor definida para los códigos de los valores perdidos.

Si un código de valor perdido entra en conflicto con alguno de los valores de categorías agrupadas de la nueva variable, el código de valor perdido de la nueva variable se recodificará a un valor no conflictivo, sumando 100 al valor de categoría agrupada superior. Por ejemplo, si el usuario define el valor 1 como valor perdido para la variable origen y la nueva variable va a contar con seis categorías agrupadas, cualquier caso con el valor 1 en la variable origen tendrá el valor 106 en la nueva variable, y 106 será definido como un valor definido como perdido por el usuario. Si el valor definido como perdido por el usuario en la variable de origen tenía definida una etiqueta de valor, dicha etiqueta se mantendrá como etiqueta de valor para el valor recodificado de la nueva variable.

*Nota:* Si la variable de origen tiene definido un rango de valores perdidos de usuario con la forma *MENOR-n*, donde *n* es un número positivo, los valores perdidos definidos por el usuario correspondientes a la nueva variable, serán números negativos.

# ***Transformaciones de los datos***

En una situación ideal, los datos brutos son perfectamente apropiados para el tipo de análisis que se desea realizar y cualquier relación existente entre las variables o es adecuadamente lineal o es claramente ortogonal. Desafortunadamente, esto ocurre pocas veces. El análisis preliminar puede revelar esquemas de codificación poco prácticos o errores de codificación, o bien pueden requerirse transformaciones de los datos para exponer la verdadera relación existente entre las variables.

Puede realizar transformaciones de los datos de todo tipo, desde tareas sencillas, como la agrupación de categorías para su análisis posterior, hasta otras más avanzadas, como la creación de nuevas variables basadas en ecuaciones complejas e instrucciones condicionales.

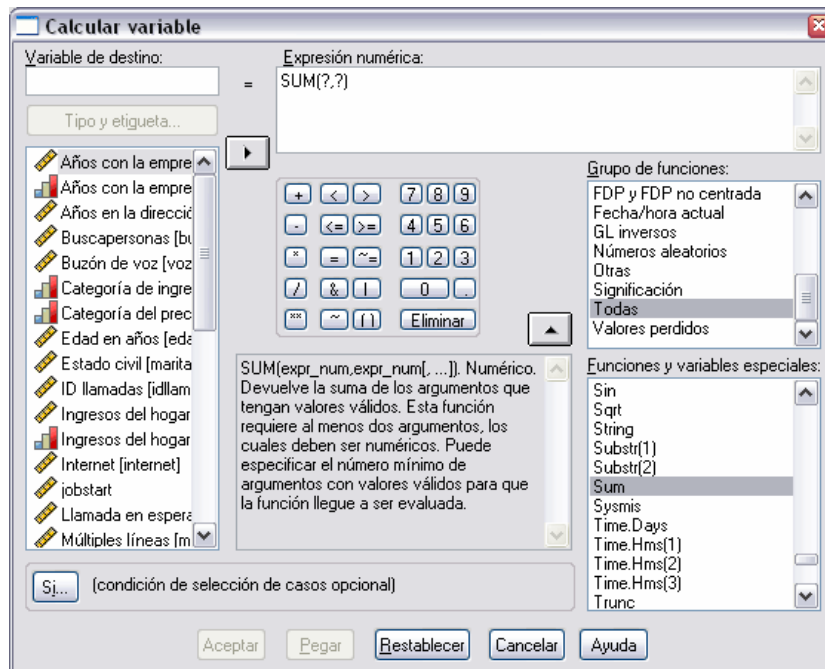
## ***Cálculo de variables***

Utilice el cuadro de diálogo Calcular para calcular los valores de una variable basándose en transformaciones numéricas de otras variables.

- Puede calcular valores para las variables numéricas o de cadena (alfanuméricas).
- Puede crear nuevas variables o bien reemplazar los valores de las variables existentes. Para las nuevas variables, también se puede especificar el tipo y la etiqueta de variable.
- Puede calcular valores de forma selectiva para subconjuntos de datos basándose en condiciones lógicas.
- Puede utilizar más de 70 funciones preincorporadas, incluyendo funciones aritméticas, funciones estadísticas, funciones de distribución y funciones de cadena.



Figura 8-1  
Cuadro de diálogo *Calcular variable*



### Para calcular variables

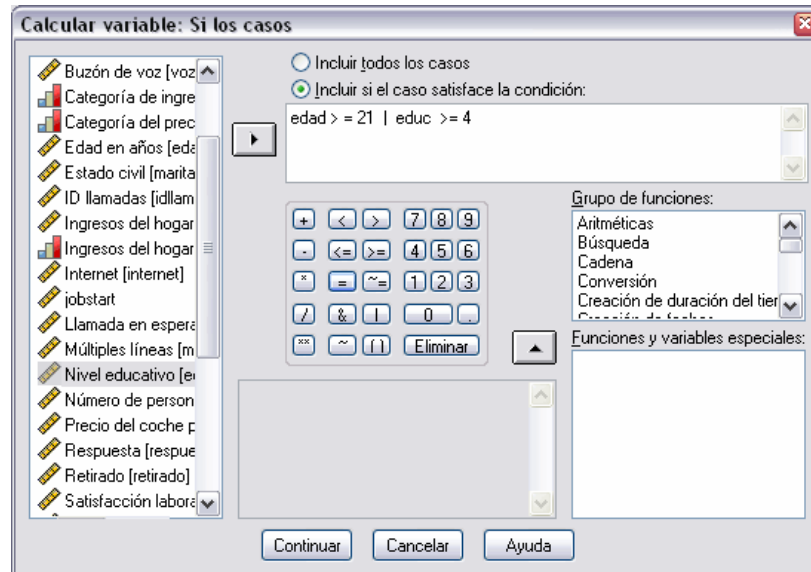
- ▶ Elija en los menús:  
Transformar  
Calcular variable...
- ▶ Escriba el nombre de una sola variable de destino. Puede ser una variable existente o una nueva que se vaya a añadir al conjunto de datos activo.
- ▶ Para crear una expresión, puede pegar los componentes en el campo Expresión o escribir directamente en dicho campo.
  - Puede pegar las funciones o las variables de sistema utilizadas habitualmente seleccionando un grupo de la lista Grupo de funciones y pulsando dos veces en la función o variable de las listas de funciones y variables especiales (o seleccione la función o variable y pulse en la flecha que se encuentra sobre la lista Grupo de funciones). Rellene los parámetros indicados mediante interrogaciones (aplicable sólo a las funciones). El grupo de funciones con la etiqueta Todo contiene una lista de todas las funciones y variables de sistema disponibles. En un área reservada del cuadro de diálogo se muestra una breve descripción de la función o variable actualmente seleccionada.
  - Las constantes de cadena deben ir entre comillas o apóstrofes.
  - Si los valores contienen decimales, debe utilizarse un punto (.) como separador decimal.
  - Para las nuevas variables de cadena, también deberán seleccionar Tipo y etiqueta para especificar el tipo de datos.

## Calcular variable: Si los casos

El cuadro de diálogo Si los casos permite aplicar transformaciones de los datos para subconjuntos de casos seleccionados utilizando expresiones condicionales. Una expresión condicional devuelve un valor *verdadero*, *falso* o *perdido* para cada caso.

Figura 8-2

Cuadro de diálogo Calcular variable: Si los casos



- Si el resultado de una expresión condicional es *verdadero*, se incluirá el caso en el subconjunto seleccionado.
- Si el resultado de una expresión condicional es *falso* o *perdido*, no se incluirá el caso en el subconjunto seleccionado.
- La mayoría de las expresiones condicionales utilizan al menos uno de los seis operadores de relación (<, >, <=, >=, =, and ~=) de la calculadora.
- Las expresiones condicionales pueden incluir nombres de variable, constantes, operadores aritméticos, funciones numéricas (y de otros tipos), variables lógicas y operadores de relación.

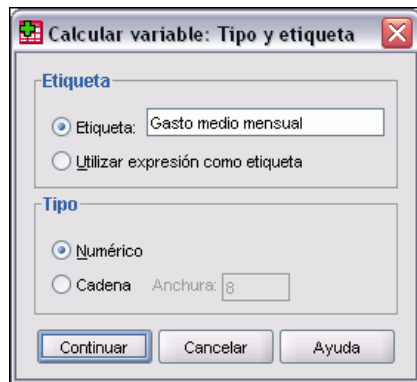
## Calcular variable: Tipo y etiqueta

Por defecto, las nuevas variables calculadas son numéricas. Para calcular una nueva variable de cadena, deberá especificar el tipo de los datos y su ancho.

**Etiqueta.** Etiqueta de variable descriptiva y opcional de hasta 255 bytes. Puede escribir el texto de una etiqueta o bien utilizar los primeros 110 caracteres de la expresión de cálculo.

**Tipo.** Las variables calculadas pueden ser numéricas o de cadena (alfanuméricas). Las variables de cadena no se pueden utilizar en cálculos aritméticos.

Figura 8-3  
Cuadro de diálogo Tipo y etiqueta



## Funciones

Se dispone de muchos tipos de funciones, entre ellos:

- Funciones aritméticas
- Funciones estadísticas
- Funciones de cadena
- Funciones de fecha y hora
- Funciones de distribución
- Funciones de variables aleatorias
- Funciones de valores perdidos
- Funciones de puntuación (sólo servidor de SPSS)

Si desea obtener más información y una descripción detallada de cada función, escriba funciones en la pestaña Índice del sistema de ayuda.

## Valores perdidos en funciones

Las funciones y las expresiones aritméticas sencillas tratan los valores perdidos de diferentes formas. En la expresión:

```
(var1+var2+var3)/3
```

El resultado es el valor perdido si un caso tiene un valor perdido para cualquiera de las tres variables.

En la expresión:

```
MEAN(var1, var2, var3)
```

El resultado es el valor perdido sólo si el caso tiene valores perdidos para las tres variables.

En las funciones estadísticas se puede especificar el número mínimo de argumentos que deben tener valores no perdidos. Para ello, escriba un punto y el número mínimo de argumentos después del nombre de la función, como en:

```
MEAN.2(var1, var2, var3)
```

## Generadores de números aleatorios

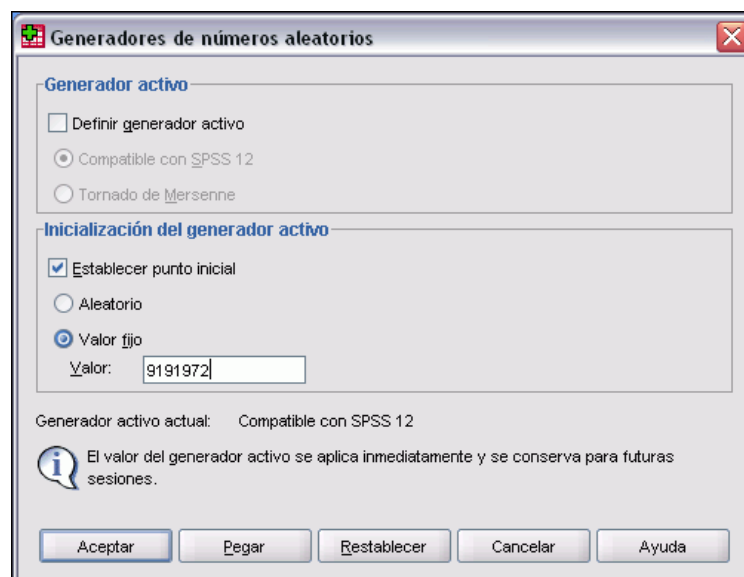
El cuadro de diálogo Generadores de números aleatorios le permite seleccionar el generador de números aleatorios y establecer el valor de secuencia de inicio de modo que pueda reproducir una secuencia de números aleatorios.

**Generador activo.** Hay dos generadores de números aleatorios disponibles:

- **Compatible con la versión 12.** El generador de números aleatorios utilizado en la versión 12 y versiones anteriores. Utilice este generador de números aleatorios si necesita reproducir los resultados aleatorizados generados por versiones previas basadas en una semilla de aleatorización especificada.
- **Tornado de Mersenne.** Un generador de números aleatorios nuevo que es más fiable en los procesos de simulación. Utilice este generador de números aleatorios si no es necesario reproducir resultados aleatorizados correspondientes a la versión 12 o anteriores.

**Inicialización del generador activo.** La semilla de aleatorización cambia cada vez que se genera un número aleatorio para utilizarlo en las transformaciones (como las funciones de distribución aleatorias), el muestreo aleatorio o la ponderación de los casos. Para replicar una secuencia de números aleatorios, establezca el valor de inicialización del punto de inicio antes de cada análisis que utilice los números aleatorios. El valor debe ser un entero positivo.

Figura 8-4  
Cuadro de diálogo Generadores de números aleatorios



Para seleccionar el generador de números aleatorios y establecer el valor de inicialización:

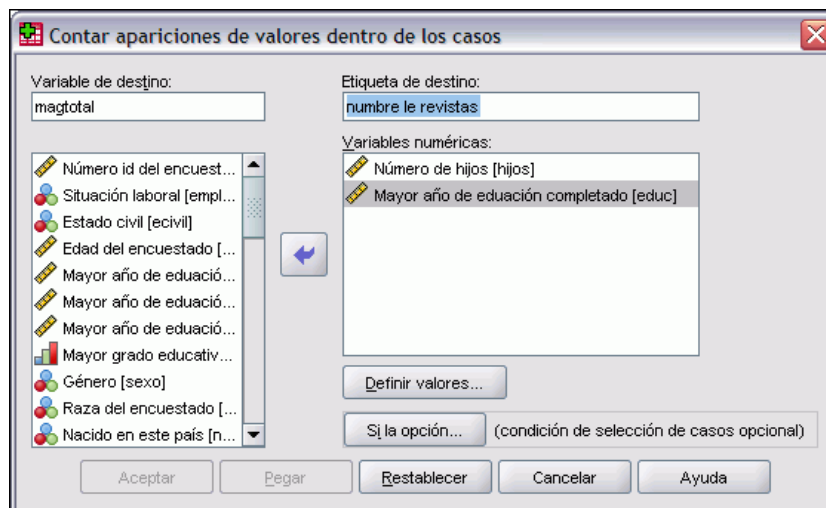
- ▶ Elija en los menús:
  - Transformar
  - Generadores de números aleatorios

## **Contar apariciones de valores dentro de los casos**

Este cuadro de diálogo crea una variable que, para cada caso, cuenta las apariciones del mismo valor, o valores, en una lista de variables. Por ejemplo, un estudio podrá contener una lista de revistas con las casillas de verificación *sí/no* para indicar qué revistas lee cada encuestado. Se podría contar el número de respuestas *sí* de cada encuestado para crear una nueva variable que contenga el número total de revistas leídas.

Figura 8-5

Cuadro de diálogo Contar apariciones de valores dentro de los casos



### **Para contar apariciones de valores dentro de los casos**

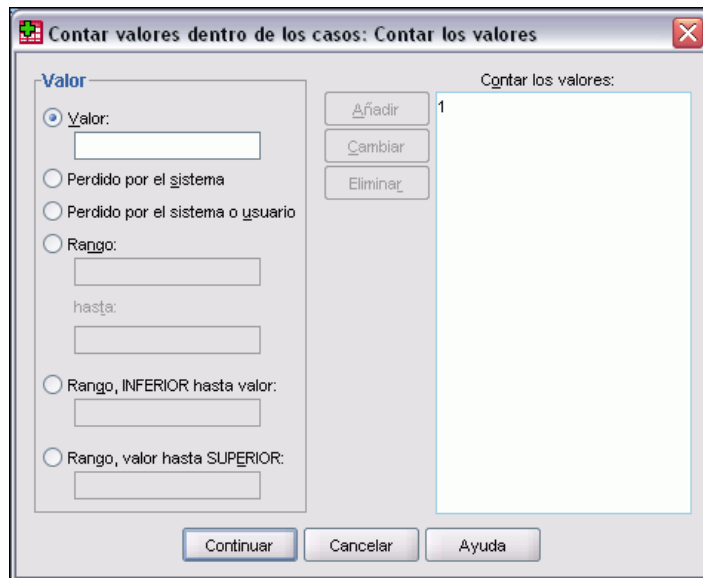
- ▶ Elija en los menús:
    - Transformar
    - Contar valores dentro de los casos...
  - ▶ Introduzca el nombre de la variable de destino.
  - ▶ Seleccione dos o más variables del mismo tipo (numéricas o de cadena).
  - ▶ Pulse en Definir valores y especifique los valores que se deben contar.
- Si lo desea, puede definir un subconjunto de casos en los que contar las apariciones de valores.

### **Contar valores dentro de los casos: Contar los valores**

El valor de la variable de destino (en el cuadro de diálogo principal) se incrementa en 1 cada vez que una de las variables seleccionadas coincide con una especificación de la lista Contar los valores. Si un caso coincide con varias de las especificaciones en cualquiera de las variables, la variable de destino se incrementa varias veces para esa variable.

Las especificaciones de valores pueden incluir valores individuales, valores perdidos o valores perdidos del sistema y rangos de valores. Los rangos incluyen sus puntos finales y los valores definidos como perdidos por el usuario que estén dentro del rango.

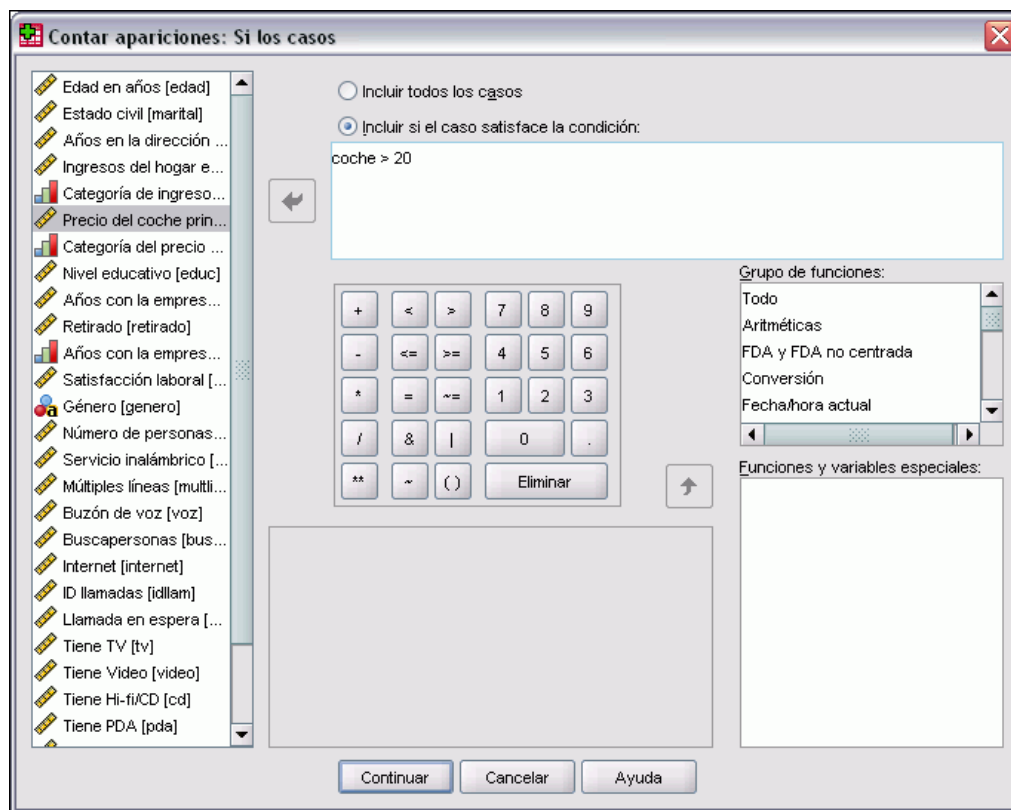
Figura 8-6  
Cuadro de diálogo Contar los valores



### **Contar apariciones: Si los casos**

El cuadro de diálogo Si los casos permite contar apariciones de valores para un subconjunto de casos seleccionado utilizando expresiones condicionales. Una expresión condicional devuelve un valor *verdadero*, *falso* o *perdido* para cada caso.

Figura 8-7  
Cuadro de diálogo Contar apariciones: Si los casos



Si desea conocer las consideraciones generales sobre el uso del cuadro de diálogo Si los casos, consulte *Calcular variable: Si los casos* en p. 138.

## ***Recodificación de valores***

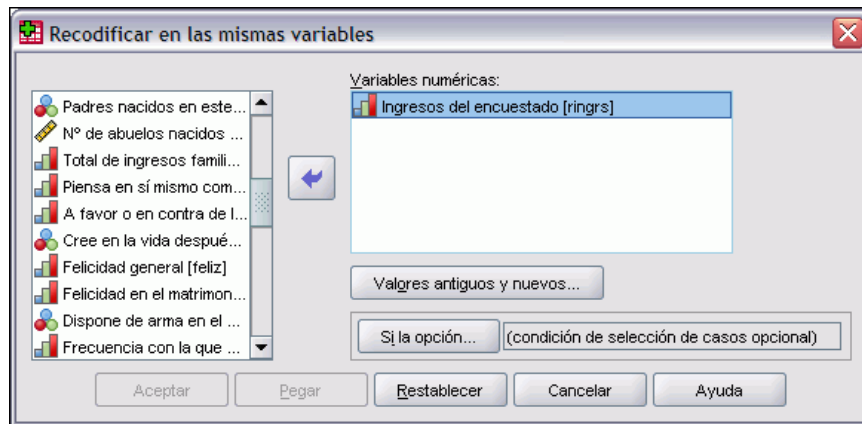
Los valores de datos se pueden modificar mediante la recodificación. Esto es particularmente útil para agrupar o combinar categorías. Puede recodificar los valores dentro de las variables existentes o crear variables nuevas que se basen en los valores recodificados de las variables existentes.

## ***Recodificar en las mismas variables***

El cuadro de diálogo *Recodificar en las mismas variables* le permite reasignar los valores de las variables existentes o agrupar rangos de valores existentes en nuevos valores. Por ejemplo, podría agrupar los salarios en categorías que sean rangos de salarios.

Puede recodificar las variables numéricas y de cadena. Si selecciona múltiples variables, todas deben ser del mismo tipo. No se pueden recodificar juntas las variables numéricas y de cadena.

Figura 8-8  
Cuadro de diálogo *Recodif. en mismas var.*



### **Para recodificar los valores de una variable**

- ▶ Elija en los menús:  
Transformar  
Recodificar en las mismas variables...
- ▶ Seleccione las variables que desee recodificar. Si selecciona múltiples variables, todas deberán ser del mismo tipo (numéricas o de cadena).
- ▶ Pulse en Valores antiguos y nuevos y especifique cómo deben recodificarse los valores.

Si lo desea, puede definir un subconjunto de los casos para su recodificación. El cuadro de diálogo Si los casos para esto es igual al que se describe para Contar apariciones.

### **Recodificar en las mismas variables: Valores antiguos y nuevos**

Este cuadro de diálogo permite definir los valores que se van a recodificar. Todas las especificaciones de valores deben pertenecer al mismo tipo de datos (numéricos o de cadena) que las variables seleccionadas en el cuadro de diálogo principal.

**Valor antiguo.** Determina el valor o los valores que se van a recodificar. Puede recodificar valores individuales, rangos de valores y valores perdidos. Los rangos y los valores perdidos del sistema no se pueden seleccionar para las variables de cadena, ya que ninguno de los conceptos es aplicable a estas variables. Los rangos incluyen sus puntos finales y los valores definidos como perdidos por el usuario que estén dentro del rango.

- **Valor.** Valor antiguo individual que se a recodificar en un valor nuevo. El tipo de datos (numérico o de cadena) del valor introducido debe coincidir con el tipo de datos de las variables desee recodificar.
- **Perdido por el sistema.** Valores asignados por SPSS cuando los valores de sus datos no están definidos de acuerdo al tipo de formato que haya especificado, cuando un campo numérico está vacío, o cuando no está definido un valor como resultado de un comando de transformación. Los valores numéricos perdidos del sistema se muestran como puntos. Las



variables de cadena no pueden tener valores perdidos del sistema, ya que es lícito cualquier carácter en las variables de cadena.

- **Perdido por el sistema o por el usuario.** Observaciones que tienen valores que, o bien han sido declarados como perdidos por el usuario o bien son desconocidos y se les ha asignado el valor perdido del sistema, lo cual se indica mediante un punto (.).
- **Amplitud.** Un rango inclusivo de valores. No está disponible para variables de cadena. Se incluirán cualesquiera valores definidos por el usuario como perdidos que se encuentren dentro del rango.
- **Todos los demás valores.** Cualquier valor no incluido en una de las especificaciones de la lista Antigo->Nuevo. Aparece en la lista Antigo->Nuevo como ELSE (otros).

**Valor nuevo.** Es el valor individual en el que se recodifica cada valor o rango de valores antiguo. Puede introducir un valor o asignar el valor perdido del sistema.

- **Valor.** Valor en el que se va a recodificar uno o más valores antiguos. El tipo de datos (numérico o de cadena) del valor introducido debe coincidir con el tipo de datos del valor antiguo.
- **Perdido por el sistema.** Recodifica el valor antiguo especificado como valor perdido por el sistema. El valor perdido por el sistema no se utiliza en los cálculos. Además, los casos con valores perdidos por el sistema se excluyen de muchos procedimientos. No está disponible para variables de cadena.

**Antigo->Nuevo.** Contiene la lista de especificaciones que se va a utilizar para recodificar la variable o las variables. Puede añadir, cambiar y borrar las especificaciones que desee. La lista se ordena automáticamente basándose en la especificación del valor antiguo y siguiendo este orden: valores únicos, valores perdidos, rangos y todos los demás valores. Si cambia una especificación de recodificación en la lista, el procedimiento volverá a ordenar la lista automáticamente, si fuera necesario, para mantener este orden.

Figura 8-9

Cuadro de diálogo Valores antiguos y nuevos

**Recodificar en las mismas variables: Valores antiguos y nuevos**

**Valor antiguo**

Valor:

Perdido por el sistema

Perdido por el sistema o usuario

Rango:

hasta

Rango, INFERIOR hasta valor:

Rango, valor hasta SUPERIOR:

Todos los demás valores

**Valor nuevo**

Valor:

Perdido por el sistema

Antigo --> Nuevo:

Añadir    Lowest thru 24999 --> 1

Cambiar    25000 thru 49999 --> 2

Eliminar    50000 thru 74999 --> 3

75000 thru Highest --> 4

Continuar    Cancelar    Ayuda

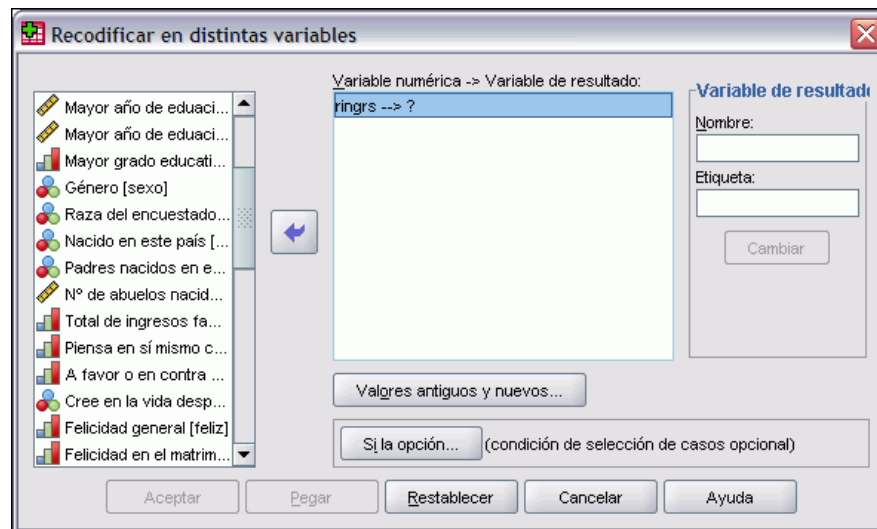
## Recodificar en distintas variables

El cuadro de diálogo Recodificar en distintas variables le permite reasignar los valores de las variables existentes o agrupar rangos de valores existentes en nuevos valores para una variable nueva. Por ejemplo, podría agrupar los salarios en una nueva variable que contenga categorías de rangos de salarios.

- Puede recodificar las variables numéricas y de cadena.
- Puede recodificar variables numéricas en variables de cadena y viceversa.
- Si selecciona múltiples variables, todas deben ser del mismo tipo. No se pueden recodificar juntas las variables numéricas y de cadena.

Figura 8-10

Cuadro de diálogo Recodif. en distintas var. (Recodificar en distintas variables)



### Para recodificar los valores de una variable en una nueva variable

- ▶ Elija en los menús:  
Transformar  
Recodificar en distintas variables...
- ▶ Seleccione las variables que desee recodificar. Si selecciona múltiples variables, todas deberán ser del mismo tipo (numéricas o de cadena).
- ▶ Introduzca el nombre de la nueva variable de resultado para cada nueva variable y pulse en Cambiar.
- ▶ Pulse en Valores antiguos y nuevos y especifique cómo deben recodificarse los valores.

Si lo desea, puede definir un subconjunto de los casos para su recodificación. El cuadro de diálogo Si los casos para esto es igual al que se describe para Contar apariciones.

## **Recodificar en distintas variables: Valores antiguos y nuevos**

Este cuadro de diálogo permite definir los valores que se van a recodificar.

**Valor antiguo.** Determina el valor o los valores que se van a recodificar. Puede recodificar valores individuales, rangos de valores y valores perdidos. Los rangos y los valores perdidos del sistema no se pueden seleccionar para las variables de cadena, ya que ninguno de los conceptos es aplicable a estas variables. Los valores antiguos deben ser del mismo tipo de datos (numéricos o de cadena) que la variable original. Los rangos incluyen sus puntos finales y los valores definidos como perdidos por el usuario que estén dentro del rango.

- **Valor.** Valor antiguo individual que se a recodificar en un valor nuevo. El tipo de datos (numérico o de cadena) del valor introducido debe coincidir con el tipo de datos de las variables desee recodificar.
- **Perdido por el sistema.** Valores asignados por SPSS cuando los valores de sus datos no están definidos de acuerdo al tipo de formato que haya especificado, cuando un campo numérico está vacío, o cuando no está definido un valor como resultado de un comando de transformación. Los valores numéricos perdidos del sistema se muestran como puntos. Las variables de cadena no pueden tener valores perdidos del sistema, ya que es lícito cualquier carácter en las variables de cadena.
- **Perdido por el sistema o por el usuario.** Observaciones que tienen valores que, o bien han sido declarados como perdidos por el usuario o bien son desconocidos y se les ha asignado el valor perdido del sistema, lo cual se indica mediante un punto (.).
- **Amplitud.** Un rango inclusivo de valores. No está disponible para variables de cadena. Se incluirán cualesquiera valores definidos por el usuario como perdidos que se encuentren dentro del rango.
- **Todos los demás valores.** Cualquier valor no incluido en una de las especificaciones de la lista Antigo->Nuevo. Aparece en la lista Antigo->Nuevo como ELSE (otros).

**Valor nuevo.** Es el valor individual en el que se recodifica cada valor o rango de valores antiguo. Los valores nuevos pueden ser numéricos o de cadena.

- **Valor.** Valor en el que se va a recodificar uno o más valores antiguos. El tipo de datos (numérico o de cadena) del valor introducido debe coincidir con el tipo de datos del valor antiguo.
- **Perdido por el sistema.** Recodifica el valor antiguo especificado como valor perdido por el sistema. El valor perdido por el sistema no se utiliza en los cálculos. Además, los casos con valores perdidos por el sistema se excluyen de muchos procedimientos. No está disponible para variables de cadena.
- **Copiar los valores antiguos.** Retiene el valor antiguo. Si algunos de los valores no requieren la recodificación, utilice esta opción para incluir los valores antiguos. Cualquier valor antiguo que no sea especificado no se incluye en la nueva variable, y los casos con esos valores se asignan al valor perdido por el sistema en la nueva variable.

**Las variables de resultado son cadenas.** Define la nueva variable recodificada como variable de cadena (alfanumérica). La variable antigua puede ser numérica o de cadena.

**Convertir cadenas numéricas en números.** Convierte los valores de cadena que contienen números en valores numéricos. A las cadenas que contengan cualquier cosa que no sean números y un carácter de signo opcional (+ ó -) se les asignará el valor perdido por el sistema.

**Antiguo→Nuevo.** Contiene la lista de especificaciones que se va a utilizar para recodificar la variable o las variables. Puede añadir, cambiar y borrar las especificaciones que desee. La lista se ordena automáticamente basándose en la especificación del valor antiguo y siguiendo este orden: valores únicos, valores perdidos, rangos y todos los demás valores. Si cambia una especificación de recodificación en la lista, el procedimiento volverá a ordenar la lista automáticamente, si fuera necesario, para mantener este orden.

Figura 8-11  
Cuadro de diálogo Valores antiguos y nuevos

## Asignar rangos a los casos

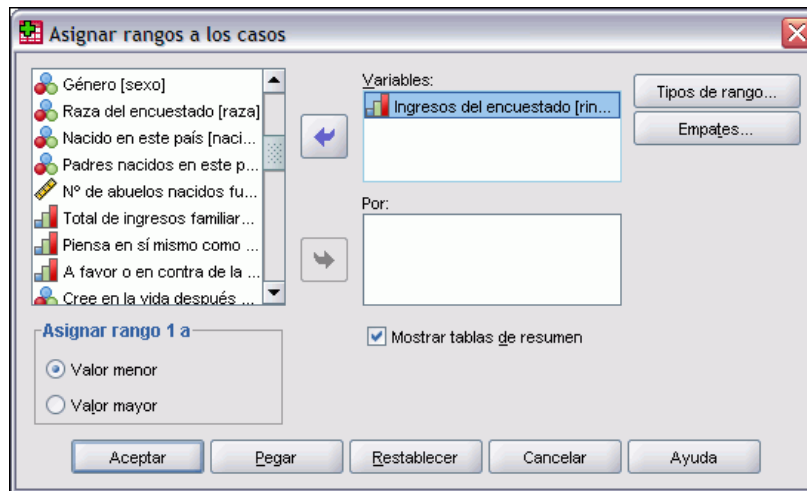
El cuadro de diálogo Asignar rangos a los casos le permite crear nuevas variables que contienen rangos, puntuaciones de Savage y normales, y los valores de los percentiles para las variables numéricas.

Los nombres de las nuevas variables y las etiquetas de variable descriptivas se generan automáticamente en función del nombre de la variable original y de las medidas seleccionadas. Una tabla de resumen presenta una lista de las variables originales, las nuevas variables y las etiquetas de variable.

Si lo desea, puede:

- Asignar los rangos a los casos en orden ascendente o descendente.
- Organizar los rangos en subgrupos seleccionando una o más variables de agrupación para la lista Por. Los rangos se calculan dentro de cada grupo, y los grupos se definen mediante la combinación de los valores de las variables de agrupación. Por ejemplo, si selecciona *sexo* y *minoría* como variables de agrupación, los rangos se calcularán para cada combinación de *sexo* y *minoría*.

Figura 8-12  
Cuadro de diálogo *Asignar rangos a los casos*



### **Para asignar rangos a los casos**

- ▶ Elija en los menús:  
Transformar  
Asignar rangos a casos...
- ▶ Seleccione la variable o variables a las que desee asignar los rangos. Sólo se pueden asignar rangos a las variables numéricas.

Si lo desea, puede asignar los rangos a los casos en orden ascendente o descendente y organizar los rangos por subgrupos.

### **Asignar rangos a los casos: Tipos**

Puede seleccionar diversos métodos de asignación de rangos. En cada método se crea una variable diferente de asignación de rangos. Los métodos de asignación de rangos incluyen rangos sencillos, puntuaciones de Savage, rangos fraccionales y percentiles. También puede crear rangos basados en estimaciones de la proporción y puntuaciones normales.

**Rango.** El rango simple. El valor de la nueva variable es igual a su rango.

**Puntuaciones de Savage.** La nueva variable contiene puntuaciones de Savage basadas en una distribución exponencial.

**Rango fraccional.** El valor de la nueva variable es igual al rango dividido por la suma de las ponderaciones de los casos no perdidos.

**Rango fraccional como porcentaje.** Cada rango se divide por el número de casos que tienen valores válidos y se multiplica por 100.

**Suma de ponderaciones de los casos.** El valor de la nueva variable es igual a la suma de las ponderaciones de los casos. La nueva variable es una constante para todos los casos del mismo grupo.

**Ntiles.** Los rangos se basan en los grupos percentiles, de forma que cada uno de los grupos contenga aproximadamente el mismo número de casos. Por ejemplo, con 4 Ntiles se asignará un rango 1 a los casos por debajo del percentil 25, 2 a los casos entre los percentiles 25 y 50, 3 a los casos entre los percentiles 50 y 75, y 4 a los casos por encima del percentil 75.

**Estimaciones de proporción.** Estimaciones de la proporción acumulada de la distribución que corresponde a un rango particular.

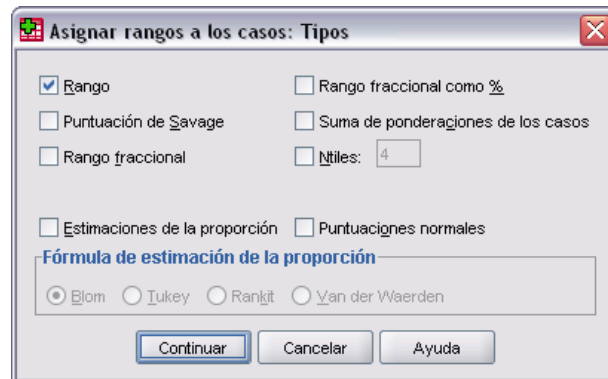
**Puntuaciones normales.** Puntuaciones z correspondientes a la proporción acumulada estimada.

**Fórmula de estimación de la proporción.** Para estimaciones de la proporción y puntuaciones normales se pueden seleccionar las fórmulas de estimación de la proporción: Blom, Tukey, Rankit o Van der Waerden.

- **Blom.** Crea nuevas variables de ordenación (rangos) que se basan en estimaciones de la proporción, las cuales utilizan la fórmula  $(r-3/8) / (w+1/4)$ , donde r es el rango y w es la suma de las ponderaciones de los casos.
- **Tukey.** Utiliza la fórmula  $(r-1/3) / (w+1/3)$ , donde r es el rango y w es la suma de las ponderaciones de los casos.
- **Rankit.** Utiliza la fórmula  $(r-1/2) / w$ , donde w es el número de observaciones y r es el rango, que va de 1 a w.
- **Van der Waerden.** La transformación de Van de Waerden, definida por la fórmula  $r/(w+1)$ , donde w es la suma de las ponderaciones de los casos y r es el rango, cuyo valor va de 1 a w.

Figura 8-13

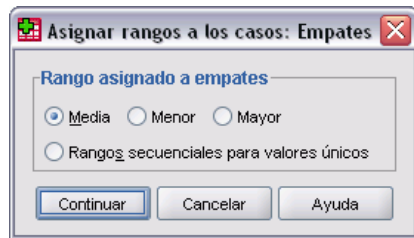
Cuadro de diálogo *Asignar rangos a los casos: Tipos*



### ***Asignar rangos a los casos: Empates***

Este cuadro de diálogo controla el método de asignación de rangos a los casos con el mismo valor en la variable original.

Figura 8-14  
Cuadro de diálogo Asignar rangos a los casos: Empates



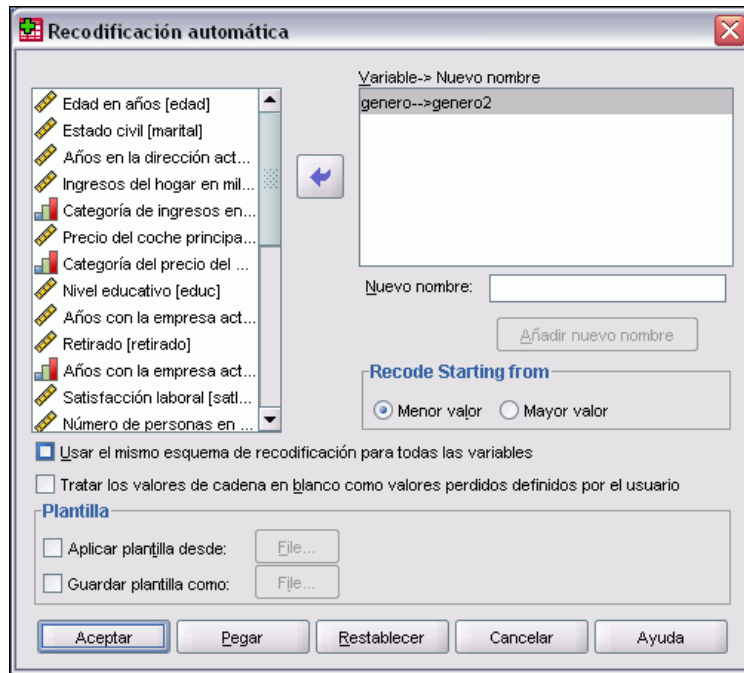
La tabla siguiente muestra cómo los distintos métodos asignan rangos a los valores empatados.

Valor	Media	Menor	Mayor	Secuencial
10	1	1	1	1
15	3	2	4	2
15	3	2	4	2
15	3	2	4	2
16	5	5	5	3
20	6	6	6	4

## ***Recodificación automática***

El cuadro de diálogo Recodificación automática le permite convertir los valores numéricos y de cadena en valores enteros consecutivos. Si los códigos de la categoría no son secuenciales, las casillas vacías resultantes reducen el rendimiento e incrementan los requisitos de memoria de muchos procedimientos. Además, algunos procedimientos no pueden utilizar variables de cadena y otros requieren valores enteros consecutivos para los niveles de los factores.

Figura 8-15  
Cuadro de diálogo Recodificación automática



- La nueva variable, o variables, creadas por la recodificación automática conservan todas las etiquetas de variable y de valor definidas de la variable antigua. Para los valores que no tienen una etiqueta de valor ya definida se utiliza el valor original como etiqueta del valor recodificado. Una tabla muestra los valores antiguos, los nuevos y las etiquetas de valor.
- Los valores de cadena se recodifican por orden alfabético, con las mayúsculas antes que las minúsculas.
- Los valores perdidos se recodifican como valores perdidos mayores que cualquier valor no perdido y conservando el orden. Por ejemplo, si la variable original posee 10 valores no perdidos, el valor perdido mínimo se recodificará como 11, y el valor 11 será un valor perdido para la nueva variable.

**Usar el mismo esquema de recodificación para todas las variables.** Esta opción le permite aplicar un único esquema de recodificación para todas las variables seleccionadas, lo que genera un esquema de codificación coherente para todas las variables nuevas.

Si selecciona esta opción, se aplican las siguientes reglas y limitaciones:

- Todas las variables deben ser del mismo tipo (numéricas o de cadena).
- Todos los valores observados para todas las variables seleccionadas se utilizan para crear un orden de valores para recodificar en enteros consecutivos.
- Los valores perdidos definidos por el usuario para las variables nuevas se basan en la primera variable de la lista con valores perdidos definidos por el usuario. El resto de los valores de las demás variables originales, excepto los valores perdidos del sistema, se consideran válidos.



**Trate los valores de cadena en blanco como valores perdidos definidos por el usuario.** En las variables de cadena, los valores en blanco o nulos no son tratados como valores **perdidos por el sistema**. Esta opción recodifica automáticamente las cadenas en blanco en **valores perdidos definidos por el usuario** valores más altos que el valor no perdido más elevado.

### **Plantillas**

Puede guardar el esquema de recodificación automática en un archivo de plantilla y, a continuación, aplicarlo a otras variables y otros archivos de datos.

Por ejemplo, puede tener un número considerable de códigos de producto alfanuméricos que se registran automáticamente en enteros cada mes, pero algunos meses se añaden códigos de productos nuevos al esquema de recodificación original. Si guarda el esquema original en una plantilla y, a continuación, la aplica a los datos nuevos que contienen el nuevo conjunto de códigos, todos los códigos nuevos encontrados en los datos se recodifican automáticamente en valores superiores al último valor de la plantilla para conservar el esquema de recodificación automática original de los códigos de productos originales.

**Guardar plantilla como.** Guarda el esquema de recodificación automática para las variables seleccionadas en un archivo de plantilla externo.

- La plantilla contiene información que asigna los valores no perdidos originales a los valores recodificados.
- En la plantilla sólo se guarda la información para los valores no perdidos. La información sobre los valores perdidos definidos por el usuario no se conserva.
- Si ha seleccionado varias variables para su recodificación, pero no ha optado por utilizar el mismo esquema de recodificación automática para todas las variables o no va a aplicar una plantilla existente como parte de la recodificación automática, la plantilla se basará en la primera variable de la lista.
- Si ha seleccionado varias variables para su recodificación, y también ha seleccionado Usar el mismo esquema de recodificación para todas las variables y/o Aplicar plantilla, la plantilla contendrá el esquema de recodificación automática combinado para todas las variables.

**Aplicar plantilla desde.** Aplica una plantilla de recodificación automática previamente guardada a las variables seleccionadas para la recodificación, añade los valores adicionales encontrados en las variables al final del esquema y conserva la relación entre los valores originales y recodificados automáticamente almacenados en el esquema guardado.

- Todas las variables seleccionadas para la recodificación deben ser del mismo tipo (numéricas o de cadena) y dicho tipo debe coincidir con el tipo definido en la plantilla.
- Las plantillas no pueden contener información sobre los valores perdidos definidos por el usuario. Los valores perdidos definidos por el usuario para las variables de destino se basan en la primera variable de la lista de variables originales con valores perdidos definidos por el usuario. El resto de los valores de las demás variables originales, excepto los valores perdidos del sistema, se consideran válidos.

- Las asignaciones de valores de la plantilla se aplican en primer lugar. Los valores restantes se recodifican en valores superiores al último valor de la plantilla, con los valores perdidos definidos por el usuario (basados en la primera variable de la lista con valores perdidos definidos por el usuario) recodificados en valores superiores al último valor válido.
- Si ha seleccionado diversas variables para su recodificación automática, la plantilla se aplica en primer lugar, seguida de una recodificación automática común combinada para todos los valores adicionales encontrados en las variables seleccionadas, lo que resulta en un único esquema de recodificación automática para todas las variables seleccionadas.

***Para recodificar valores numéricos o de cadena en valores enteros consecutivos***

- ▶ Elija en los menús:
  - Transformar
  - Recodificación automática...
- ▶ Seleccione la variable o variables que desee recodificar.
- ▶ Para cada variable seleccionada, introduzca un nombre para la nueva variable y pulse en Nuevo nombre.

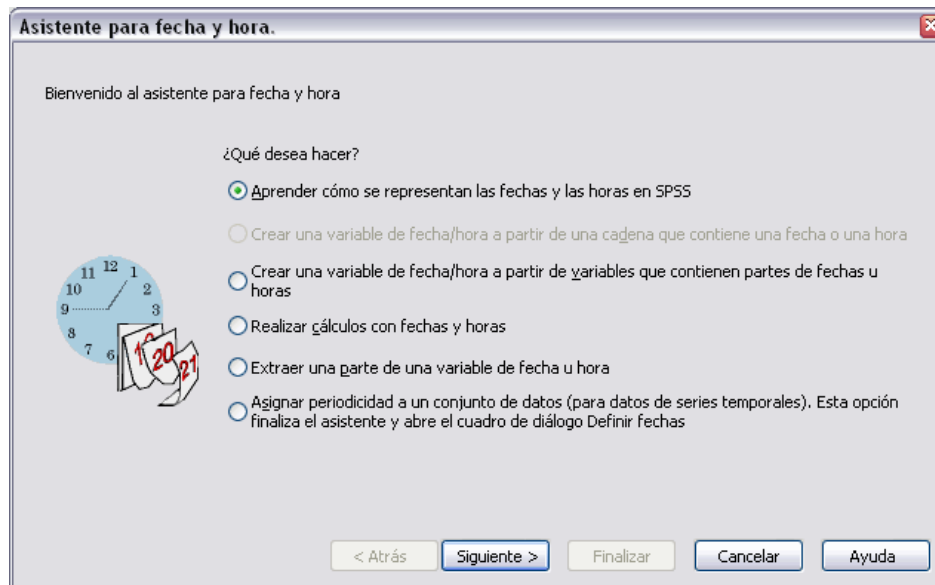
## ***Asistente para fecha y hora***

El Asistente para fecha y hora simplifica ciertas tareas comunes asociadas a las variables de fecha y hora.

***Para usar el Asistente para fecha y hora***

- ▶ Elija en los menús:
  - Transformar
  - Asistente para fecha y hora...
- ▶ Seleccione la tarea que desee realizar y siga los pasos para definir la tarea.

Figura 8-16  
Pantalla principal del Asistente para fecha y hora



- **Aprender cómo se representan las fechas y las horas.** Esta opción ofrece una pantalla en la que se presenta una breve descripción de las variables de fecha/hora en SPSS . El botón *Ayuda* también proporciona un vínculo para obtener información más detallada.
- **Crear una variable de fecha/hora a partir de una cadena que contiene una fecha o una hora.** Use esta opción para crear una variable de fecha/hora a partir de una variable de cadena. Por ejemplo, dispone de una variable de cadena que representa fechas con el formato mm/dd/aaaa y desea crear una variable de fecha/hora a partir de ella.
- **Crear una variable de fecha/hora fusionando variables que contengan partes diferentes de la fecha u hora.** Esta opción permite construir una variable de fecha/hora a partir de un conjunto de variables existentes. Por ejemplo, dispone de una variable que representa el mes (como un número entero), una segunda que representa el día del mes y una tercera que representa el año. Se pueden combinar estas variables en una única variable de fecha/hora.
- **Realizar cálculos con fechas y horas.** Use esta opción para añadir o sustraer valores a variables de fecha/hora. Por ejemplo, puede calcular la duración de un proceso sustrayendo una variable que represente la hora de comienzo del proceso de otra variable que represente la hora de finalización del proceso.
- **Extraer una parte de una variable de fecha/hora.** Esta opción permite extraer parte de una variable de fecha/hora, como el día del mes de una variable de fecha/hora, con el formato mm/dd/aaaa.
- **Asignar periodicidad a un conjunto de datos.** Esta opción presenta el cuadro de diálogo Definir fechas, que se usa para crear variables de fecha/hora compuestas por un conjunto de fechas secuenciales. Esta característica se usa generalmente para asociar fechas con datos de serie temporal.

*Nota:* las tareas se desactivan cuando el conjunto de datos carece de los tipos de variables necesarios para completar la tarea. Por ejemplo, si el conjunto de datos no contiene variables de cadena, la tarea de creación de una variable de fecha/hora a partir de una cadena no se aplica y se desactiva.

## **Fechas y horas en SPSS**

Las variables que representan fechas y horas en SPSS tienen un tipo de variable numérico, con formatos de presentación que se corresponden con los formatos específicos de fecha/hora. Estas variables se denominan generalmente variables de fecha/hora. Se distingue entre variables de fecha/hora que realmente representan fechas y aquellas que representan una duración temporal independiente de cualquier fecha, como 20 horas, 10 minutos y 15 segundos. Éstas últimas se denominan generalmente variables de duración, mientras que las primeras se conocen como variables de fecha o de fecha/hora. Para obtener una lista completa de los formatos de presentación, consulte el apartado sobre fecha y hora en la sección de elementos universales de la referencia de sintaxis de comandos (*Command Syntax Reference*).

**Variables de fecha y de fecha/hora** Las variables de fecha tienen un formato que representa una fecha, como mm/dd/aaaa. Las variables de fecha/hora tienen un formato que representa una fecha y una hora, como dd-mmm-aaaa hh:mm:ss. Internamente, las variables de fecha y de fecha/hora se almacenan como el número de segundos a partir del 14 de octubre de 1582. Las variables de fecha y de fecha/hora se denominan a menudo variables con formato de fecha.

- Las especificaciones de año reconocidas son tanto de dos como de cuatro dígitos. Por defecto, los años representados por dos dígitos representan un intervalo que comienza 69 años antes de la fecha actual y finaliza 30 años después de la fecha actual. Este intervalo está determinado por la configuración de las Opciones y se puede modificar (en el menú Edición, seleccione Opciones y pulse en la pestaña Datos).
- Los delimitadores que se pueden usar en los formatos de día-mes-año son guiones, puntos, comas, barras y espacios en blanco.
- Los meses se pueden representar en dígitos, números romanos o abreviaturas de tres caracteres, y se pueden escribir con el nombre completo. Los nombres de los meses expresados con abreviaturas de tres letras y nombres completos deben estar en inglés, ya que no se reconocen los nombres de meses en otros idiomas.

**Variables de duración.** Las variables de duración tienen un formato que representa una duración de tiempo, como hh:mm. Se almacenan internamente como segundos sin hacer referencia a ninguna fecha en particular.

- En las especificaciones de tiempo (se aplican a las variables de fecha/hora y de duración), los dos puntos se pueden usar como delimitadores entre horas, minutos y segundos. Las horas y los minutos son valores necesarios, pero los segundos son opcionales. Para separar los segundos de las fracciones de segundo, es necesario utilizar un punto. Las horas pueden tener una magnitud ilimitada, pero el valor máximo de los minutos es 59 y el de los segundos, 59.999...

**Fecha y hora actuales.** La variable del sistema  $\$TIME$  contiene la fecha y hora actuales. Representa el número de segundos transcurridos desde el 14 de octubre de 1582 hasta la fecha y la hora en que se ejecute el comando de transformación que la use.

### Creación de una variable de fecha/hora a partir de una cadena

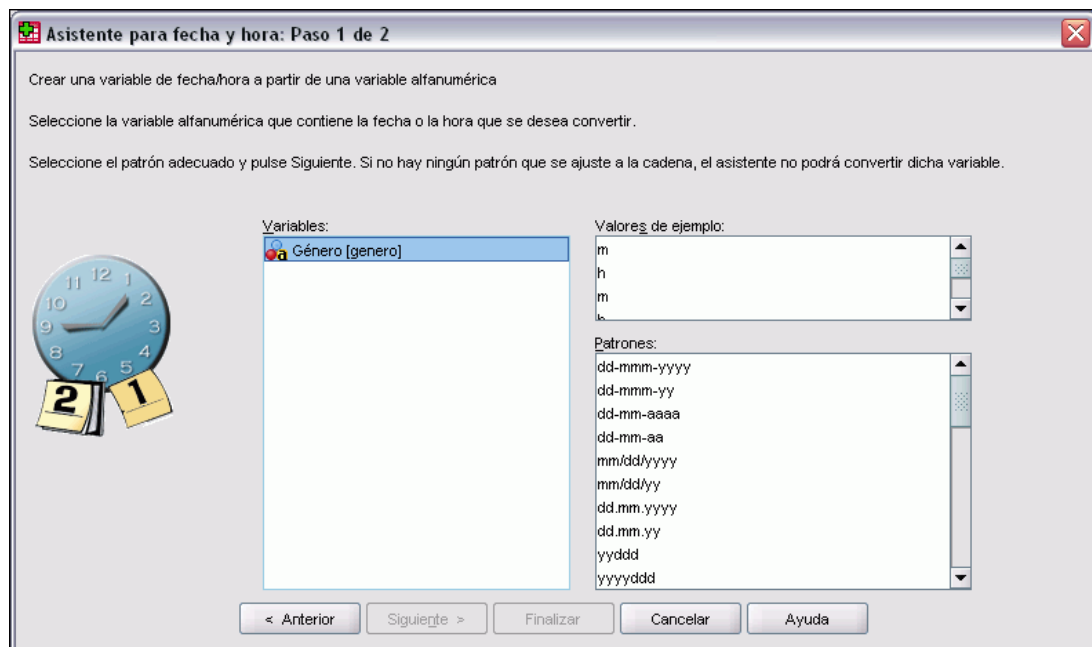
Para crear una variable de fecha/hora a partir de una variable de cadena:

- Seleccione Crear una variable de fecha/hora a partir de una cadena que contiene una fecha o una hora en la pantalla principal del Asistente para fecha y hora.

### Selección de una variable de cadena para convertir en una variable de fecha/hora

Figura 8-17

Creación de una variable de fecha/hora a partir de una variable de cadena, paso 1



- En la lista Variables, seleccione la variable de cadena que desee convertir. Observe que la lista sólo contiene variables de cadena.
- En la lista Patrones, seleccione el patrón que coincida con el modo en que la variable de cadena representa las fechas. La lista Valores de ejemplo muestra los valores reales de la variable seleccionada en el archivo de datos. Los valores de la variable de cadena que no se ajusten al patrón seleccionado darán como resultado un valor perdido del sistema para la nueva variable.

### **Especificación del resultado de convertir la variable de cadena en variable de fecha/hora**

Figura 8-18

Creación de una variable de fecha/hora a partir de una variable de cadena, paso 2

- ▶ Escriba un nombre para la variable de resultado. Éste no puede coincidir con el de una variable existente.

Si lo desea, puede:

- Seleccionar un formato de fecha/hora para la nueva variable en la lista Formato de resultado.
- Asignar una etiqueta de variable descriptiva a la nueva variable.

### **Creación de una variable de fecha/hora a partir de un conjunto de variables**

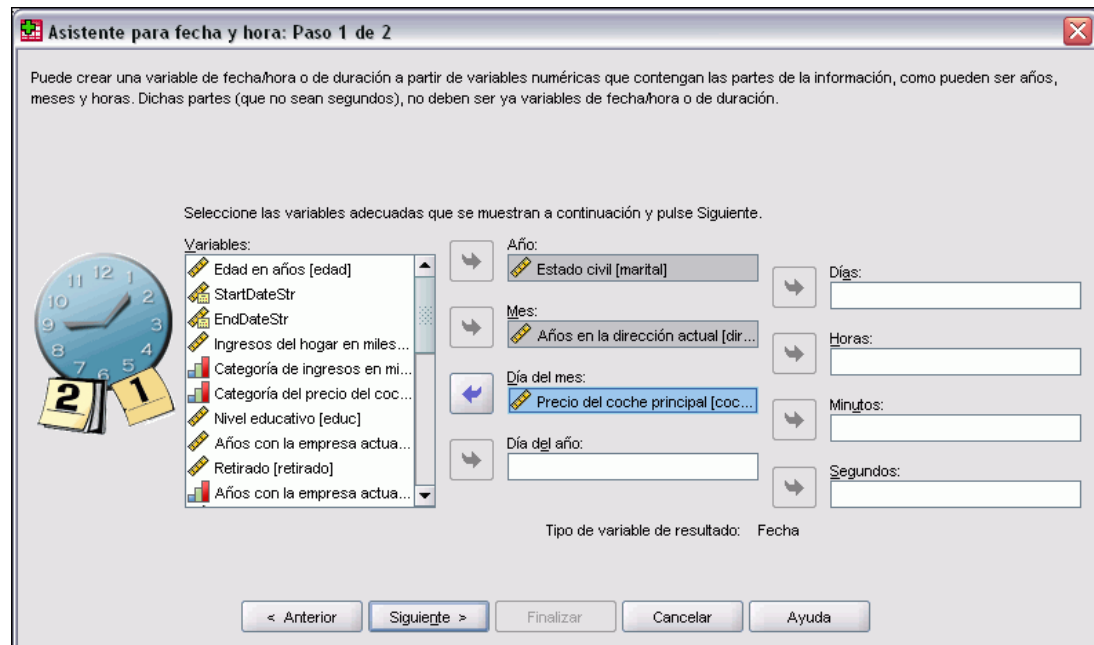
Para fusionar un conjunto de variables existentes en una única variable de fecha/hora:

- ▶ Seleccione Crear una variable de fecha/hora fusionando variables que contengan partes diferentes de la fecha u hora en la pantalla principal del Asistente para fecha y hora.

### Selección de variables que fusionar en una única variable de fecha/hora

Figura 8-19

Creación de una variable de fecha/hora a partir de un conjunto de variables, paso 1

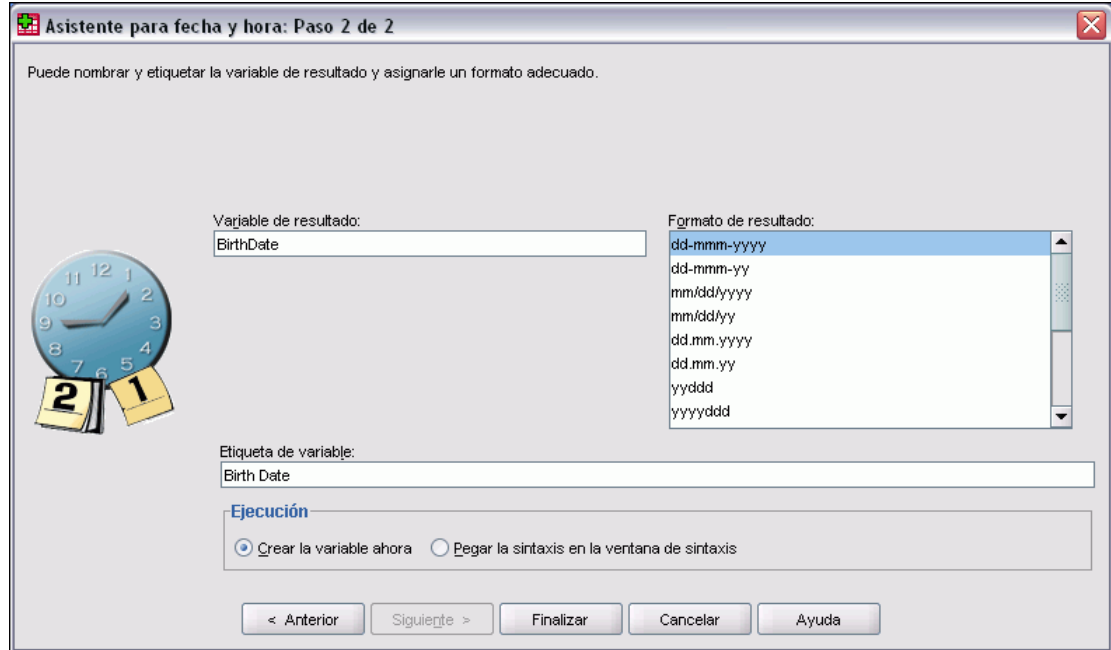


- ▶ Seleccione las variables que representen las diferentes partes de la fecha/hora.
  - Algunas combinaciones de selecciones no están permitidas. Por ejemplo, la creación de una variable de fecha/hora a partir de un valor de Año y Día del mes no es válida porque, una vez seleccionado Año, es necesario especificar una fecha completa.
  - No se puede utilizar una variable de fecha/hora existente como una de las partes de la variable de fecha/hora final que se está creando. Las variables que componen las partes de la nueva variable de fecha/hora deben ser números enteros. La excepción es el uso permitido de una variable de fecha/hora existente como la parte de los segundos de la nueva variable. Puesto que se permite el uso de fracciones de segundos, las variables utilizadas para los segundos no tiene que ser obligatoriamente un número entero.
  - Los valores de cualquier parte de la nueva variable que no se ajusten al rango permitido darán como resultado un valor perdido del sistema para la nueva variable. Por ejemplo, si se usa inadvertidamente una variable que representa un día del mes como valor de Mes, todos los casos en que el valor del día del mes pertenezca al rango 14–31 se considerarán valores perdidos del sistema para la nueva variable, puesto que el rango válido para los meses en SPSS es 1–13.

### **Especificación de variable de fecha/hora creada fusionando variables**

Figura 8-20

Creación de una variable de fecha/hora a partir de un conjunto de variables, paso 2



- ▶ Escriba un nombre para la variable de resultado. Éste no puede coincidir con el de una variable existente.
- ▶ Seleccione un formato de fecha/hora de la lista Formato de resultado.
  - Si lo desea, puede:
    - Asignar una etiqueta de variable descriptiva a la nueva variable.

### **Adición o sustracción de valores a partir de variables de fecha/hora**

Para añadir o sustraer valores a variables de fecha/hora:

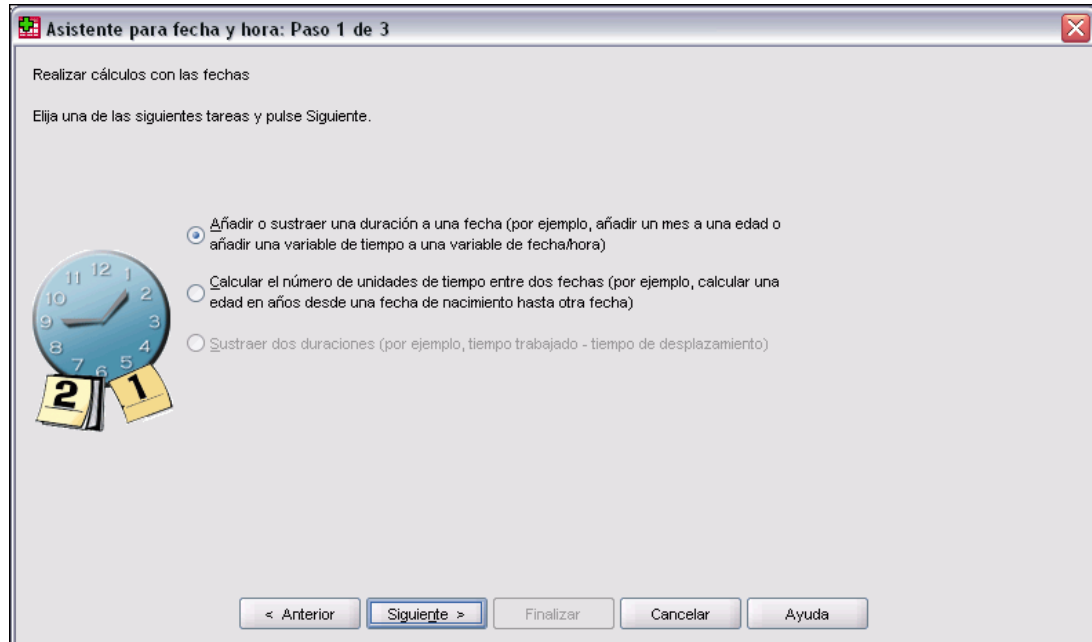
- ▶ Seleccione Calcular con fechas y horas en la pantalla principal del Asistente para fecha y hora.



### Selección del tipo de cálculo que realizar con las variables de fecha/hora

Figura 8-21

Adición o sustracción de valores de variables de fecha/hora, paso 1



- **Añadir o sustraer una duración a una fecha.** Use esta opción para añadir o sustraer valores a una variable con formato de fecha. Si lo desea, puede añadir o sustraer duraciones que sean valores fijos, como 10 días, o los valores de una variable numérica (por ejemplo, una variable que represente años).
- **Calcular el número de unidades de tiempo entre dos fechas.** Use esta opción para obtener la diferencia entre dos fechas medidas en una unidad seleccionada. Por ejemplo, puede obtener el número de años o el número de días que separan dos fechas.
- **Sustraer dos duraciones.** Use esta opción para obtener la diferencia entre dos variables con formatos de duración, como hh:mm o hh:mm:ss.

*Nota:* las tareas se desactivan cuando el conjunto de datos carece de los tipos de variables necesarios para completar la tarea. Por ejemplo, si el conjunto de datos no contiene dos variables con formatos de duración, la tarea de sustracción de dos duraciones no se aplica y se desactiva.

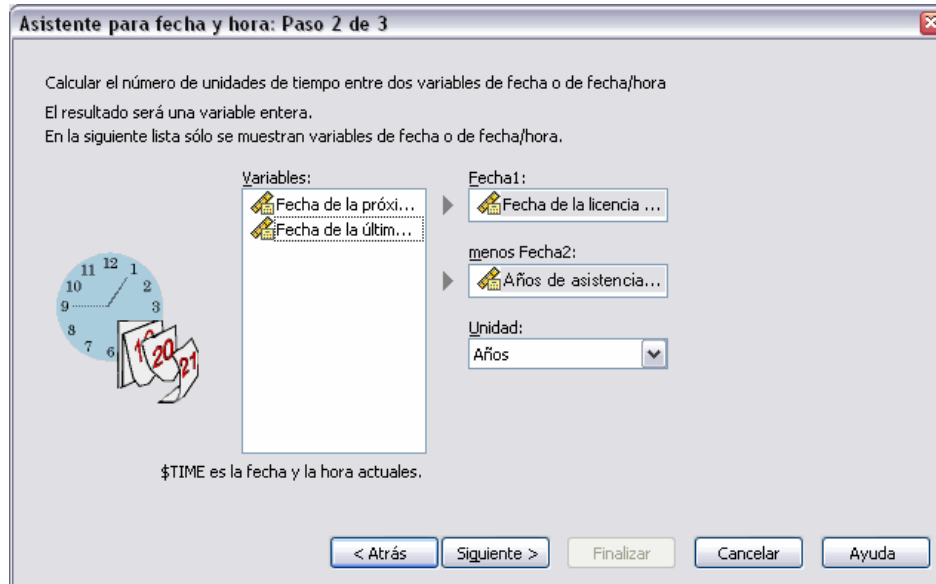
### Adición o sustracción de una duración a una fecha

Para añadir o sustraer una duración a una variable con formato de fecha:

- Seleccione Añadir o sustraer una duración a una fecha en la pantalla del Asistente para fecha y hora denominada *Realizar cálculos con las fechas*.

### Selección de variable de fecha/hora y duración que añadir o sustraer

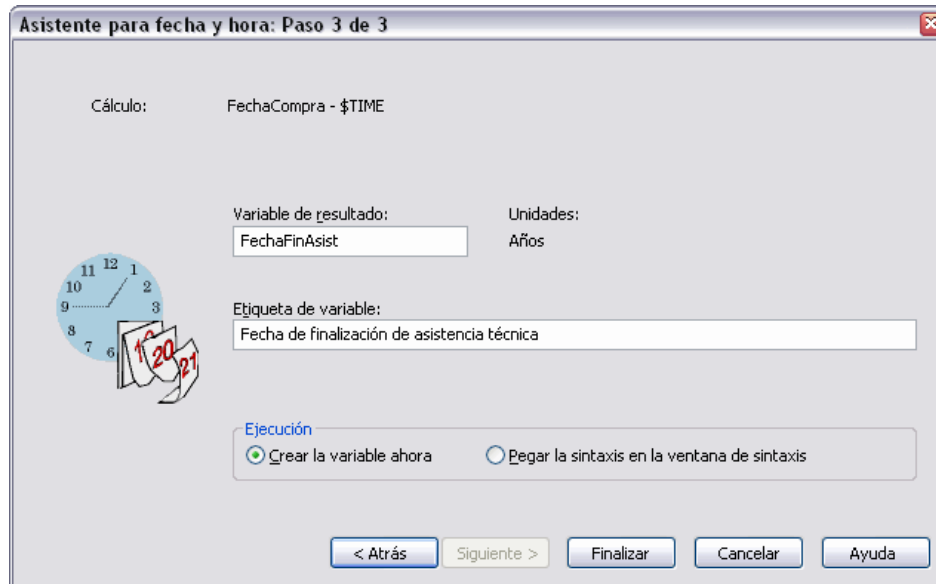
Figura 8-22  
Adición o sustracción de duración, paso 2



- ▶ Seleccione una variable de fecha (u hora).
- ▶ Seleccione una variable de duración o especifique un valor para Constante de duración. Las variables utilizadas para las duraciones no pueden ser variables de fecha o de fecha/hora. Pueden ser variables de duración o variables numéricas simples.
- ▶ Seleccione la unidad que represente la duración en la lista desplegable. Seleccione Duración si se usa una variable y ésta tiene el formato de una duración, como hh:mm o hh:mm:ss.

### **Especificación de los resultados de la adición o sustracción de una duración a una variable de fecha/hora**

Figura 8-23  
Adición o sustracción de duración, paso 3



- ▶ Escriba un nombre para la variable de resultado. Éste no puede coincidir con el de una variable existente.
- Si lo desea, puede:
- Asignar una etiqueta de variable descriptiva a la nueva variable.

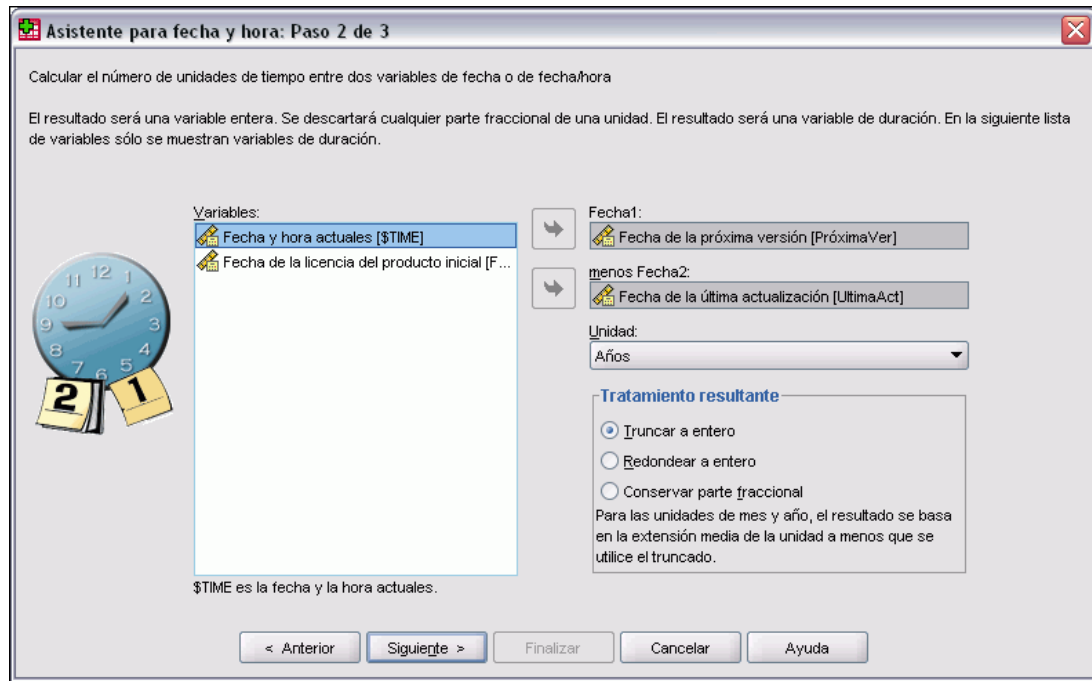
### **Sustracción de variables con formato de fecha**

Para sustraer dos variables con formato de fecha:

- ▶ Seleccione Calcular el número de unidades de tiempo entre dos fechas en la pantalla del Asistente para fecha y hora denominada *Realizar cálculos con las fechas*.

### Selección de variables con formato de fecha que sustraer

Figura 8-24  
Sustracción de fechas, paso 2



- ▶ Seleccione las variables que se van a sustraer.
- ▶ Seleccione la unidad del resultado en la lista desplegable.
- ▶ Seleccione cómo debería calcularse el resultado (tratamiento de resultado).

#### Tratamiento de resultado

Las siguientes opciones están disponibles para el cálculo del resultado:

- **Truncar a número entero.** Se ignora cualquier parte fraccional del resultado. Por ejemplo, si se resta 28/10/2006 de 21/10/2007 se obtiene un resultado de 0 para los años y 11 para los meses.
- **Redondear a número entero.** El resultado se redondea al número entero más cercano. Por ejemplo, si se resta 28/10/2006 de 21/10/2007 se obtiene un resultado de 1 para los años y 12 para los meses.
- **Conservar parte fraccional.** Se conserva el valor completo, sin truncar ni redondear el resultado. Por ejemplo, si se resta 28/10/2006 de 21/10/2007 se obtiene un resultado de 0,98 para los años y 11,76 para los meses.

En el caso del redondeo y la conservación fraccional, el resultado para los años se basa en el número medio de días de un año (365,25), y el resultado para los meses se basa en el número medio de días de un mes (30,4375). Por ejemplo, si se resta 1/2/2007 de 1/3/2007 (formato d/m/a) devuelve un resultado fraccional de 0,92 meses; en cambio, si se resta 1/3/2007 de 1/2/2007 devuelve una diferencia fraccional de 1,02 meses. Esto también afecta a los valores calculados en

periodos de tiempo que incluyen años bisiestos. Por ejemplo, si se resta 1/2/2008 de 1/3/2008 devuelve una diferencia fraccional de 0,95 meses, en comparación con los 0,92 para el mismo periodo de tiempo en un año no bisiesto.

Fecha 1	Fecha 2	Años			Meses		
		Truncado	Redondeado	Fracción	Truncado	Redondeado	Fracción
10/21/2006	10/28/2007	1	1	1.02	12	12	12.22
10/28/2006	10/21/2007	0	1	.98	11	12	11.76
2/1/2007	3/1/2007	0	0	.08	1	1	.92
2/1/2008	3/1/2008	0	0	.08	1	1	.95
3/1/2007	4/1/2007	0	0	.08	1	1	1.02
4/1/2007	5/1/2007	0	0	.08	1	1	.99

### Especificación del resultado de la sustracción de dos variables con formato de fecha

Figura 8-25  
Sustracción de fechas, paso 3

- Escriba un nombre para la variable de resultado. Éste no puede coincidir con el de una variable existente.

Si lo desea, puede:

- Asignar una etiqueta de variable descriptiva a la nueva variable.

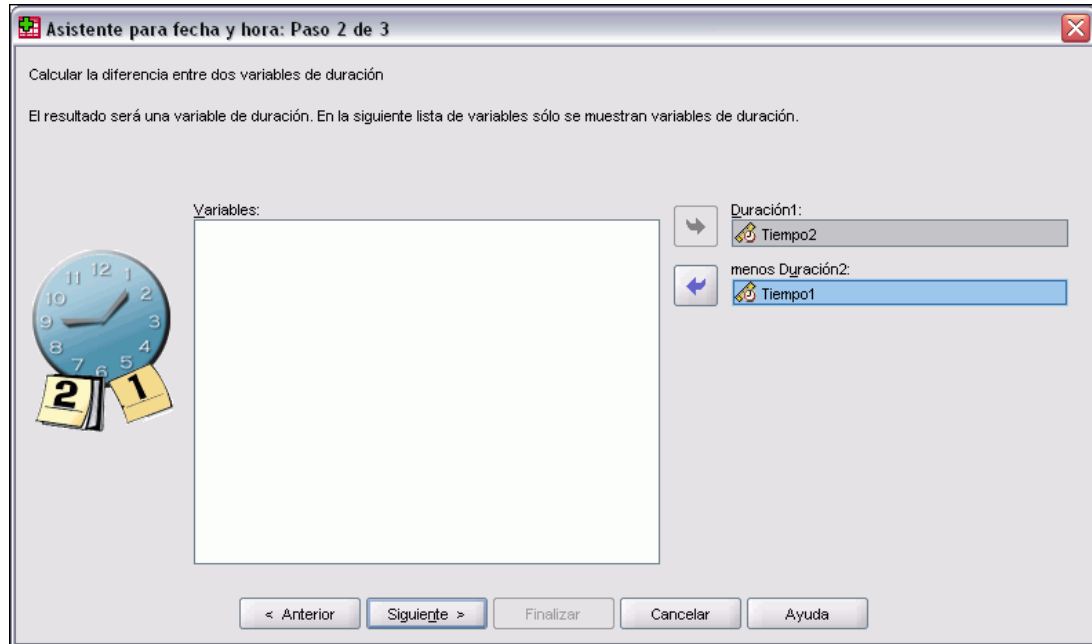
### Sustracción de variables de duración

Para sustraer dos variables de duración:

- Seleccione Sustraer dos duraciones en la pantalla del Asistente para fecha y hora denominada *Realizar cálculos con las fechas*.

**Selección de las variables de duración que sustraer**

Figura 8-26  
Sustracción de dos duraciones, paso 2



- Seleccione las variables que se van a sustraer.

### Especificación del resultado de la sustracción de dos variables de duración

Figura 8-27  
Sustracción de dos duraciones, paso 3

- ▶ Escriba un nombre para la variable de resultado. Éste no puede coincidir con el de una variable existente.
- ▶ Seleccione un formato de duración de la lista Formato de resultado.
  - Si lo desea, puede:
    - Asignar una etiqueta de variable descriptiva a la nueva variable.

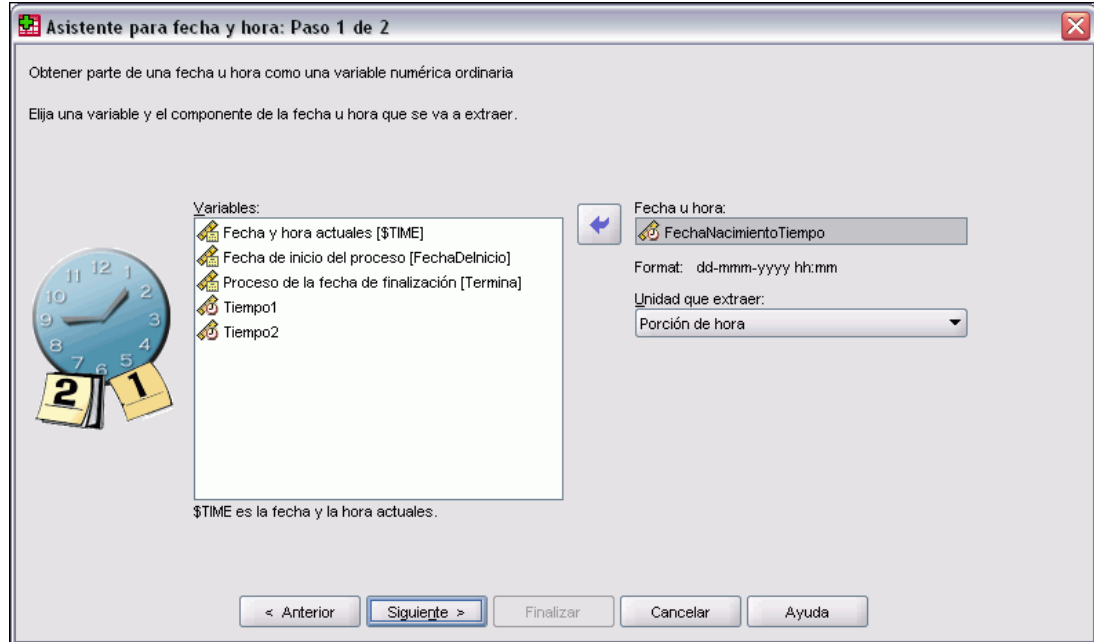
### Extracción de parte de una variable de fecha/hora

Para extraer un componente, como puede ser el año, de una variable de fecha/hora:

- ▶ Seleccione Extraer una parte de una variable de fecha u hora en la pantalla principal del Asistente para fecha y hora.

**Selección de componente que extraer de una variable de fecha/hora**

Figura 8-28  
Extracción de parte de una variable de fecha/hora, paso 1

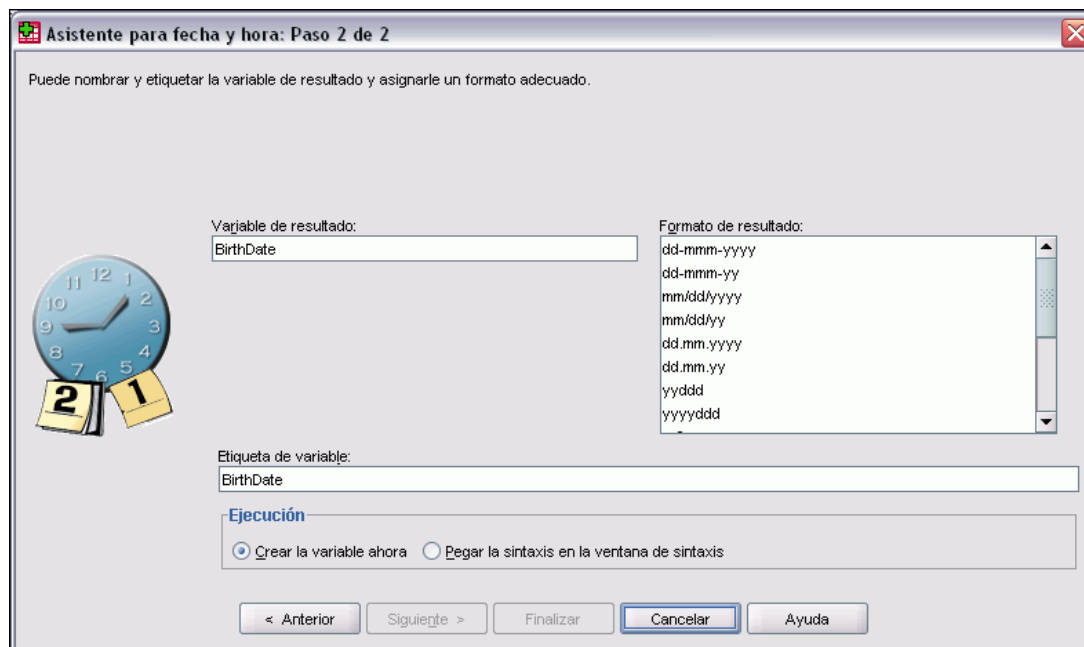


- ▶ Seleccione la variable que contiene la parte de fecha u hora que desee extraer.
- ▶ En la lista desplegable, seleccione la parte de la variable que se va a extraer. Si lo desea, puede extraer información de fechas que no sea explícitamente parte de la fecha que se muestra, por ejemplo, un día de la semana.



### Especificación del resultado de la extracción de un componente de una variable de fecha/hora

Figura 8-29  
Extracción de parte de una variable de fecha/hora, paso 2



- ▶ Escriba un nombre para la variable de resultado. Éste no puede coincidir con el de una variable existente.
- ▶ Si está extrayendo la parte de fecha o de hora de una variable de fecha/hora, debe seleccionar un formato de la lista Formato de resultado. En los casos en que el formato de resultado no es necesario se desactivará la lista Formato de resultado.

Si lo desea, puede:

- Asignar una etiqueta de variable descriptiva a la nueva variable.

## Transformaciones de los datos de serie temporal

Se incluyen diversas transformaciones de datos de gran utilidad en los análisis de series temporales:

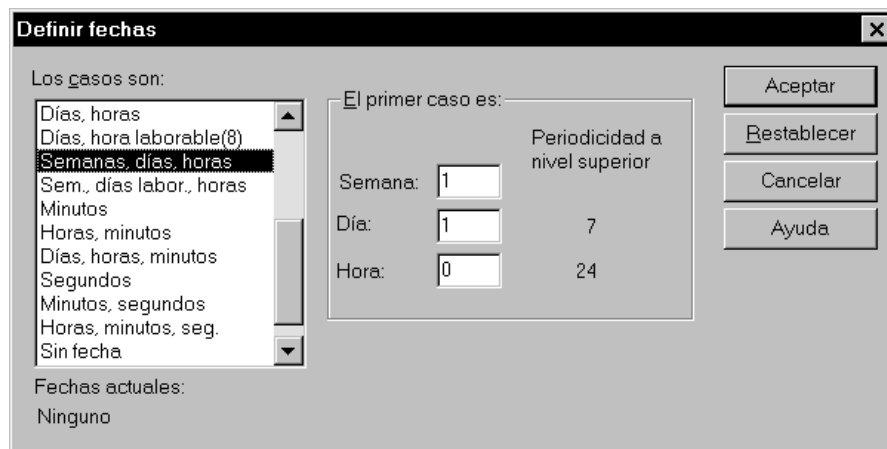
- Generar variables de fecha para establecer la periodicidad y distinguir entre los períodos históricos, de validación y de predicción.
- Elaborar nuevas variables de series temporales como funciones de variables de series temporales existentes.
- Reemplazar valores definidos como perdidos por el usuario y perdidos del sistema con estimaciones basadas en uno de los diversos métodos existentes.

Una **serie temporal** se obtiene midiendo una variable (o un conjunto de variables) de manera regular a lo largo de un período de tiempo. Las transformaciones de los datos de serie temporal suponen una estructura de archivo de datos en la que cada caso (fila) representa un conjunto de observaciones para un momento diferente y la duración del tiempo entre los casos es uniforme.

## Definir fechas

El cuadro de diálogo Definir fechas permite generar variables de fecha que se pueden utilizar para establecer la periodicidad de una **serie temporal** y para etiquetar los resultados de los análisis de series temporales.

Figura 8-30  
Cuadro de diálogo Definir fechas



**Los casos son.** Define el intervalo de tiempo utilizado para generar las fechas.

- Sin fecha elimina las variables de fecha definidas anteriormente. Se eliminarán todas las variables con los nombres siguientes: *año\_*, *trimestre\_*, *mes\_*, *semana\_*, *día\_*, *hora\_*, *minuto\_*, *segundo\_* y *fecha\_*.
- Personalizado indica la presencia de variables de fecha personalizadas, creadas con la sintaxis de comandos (por ejemplo, una semana de cuatro días laborables). Este elemento simplemente refleja el estado actual del conjunto de datos activo. Su selección en la lista no produce ningún efecto.

**El primer caso es.** Define el valor de la fecha inicial, que se asigna al primer caso. A los casos subsiguientes se les asignan valores secuenciales, basándose en el intervalo de tiempo.

**Periodicidad a nivel superior.** Indica la variación cíclica repetitiva, como el número de meses de un año o el número de días de la semana. El valor mostrado indica el valor máximo que se puede introducir.

Para cada componente utilizado para definir la fecha, se crea una nueva variable numérica. Los nombres de las nuevas variables terminan con un carácter de subrayado. A partir de los componentes también se crea una variable de cadena descriptiva, *fecha\_*. Por ejemplo, si selecciona Semanas, días, horas, se crearán cuatro nuevas variables: *semana\_*, *día\_*, *hora\_* y *fecha\_*.

Si ya se han definido variables de fecha, éstas serán reemplazadas cuando se definan nuevas variables de fecha con los mismos nombres que las existentes.

### **Para definir fechas para los datos de serie temporal**

- ▶ Elija en los menús:
  - Datos
  - Definir fechas...
- ▶ Seleccione un intervalo de tiempo en la lista Los casos son.
- ▶ Introduzca el valor o los valores que definen la fecha inicial en El primer caso es, que determina la fecha asignada al primer caso.

### **Variables de fecha frente a variables con formato de fecha**

Las variables de fecha creadas con Definir fechas no deben confundirse con las variables con formato de fecha, que se definen en Vista de variables del Editor de datos. Las variables de fecha se emplean para establecer la periodicidad de los datos de serie temporal; mientras que las variables con formato de fecha representan fechas y horas mostradas en varios formatos de fecha y hora. Las variables de fecha son números enteros sencillos que representan el número de días, semanas, horas, etc., a partir de un punto inicial especificado por el usuario. Internamente, la mayoría de las variables con formato de fecha se almacenan como el número de segundos transcurridos desde el 14 de octubre de 1582.

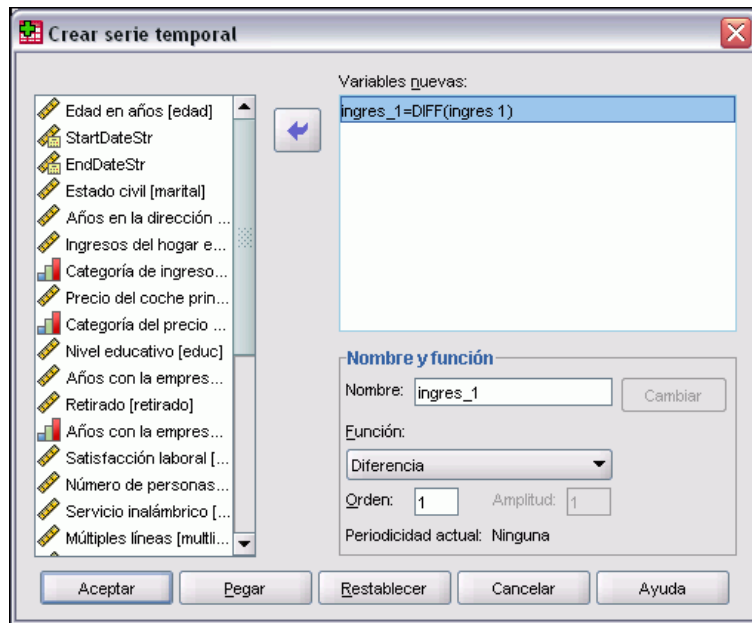
### **Crear serie temporal**

El cuadro de diálogo Crear serie temporal crea nuevas variables basadas en funciones de variables de **series temporales** numéricas existentes. Estos valores transformados son de gran utilidad en muchos procedimientos de análisis de series temporales.

Los nombres por defecto de las nuevas variables se componen de los seis primeros caracteres de las variables existentes utilizadas para crearlas, seguidos por un carácter de subrayado y un número secuencial. Por ejemplo, para la variable *precio*, el nombre de la nueva variable sería *precio\_1*. Las nuevas variables conservarían cualquier etiqueta de valor definida de las variables originales.

Las funciones disponibles para crear variables de series temporales incluyen las funciones de diferencias, medias móviles, medianas móviles, retardo y adelanto.

Figura 8-31  
Cuadro de diálogo Crear serie temporal



### **Para crear una nueva variable de serie temporal**

- ▶ Elija en los menús:  
Transformar  
Crear serie temporal...
- ▶ Seleccione la función de serie temporal que desea utilizar para transformar la variable o variables originales.
- ▶ Seleccione la variable o variables a partir de las cuales desee crear nuevas variables de serie temporal. Sólo se pueden utilizar variables numéricas.

Si lo desea, puede:

- Introducir nombres de variables, para omitir los nombres por defecto de las nuevas variables.
- Cambiar la función para una variable seleccionada.

### **Funciones de transformación de series temporales**

**Diferencia.** Diferencia no estacional entre valores sucesivos de la serie. El orden es el número de valores previos utilizados para calcular la diferencia. Dado que se pierde una observación para cada orden de diferencia, aparecerán valores perdidos del sistema al comienzo de la serie. Por ejemplo, si el orden de diferencia es 2, los primeros dos casos tendrán el valor perdido del sistema para la nueva variable.

**Diferencia estacional.** Diferencia los valores de la serie respecto a los valores de la propia serie distanciados un orden (un lapso) de valores constante. El orden se basa en la periodicidad definida actualmente. Para calcular diferencias estacionales debe haber definido variables de fecha (menú

Datos, Definir fechas) que incluyan un componente estacional (como por ejemplo los meses del año). El orden es el número de períodos estacionales utilizados para calcular la diferencia. El número de casos con el valor perdido del sistema al comienzo de la serie es igual a la periodicidad multiplicada por el orden de la diferencia estacional. Por ejemplo, si la periodicidad actual es 12 y el orden es 2, los primeros 24 casos tendrán el valor perdido del sistema para la nueva variable.

**Media móvil centrada.** Se utiliza el promedio de un rango de los valores de la serie, que rodean e incluyen al valor actual. La amplitud es el número de valores de la serie utilizados para calcular el promedio. Si la amplitud es par, la media móvil se calcula con el promedio de cada par de medias no centradas. Número de casos con el valor perdido del sistema al comienzo y al final de la serie para una amplitud de  $n$  es igual a  $n/2$  para los valores de la amplitud par y  $(n-1)/2$  para los valores de la amplitud impar. Por ejemplo, si la amplitud es 5, el número de casos con el valor perdido del sistema al comienzo y al final de la serie es 2.

**Media móvil anterior.** Se utiliza el promedio de un rango de las observaciones precedentes. La amplitud es el número de valores precedentes de la serie utilizados para calcular el promedio. El número de casos con el valor perdido del sistema al comienzo de la serie es igual al valor de la amplitud.

**Medianas móviles.** Se utiliza la mediana de un rango de los valores de la serie, que rodean e incluyen al valor actual. La amplitud es el número de valores de la serie utilizados para calcular la mediana. Si la amplitud es par, la mediana se calcula con el promedio de cada par de medianas no centradas. Número de casos con el valor perdido del sistema al comienzo y al final de la serie para una amplitud de  $n$  es igual a  $n/2$  para los valores de la amplitud par y  $(n-1)/2$  para los valores de la amplitud impar. Por ejemplo, si la amplitud es 5, el número de casos con el valor perdido del sistema al comienzo y al final de la serie es 2.

**Suma acumulada.** Cada valor de la serie se sustituye por la suma acumulada de los valores precedentes, incluyendo el valor actual.

**Retardo.** Cada valor de la serie se sustituye por el valor del caso precedente, en el orden especificado. El orden especifica a qué distancia se encuentra el caso precedente. El número de casos con el valor perdido del sistema al comienzo de la serie es igual al valor del orden.

**Adelanto.** Cada valor de la serie se sustituye por el valor de un caso posterior, en el orden especificado. El orden especifica a qué distancia se encuentra el caso posterior. El número de casos con el valor perdido del sistema al final de la serie es igual al valor del orden.

**Suavizado.** Los nuevos valores de la serie se basan en un suavizador de datos compuesto. El suavizador comienza con una mediana móvil de 4, que se centra por una mediana móvil de 2. A continuación, se vuelven a suavizar estos valores aplicando una mediana móvil de 5, una mediana móvil de 3 y los promedios ponderados móviles (hanning). Los residuos se calculan sustrayendo la serie suavizada de la serie original. Después se repite todo el proceso sobre los residuos calculados. Por último, los residuos suavizados se calculan sustrayendo los valores suavizados obtenidos la primera vez que se realizó el proceso. A esto se le denomina a veces **suavizado T4253H**.

## Reemplazar los valores perdidos

Las observaciones perdidas pueden causar problemas en los análisis y algunas medidas de series temporales no se pueden calcular si hay valores perdidos en la serie. En ocasiones el valor para una observación concreta no se conoce. Además, los datos perdidos pueden ser el resultado de lo siguiente:

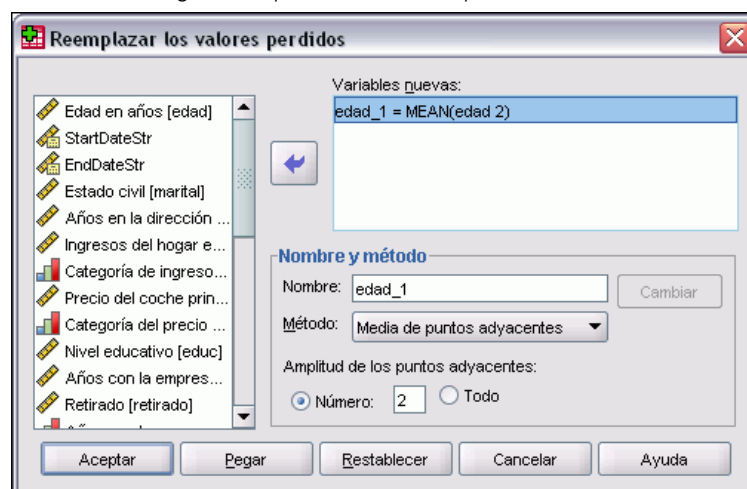
- Cada grado de diferenciación reduce la longitud de una serie en 1.
- Cada grado de diferenciación estacional reduce la longitud de una serie en una estación.
- Si genera una serie nueva que contenga predicciones que sobrepasen el final de la serie existente (al pulsar en el botón Guardar y realizar las selecciones adecuadas), la serie original y la serie residual generada incluirán datos perdidos para las observaciones nuevas.
- Algunas transformaciones (por ejemplo, la transformación logarítmica) generan datos perdidos para determinados valores de la serie original.

Los valores perdidos al principio o fin de una serie no suponen un problema especial; sencillamente acortan la longitud útil de la serie. Las discontinuidades que aparecen en mitad de una serie (datos *incrustados* perdidos) pueden ser un problema mucho más grave. El alcance del problema depende del procedimiento analítico que se utilice.

El cuadro de diálogo Reemplazar valores perdidos permite crear nuevas **variables de series temporales** a partir de otras existentes, reemplazando los valores perdidos por estimaciones calculadas mediante uno de los distintos métodos posibles. Los nombres por defecto de las nuevas variables se componen de los seis primeros caracteres de las variables existentes utilizadas para crearlas, seguidos por un carácter de subrayado y un número secuencial. Por ejemplo, para la variable *precio*, el nombre de la nueva variable sería *precio\_1*. Las nuevas variables conservarían cualquier etiqueta de valor definida de las variables originales.

Figura 8-32

Cuadro de diálogo Reemplazar los valores perdidos



**Para reemplazar los valores perdidos para las variables de series temporales**

- ▶ Elija en los menús:
  - Transformar
  - Reemplazar valores perdidos...
- ▶ Seleccione el método de estimación que desee utilizar para reemplazar los valores perdidos.
- ▶ Seleccione la variable o variables para las que desea reemplazar los valores perdidos.
  - Si lo desea, puede:
    - Introducir nombres de variables, para omitir los nombres por defecto de las nuevas variables.
    - Cambiar el método de estimación para una variable seleccionada.

**Métodos de estimación para reemplazar los valores perdidos**

**Media de la serie.** Sustituye los valores perdidos con la media de la serie completa.

**Media de puntos adyacentes.** Sustituye los valores perdidos por la media de los valores válidos circundantes. La amplitud de los puntos adyacentes es el número de valores válidos, por encima y por debajo del valor perdido, utilizados para calcular la media.

**Mediana de puntos adyacentes.** Sustituye los valores perdidos por la mediana de los valores válidos circundantes. La amplitud de los puntos adyacentes es el número de valores válidos, por encima y por debajo del valor perdido, utilizados para calcular la mediana.

**Interpolación lineal.** Sustituye los valores perdidos utilizando una interpolación lineal. Se utilizan para la interpolación el último valor válido antes del valor perdido y el primer valor válido después del valor perdido. Si el primer o el último caso de la serie tiene un valor perdido, el valor perdido no se sustituye.

**Tendencia lineal en el punto.** Reemplaza los valores perdidos de la serie por la tendencia lineal en ese punto. Se hace una regresión de la serie existente sobre una variable índice escalada de 1 a  $n$ . Los valores perdidos se sustituyen por sus valores pronosticados.

**Puntuación de datos con modelos predictivos**

El proceso de aplicar un modelo predictivo a un conjunto de datos se denomina **puntuación** de los datos. SPSS, Clementine y AnswerTree disponen de procedimientos para la creación de modelos predictivos como los modelos de regresión, conglomeración, árbol y red neuronal. Una vez creado un modelo, sus especificaciones se pueden guardar como un archivo XML que contenga toda la información necesaria para reconstruirlo. La versión de servidor de SPSS permite leer un archivo de modelo XML y aplicar el modelo a un conjunto de datos.

**Ejemplo.** El riesgo de una solicitud de crédito se evalúa basándose en diversos aspectos del solicitante y el préstamo en cuestión. La puntuación del crédito obtenida a partir del modelo de riesgo se usa para aceptar o rechazar la solicitud de préstamo.

La puntuación se trata como una transformación de los datos. El modelo se expresa internamente como un conjunto de transformaciones numéricas que se deben aplicar a un determinado conjunto de variables (las variables predictoras especificadas en el modelo), con el fin de obtener un

resultado predictivo. En este sentido, el proceso de puntuación de los datos con un modelo dado es, inherentemente, igual que la aplicación de cualquier función, como puede ser una función de raíz cuadrada, a un conjunto de datos.

La puntuación sólo está disponible con el servidor de SPSS y puede realizarse interactivamente por los usuarios que trabajan en análisis en modo distribuido. Para puntuar archivos de datos de gran tamaño, es posible que desee usar SPSS Batch Facility, un ejecutable independiente que se distribuye junto con la versión de servidor de SPSS. Para obtener información acerca del uso de SPSS Batch Facility, consulte el manual del usuario de SPSS Batch Facility (*SPSS Batch Facility User's Guide*), que se incluye como archivo PDF en el CD del producto de la versión de servidor de SPSS.

El proceso de puntuación consta de:

- ▶ [Cargar](#) un modelo de un archivo en formato XML (PMML).
- ▶ Calcular las puntuaciones como una variable nueva, mediante las funciones `ApplyModel` o `StrApplyModel` disponibles en el cuadro de diálogo [Calcular variable](#).

Para obtener más información sobre las funciones `ApplyModel` o `StrApplyModel`, consulte las expresiones de puntuación de la sección sobre expresiones de transformación de la referencia de sintaxis de comandos (*Command Syntax Reference*).

En la siguiente tabla se muestran los procedimientos que admiten la exportación de especificaciones de modelo a XML. Los modelos exportados se pueden utilizar con SPSS Server para puntuar nuevos datos como se describe a continuación. La lista completa de tipos de modelos que se pueden puntuar con SPSS Server se encuentra en la descripción de la función `ApplyModel`.

Nombre de procedimiento	Nombre de comando	Opción
Discriminante	DISCRIMINANT	Base
Regresión lineal	REGRESSION	Base
Conglomerado en dos fases	TWOSTEP CLUSTER	Base
Modelos lineales generalizados	GENLIN	Modelos avanzados
Modelo lineal general de muestras complejas	CSGLM	Muestras complejas
Regresión logística de muestras complejas	CSLOGISTIC	Muestras complejas
Regresión ordinal de muestras complejas	CSORDINAL	Muestras complejas
Regresión logística	LOGISTIC REGRESSION	Regresión
Regresión logística multinomial	NOMREG	Regresión
Árbol de clasificación	TREE	Árbol

### ***Carga de un modelo guardado***

El cuadro de diálogo Cargar modelo permite cargar modelos predictivos guardados en formato XML (PMML) y sólo está disponible cuando se trabaja en análisis en modo distribuido. La carga de modelos es un primer paso necesario para puntuar los datos con ellos.



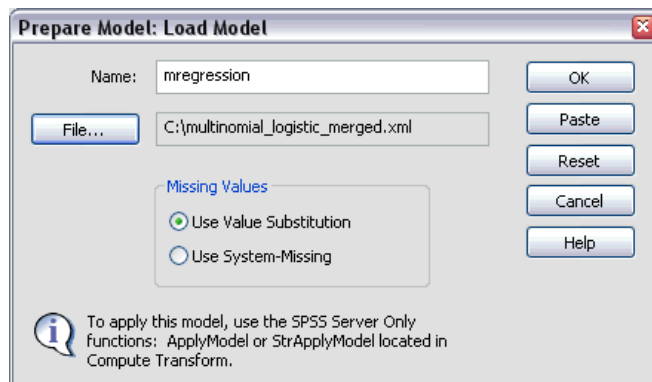
Figura 8-33  
Resultado de la carga de modelos

SPSS Variables	Model Variables					
	Name	Label	Type	Width	Role	Measurement
Age_Group	Age_Group	Age group	Numeric	8	Factor	Nominal
Frequency	Frequency	Number of purchases	Numeric	7	Covariate	Continuous
Gender	Gender	Gender	String	6	Factor	Nominal
Has_Broadband	Has_Broadband	Has home broadband connection	Numeric	10	Factor	Nominal
Has_Child	Has_Child	Has children	Numeric	10	Factor	Nominal
Income_Group	Income_Group	Income group	Numeric	10	Factor	Nominal
Log_Amount	Log_Amount	Log-transformed amount	Numeric	8	Covariate	Continuous
Recency	Recency	Days since last purchase	Numeric	10	Covariate	Continuous
Response	Response	Response	Numeric	10	Target	Nominal

### Para cargar un modelo

- ▶ Elija en los menús:  
Transformar  
Preparar modelo  
Cargar modelo...

Figura 8-34  
Cuadro de diálogo Preparar modelo: Cargar modelo



- ▶ Escriba un nombre para asociarlo a este modelo. Cada modelo cargado debe tener un nombre exclusivo.
- ▶ Pulse en Archivo y seleccione un archivo de modelo. El cuadro de diálogo Abrir archivo resultante muestra los archivos disponibles en análisis en modo distribuido. Esto incluye los archivos del equipo donde se instaló el servidor de SPSS y los archivos del equipo local que se encuentran en carpetas compartidas o en unidades compartidas.

*Nota:* Al puntuar datos, el modelo se aplica a las variables del conjunto de datos activo con los mismos nombres que las variables del archivo de modelo. Puede asignar variables del modelo original a otras variables del conjunto de datos activo mediante la sintaxis de comandos (consulte el comando MODEL HANDLE).

**Nombre.** Nombre utilizado para identificar este modelo. Las normas para los nombres de modelo válidos son las mismas que para los [nombres de variable](#) (consulte Nombres de variable en Capítulo 5 en p. 82), con la adición del carácter \$ como un primer carácter permitido. Este nombre se utiliza para especificar el modelo cuando se puntúan los datos con las funciones ApplyModel o StrApplyModel.

**Archivo.** El archivo XML (PMML) que contiene las especificaciones del modelo.

### **Valores perdidos**

Este grupo de opciones controla el tratamiento de los valores perdidos, que se encuentran durante el proceso de puntuación, para las variables predictoras del modelo. Un valor perdido en el contexto de puntuación hace referencia a:

- Una variable predictora no contiene ningún valor. Para las variables numéricas, esto significa el valor perdido del sistema. Para las variables de cadena, esto significa una cadena nula.
- El valor se ha definido como perdido por el usuario, en el modelo, para el predictor dado. Los valores definidos como perdidos por el usuario en el conjunto de datos activo, pero no en el modelo, no se tratan como valores perdidos en el proceso de puntuación.
- La variable predictora es categórica y el valor no es una de las categorías definidas en el modelo.

**Usar sustitución de valores.** Puede utilizar la sustitución de valores cuando puntúe casos con valores perdidos. El método para determinar el valor que sustituye a un valor perdido depende del tipo de modelo predictivo.

- **SPSS Modelos.** Para las variables independientes en modelos de regresión lineal y discriminantes, si se especificó la sustitución por el valor medio para los valores perdidos cuando se generó y guardó el modelo, dicho valor medio se utiliza en lugar del valor perdido en el cálculo de puntuación y la puntuación continua. Si el valor medio no está disponible, el valor perdido del sistema se devuelve.
- **Modelos de AnswerTree y modelos del comando TREE.** Para los modelos CHAID y CHAID exhaustivo, se selecciona el nodo filial de mayor tamaño para una variable de segmentación perdida. El nodo filial de mayor tamaño es el que tiene mayor población entre los nodos filiales que utilizan los casos de muestra de aprendizaje. Para los modelos C&RT y QUEST, las variables de segmentación sustitutas (si las hay) se utilizan primero. (Las divisiones sustitutas son divisiones que intentan coincidir con la división original tanto como sea posible utilizando predictores alternativos.) Si no se especifica ninguna división sustituta o todas las variables de segmentación sustitutas corresponden a valores perdidos, se utiliza el nodo filial de mayor tamaño.
- **Modelos Clementine.** Los modelos de regresión lineal se tratan como se describe en los modelos de SPSS. Los modelos de regresión logística se tratan como se describe en los modelos de regresión logística. Los modelos de árbol C&RT se tratan como se describe para los modelos C&RT bajo los modelos de AnswerTree.
- **Modelos de regresión logística.** Para las covariables de los modelos de regresión logística, si un valor medio del predictor se incluye como parte del modelo guardado, este valor medio se utiliza en lugar del valor perdido en el cálculo de puntuación y la puntuación continua. Si el predictor es categórico (por ejemplo, un factor en un modelo de regresión logística) o si el valor medio no está disponible, se devolverá el valor perdido del sistema.

**Usar perdido por el sistema.** Devuelve el valor perdido por el sistema al puntuar un caso con un valor perdido.

### ***Visualización de una lista de modelos cargados***

Puede obtener una lista de los modelos cargados actualmente. En los menús (sólo disponible en el análisis en modo distribuido) elija:

Transformar  
Preparar modelo  
Listar modelos

Esto generará una tabla de asas de modelo. La tabla contiene una lista de todos los modelos cargados actualmente e incluye el nombre (denominado asa de modelo) asignado al modelo, el tipo de modelo, la ruta de acceso al archivo de modelo y el método para tratar los valores perdidos.

Figura 8-35  
*Lista de modelos cargados*

Model Handles				
Name	Type	Input File Name	Validation handling	Missing handling
mregression	NOMREG	C:\models\mregression.xml	ON	SUBSTITUTE
tree	TREE	C:\models\tree.xml	ON	SUBSTITUTE

### ***Funciones adicionales con sintaxis de comandos***

En el cuadro de diálogo Cargar modelo, puede pegar las selecciones en una ventana de sintaxis y editar la sintaxis de comandos MODEL HANDLE resultante. Esto permite:

- Asignar variables del modelo original a distintas variables del conjunto de datos activo (con el subcomando MAP). Por defecto, el modelo se aplica a las variables del conjunto de datos activo con los mismos nombres que las variables del archivo de modelo.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

# Gestión y transformación de los archivos

Los archivos de datos no siempre están organizados de la forma ideal para las necesidades específicas del usuario. Puede que le interese combinar archivos de datos, organizar los datos en un orden diferente, seleccionar un subconjunto de casos o cambiar la unidad de análisis agrupando casos. Entre la amplia gama de posibilidades de transformación de archivos disponibles se encuentran las siguientes:

**Ordenar datos.** Puede ordenar los casos en función del valor de una o más variables.

**Transponer casos y variables.** El formato de archivo de datos de SPSS lee las filas como casos y las columnas como variables. Para los archivos de datos en los que el orden está invertido, se pueden intercambiar las filas y las columnas para leer los datos en el formato correcto.

**Fundir archivos.** Puede fundir dos o más archivos de datos. Es posible combinar archivos con las mismas variables pero con casos distintos, o con los mismos casos pero variables diferentes.

**Seleccionar subconjuntos de casos.** Puede restringir el análisis a un subconjunto de casos o efectuar análisis simultáneos de subconjuntos diferentes.

**Agregar datos.** Puede cambiar la unidad de análisis agregando casos basados en el valor de una o más variables de agrupación.

**Ponderar datos.** Puede ponderar los casos para un análisis basado en el valor de una variable de ponderación.

**Reestructurar datos.** Puede reestructurar los datos para crear un único caso (registro) a partir de varios casos o crear varios casos a partir de un único caso.

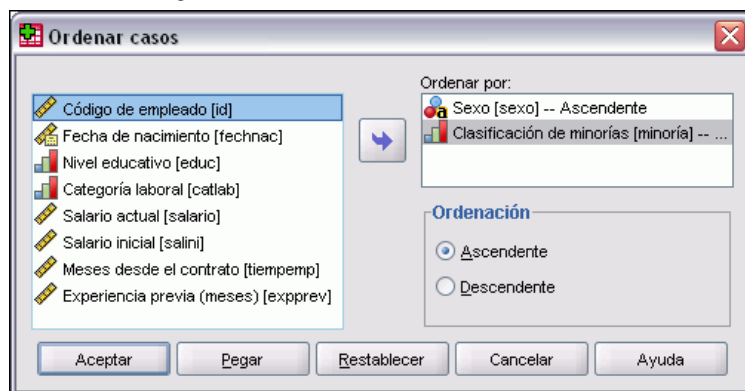
## Ordenar casos

Este cuadro de diálogo ordena los casos (las filas) del archivo de datos basándose en los valores de una o más variables de ordenación. Puede ordenar los casos en orden ascendente o descendente.

- Si selecciona más de una variable de ordenación, los casos se ordenarán por variable dentro de las categorías de la variable anterior de la lista Ordenar por. Por ejemplo, si selecciona *Sexo* como la primera variable de ordenación y *Minoría* como la segunda, los casos se ordenarán por minorías dentro de cada categoría de sexo.
- La secuencia de ordenación está basada en el orden definido de forma regional (y no tiene por qué ser igual al orden numérico de los códigos de caracteres). La configuración regional por defecto es la configuración regional del sistema operativo. Puede controlar la configuración

regional con el ajuste Idioma de la pestaña General del cuadro de diálogo Opciones (menú Edición).

Figura 9-1  
Cuadro de diálogo Ordenar casos



### Para ordenar casos

- ▶ Elija en los menús:  
Datos  
Ordenar casos...
- ▶ Seleccione una o más variables de ordenación.

## Ordenación de variables

Puede ordenar las variables del conjunto de datos activo en función de los valores de cualquiera de los atributos de variable (por ejemplo, nombre de la variable, tipo de variable, tipo de datos, nivel de medida), incluidos los atributos de variable personalizados.

- Los valores se pueden ordenar en orden ascendente o descendente.
- Puede guardar la variable original (ordenada previamente) en un atributo de variable personalizado.
- La ordenación por valores de atributos de variable personalizados se limita a los atributos de variable personalizados que están visibles actualmente en la Vista de variables.

Si desea obtener más información sobre los atributos de variable personalizados, consulte [Atributos de variable personalizados](#).

### Para ordenar variables

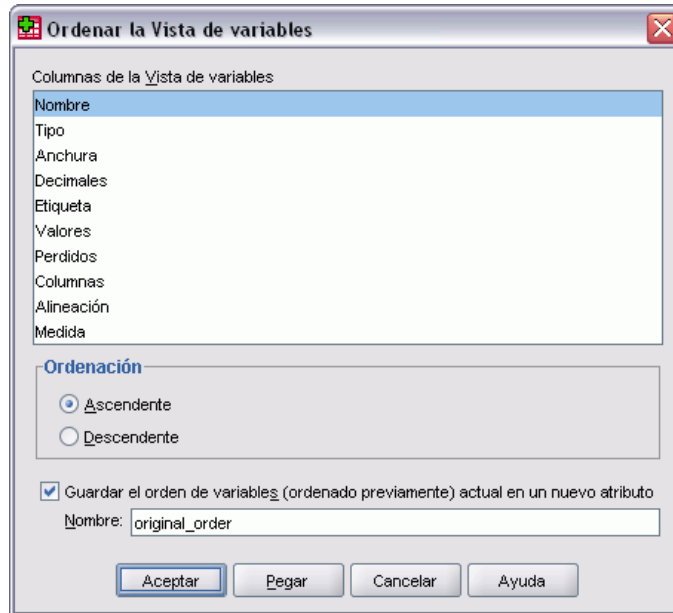
En la Vista de variables del Editor de datos:

- ▶ Pulse con el botón derecho del ratón en el encabezado de columna Atributo y elija en el menú contextual Ordenar de forma ascendente u Ordenar de forma descendente.

o

- ▶ En los menús de cualquier ventana, seleccione:  
Datos  
Ordenar variables
- ▶ Seleccione el atributo que desea utilizar para ordenar variables.
- ▶ Seleccione el orden de orden (ascendente o descendente).

Figura 9-2  
Cuadro de diálogo Ordenar variables



- La lista de atributos de variable coincide con los nombres de la columna de atributos representada en la Vista de variables del Editor de datos.
- Puede guardar la variable original (ordenada previamente) en un atributo de variable personalizado. Para cada variable, el valor del atributo es un valor entero que indica su posición antes de la ordenación; de manera que al ordenar las variables en función del valor de dicho atributo personalizado se puede restaurar su orden original.

## Transponer

Transponer crea un archivo de datos nuevo en el que se transponen las filas y las columnas del archivo de datos original de manera que los casos (las filas) se convierten en variables, y las variables (las columnas) se convierten en casos. También crea automáticamente nombres de variable y presenta una lista de dichos nombres.

- Se crea automáticamente una nueva variable de cadena, *case\_lbl*, que contiene el nombre de variable original.

- Si el conjunto de datos activo contiene una variable de identificación o de nombre con valores únicos, podrá utilizarla como variable de nombre: sus valores se emplearán como nombres de variable en el archivo de datos transpuesto. Si se trata de una variable numérica, los nombres de variable comenzarán por la letra V, seguida de un valor numérico.
- Los valores perdidos definidos por el usuario se convierten en el valor perdido del sistema en el archivo de datos transpuesto. Para conservar cualquiera de estos valores, se debe cambiar la definición de los valores perdidos en la Vista de variables del Editor de datos.

#### **Para transponer variables y casos**

- ▶ Elija en los menús:  
Datos  
Transponer...
- ▶ Seleccione la variable o variables que desee transponer en casos.

## **Fusión de archivos de datos**

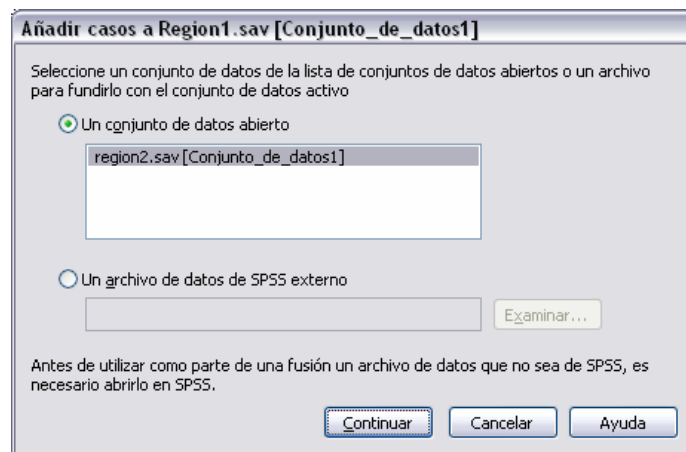
Es posible unir los datos de dos archivos de dos maneras diferentes. Tiene la posibilidad de:

- Fundir el conjunto de datos activo con otro conjunto de datos abierto o archivo de datos con formato SPSS que contenga las mismas variables pero diferentes casos.
- Fundir el conjunto de datos activo con otro conjunto de datos abierto o archivo de datos con formato SPSS que contenga los mismos casos pero diferentes variables.

#### **Para fundir archivos**

- ▶ Elija en los menús:  
Datos  
Fundir archivos
- ▶ Seleccione Añadir casos o Añadir variables.

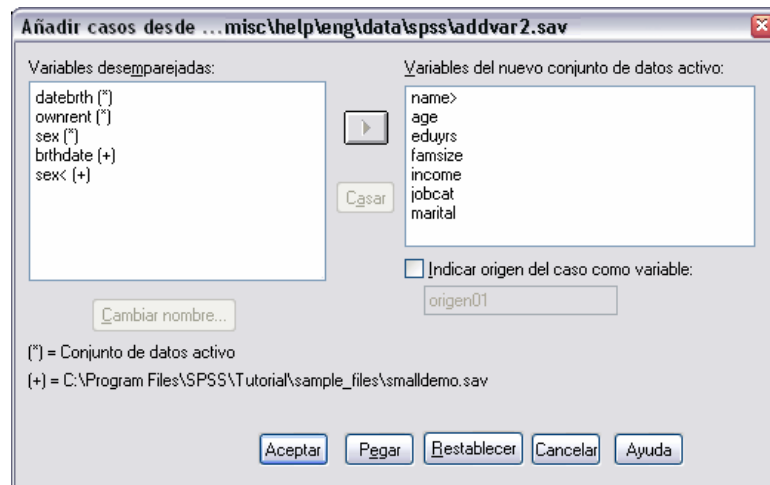
Figura 9-3  
*Selección de archivos para fundir*



## Añadir casos

Añadir casos fusiona el conjunto de datos activo con un segundo conjunto de datos o un archivo de datos con formato SPSS que contenga las mismas variables (columnas) pero diferentes casos (filas). Por ejemplo, podría registrar la misma información de los clientes de dos zonas de venta diferentes y conservar los datos de cada zona en archivos distintos. El segundo conjunto de datos puede ser un archivo de datos con formato SPSS externo o un conjunto de datos disponible en la sesión actual.

Figura 9-4  
Cuadro de diálogo Añadir casos



**Variables desemparejadas.** Muestra las variables que se van a excluir del nuevo archivo de datos fusionado. Las variables del conjunto de datos activo se identifican mediante un asterisco (\*). Las variables del otro conjunto de datos se identifican con un signo más (+). Por defecto, la lista contiene:

- Las variables de cualquiera de los archivos de datos que no coincidan con un nombre de variable del otro archivo. Puede crear pares a partir de variables desemparejadas e incluirlos en el nuevo archivo fusionado.
- Las variables definidas como datos numéricos en un archivo y como datos de cadena en el otro. Las variables numéricas no pueden fusionarse con variables de cadena.
- Variables de cadena de longitud diferente. El ancho definido de una variable de cadena debe ser el mismo en ambos archivos de datos.

**Variables del nuevo conjunto de datos activo.** Variables que se van a incluir en el nuevo archivo de datos fusionado. Por defecto, la lista incluye todas las variables que coinciden en el nombre y el tipo de datos (numéricos o de cadena).

- Puede eliminar de la lista las variables que no desee incluir en el archivo fusionado.
- Las variables desemparejadas incluidas en el archivo fusionado contendrán los datos perdidos para los casos del archivo que no contiene esa variable.

**Indicar origen del caso como variable.** Indica, para cada caso, el archivo de datos de origen. Esta variable toma un valor 0 para los casos del archivo de datos de trabajo y un valor 1 para los casos del archivo de datos externo.



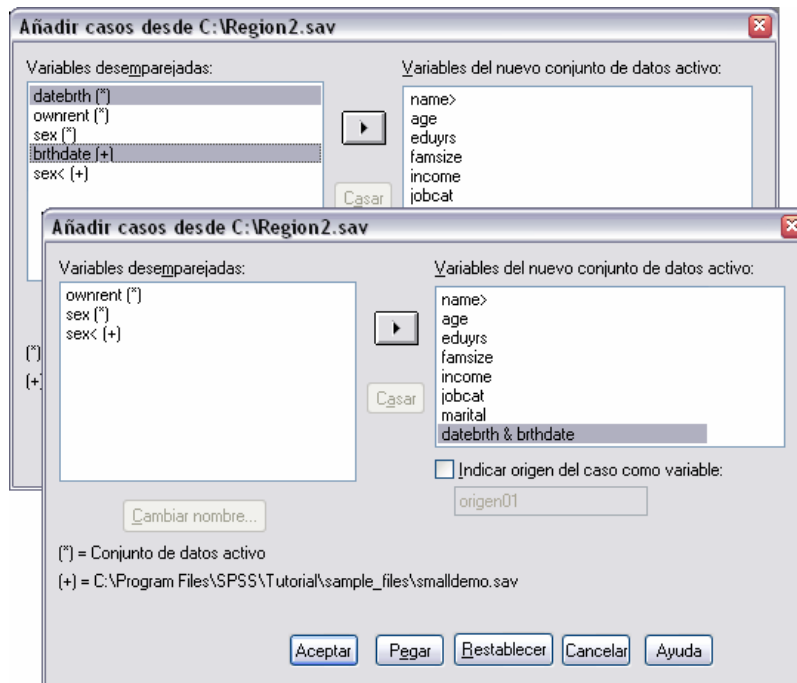
**Para fundir archivos de datos con las mismas variables y casos diferentes**

- ▶ Abra al menos uno de los archivos de datos que desea fusionar. Si tiene varios conjuntos de datos abiertos, convierta uno de los conjuntos de datos que desea fusionar en el conjunto de datos activo. Los casos de este archivo aparecerán primero en el nuevo archivo de datos fusionado.
- ▶ Elija en los menús:
  - Datos
  - Fundir archivos
  - Añadir casos...
- ▶ Seleccione el conjunto de datos o el archivo de datos con formato SPSS que va a fusionar con el conjunto de datos activo.
- ▶ Elimine de la lista Variables del nuevo conjunto de datos activo cualquier variable que no desee incluir.
- ▶ Añada parejas de variables de la lista Variables desemparejadas que representen la misma información registrada con nombres diferentes en los dos archivos. Por ejemplo, la fecha de nacimiento podría tener el nombre de variable *fechnac* en un archivo y *nacfech* en el otro.

**Para seleccionar una pareja de variables desemparejadas**

- ▶ Pulse en una de las variables en la lista Variables desemparejadas.
- ▶ Mantenga pulsada la tecla Ctrl mientras selecciona la otra variable de la lista con el ratón (pulse al mismo tiempo la tecla Ctrl y el botón izquierdo del ratón).
- ▶ Pulse en Casar para desplazar el par de variables a la lista Variables del nuevo conjunto de datos activo. (El nombre de variable del conjunto de datos activo se empleará como el nombre de variable en el archivo fusionado.)

Figura 9-5  
Selección de parejas de variables mediante la tecla Ctrl



### **Añadir casos: Cambiar nombre**

Puede cambiar los nombres de las variables del conjunto de datos activo o de otro conjunto de datos antes de desplazarlas desde la lista de variables desparejadas a la lista de variables que se van a incluir en el archivo de datos fusionado. Cambiar el nombre de las variables le permite:

- Utilizar el nombre de variable del otro conjunto de datos en lugar del nombre del conjunto de datos activo para las parejas de variables.
- Incluir dos variables con el mismo nombre pero de diferentes tipos o longitudes de cadena. Por ejemplo, para incluir la variable numérica *sexo* del conjunto de datos activo y la variable de cadena *sexo* del otro conjunto de datos, primero se debe cambiar el nombre de una de ellas.

### **Añadir casos: Información del diccionario**

Toda información del diccionario (etiquetas de variable y de valor, valores definidos como perdidos por el usuario, formatos de presentación) existente en el conjunto de datos activo se aplicará al archivo de datos fusionado.

- Si alguna información del diccionario sobre una variable no está definida en el conjunto de datos activo, se utilizará la información del diccionario del otro conjunto de datos.
- Si el conjunto de datos activo contiene cualquier etiqueta de valor definida o valores definidos como perdidos por el usuario para una variable, se ignorará cualquier otra etiqueta de valor o valor definido como perdido por el usuario para esa variable en el otro conjunto de datos.

## Fusión de más de dos orígenes de datos

Puede fusionar hasta 50 conjuntos de datos y/o archivos de datos con la sintaxis de comandos. Si desea obtener información, consulte el comando `ADD FILES` en la referencia de sintaxis de comandos (*Command Syntax Reference*) (disponible en el menú Ayuda).

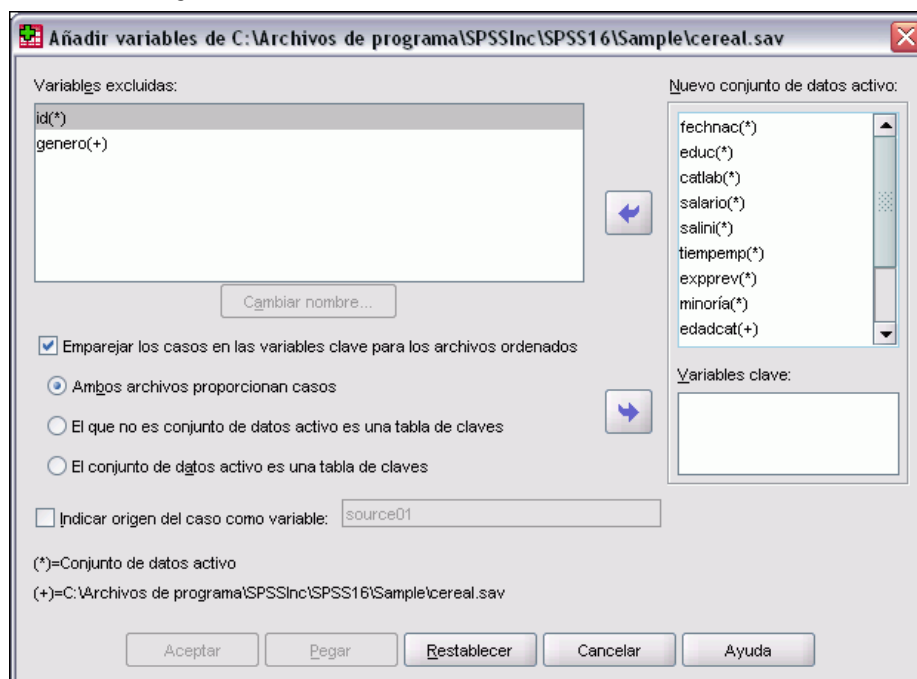
## Añadir variables

Añadir variables fusiona el conjunto de datos activo con otro conjunto de datos abierto o un archivo de datos con formato SPSS que contenga los mismos casos (filas) pero diferentes variables (columnas). Por ejemplo, es posible que desee fusionar un archivo de datos que contenga los resultados previos de la prueba con otro que contenga los resultados posteriores.

- Los casos deben estar ordenados de la misma manera en ambos conjuntos de datos.
- Si se utilizan una o más variables clave para emparejar los casos, los dos conjuntos de datos deben estar ordenados por orden ascendente de la variable o variables clave.
- Los nombres de las variables del segundo archivo de datos que son duplicados de los del conjunto de datos activo se excluyen por defecto, ya que Añadir variables supone que estas variables contienen información duplicada.

**Indicar origen del caso como variable.** Indica, para cada caso, el archivo de datos de origen. Esta variable toma un valor 0 para los casos del archivo de datos de trabajo y un valor 1 para los casos del archivo de datos externo.

Figura 9-6  
Cuadro de diálogo Añadir variables



**Variables excluidas.** Muestra las variables que se van a excluir del nuevo archivo de datos fusionado. Por defecto, la lista contiene los nombres de variable de otro conjunto de datos que son duplicados de los del conjunto de datos. Las variables del conjunto de datos activo se identifican mediante un asterisco (\*). Las variables del otro conjunto de datos se identifican con un signo más (+). Si desea incluir en el archivo fusionado una variable excluida con un nombre duplicado, cámbiele el nombre y añádala a la lista de variables que se van a incluir.

**Nuevo conjunto de datos activo.** Variables que se van a incluir en el nuevo conjunto de datos fusionado. Por defecto, se incluyen en la lista todos los nombres de variable únicos que existan en ambos conjuntos de datos.

**Variables clave.** Si algunos casos de un conjunto de datos no se emparejan con los del otro conjunto de datos (es decir, si faltan casos en un conjunto de datos), utilice las variables clave para identificar y emparejar correctamente los casos de ambos conjuntos de datos. También puede utilizar las variables clave según tablas de claves.

- Las variables clave deben tener los mismos nombres en ambos conjuntos de datos.
- Ambos conjuntos de datos deben estar ordenados según el orden ascendente de las variables clave, y el orden de las variables de la lista Variables clave debe ser igual a su secuencia de ordenación.
- Los casos que no se emparejan con las variables clave se incluyen en el archivo fusionado, pero no se funden con los casos del otro archivo. Los casos no emparejados sólo contienen valores para las variables del archivo de procedencia; las variables del otro archivo contienen el valor perdido del sistema.

**El que no es conjunto de datos activo (o el conjunto de datos activo) es una tabla de claves.** Una tabla de claves, o **tabla de referencia**, es un archivo en el que los datos de cada “caso” se pueden aplicar a varios casos del otro archivo de datos. Por ejemplo, si un archivo contiene información sobre los diferentes miembros de la familia (como el sexo, la edad, la formación) y el otro contiene información global (como los ingresos totales, el número de miembros o la ubicación), se puede utilizar el archivo global como una tabla de referencia y aplicar los datos comunes de la familia a cada uno de sus miembros en el archivo fusionado.

#### ***Para fundir archivos con los mismos casos pero variables diferentes***

- ▶ Abra al menos uno de los archivos de datos que desea fusionar. Si tiene varios conjuntos de datos abiertos, convierta uno de los conjuntos de datos que desea fusionar en el conjunto de datos activo.
- ▶ Elija en los menús:
  - Datos
  - Fundir archivos
  - Añadir variables...
- ▶ Seleccione el conjunto de datos o el archivo de datos con formato SPSS que va a fusionar con el conjunto de datos activo.

#### ***Para seleccionar variables clave***

- ▶ Seleccione las variables entre las variables del archivo externo (marcadas con el signo +) en la lista Variables excluidas.

- ▶ Seleccione Emparejar los casos en las variables clave para los archivos ordenados.
- ▶ Añada las variables a la lista Variables clave.

Las variables clave deben existir en el conjunto de datos activo y en el otro conjunto de datos. Ambos conjuntos de datos deben estar ordenados según el orden ascendente de las variables clave, y el orden de las variables de la lista Variables clave debe ser igual a su secuencia de ordenación.

### ***Añadir variables: Cambiar nombre***

Puede cambiar los nombres de las variables del conjunto de datos activo o de otro conjunto de datos antes de desplazarlas a la lista de variables que se van a incluir en el archivo de datos fusionado. Esta medida es especialmente útil cuando se desea incluir dos variables que tienen el mismo nombre pero información diferente en los dos archivos.

### ***Fusión de más de dos orígenes de datos***

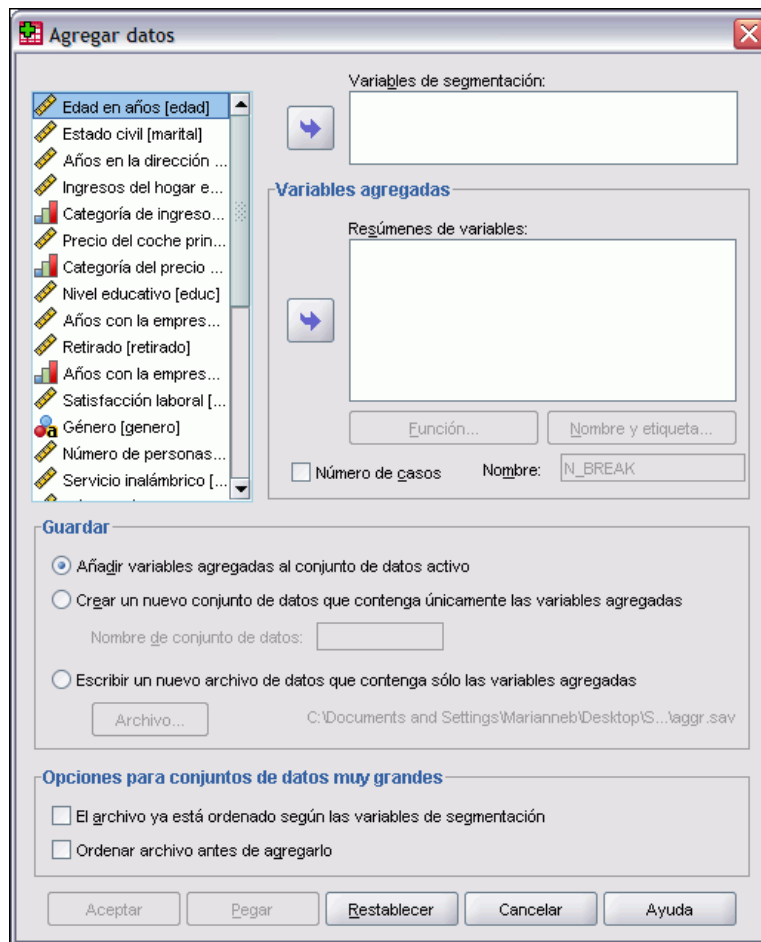
Puede fusionar hasta 50 conjuntos de datos y/o archivos de datos con la sintaxis de comandos. Si desea obtener información, consulte el comando `MATCH FILES` en la referencia de sintaxis de comandos (*Command Syntax Reference*) (disponible en el menú Ayuda).

### ***Agregar datos***

Agregar datos agrega grupos de casos en el conjunto de datos activo en casos individuales y crea un archivo nuevo agregado o variables nuevas en el conjunto de datos activo que contiene los datos agregados. Los casos se agregan en función del valor de una o más variables de segmentación (agrupación).

- Si crea un archivo de datos agregado nuevo, dicho archivo de datos nuevo contiene un caso para cada grupo definido por las variables de segmentación. Por ejemplo, si hay una variable de segmentación con dos valores, el archivo de datos nuevo contiene sólo dos casos.
- Si añade variables agregadas al conjunto de datos activo, no se agrega el archivo de datos. Cada caso con los mismos valores de variables de segmentación recibe los mismos valores para las nuevas variables agregadas. Por ejemplo, si *sexo* es la única variable de segmentación, todos los hombres reciben el mismo valor para la variable agregada nueva que representa la edad media.

Figura 9-7  
Cuadro de diálogo Agregar datos



**Variabes de segmentación.** Los casos se agrupan en función de los valores de las variables de segmentación. Cada combinación única de valores de variables de segmentación define un grupo. Al crear un archivo de datos agregados nuevo, todas las variables de segmentación se guardan en el archivo nuevo con sus nombres y la información del diccionario. La variable de segmentación puede ser tanto numérica como de cadena.

**Variabes agregadas.** Las variables de origen se utilizan con funciones agregadas para crear variables agregadas nuevas. El nombre de la variable agregada viene seguido de una etiqueta de variable opcional, el nombre de la función de agregación y el nombre de la variable de origen entre paréntesis.

Puede anular los nombres por defecto de las variables agregadas con nuevos nombres de variable, proporcionar etiquetas de variable descriptivas y cambiar las funciones empleadas para calcular los valores de los datos agregados. También puede crear una variable que contenga el número de casos en cada grupo de segmentación.

### **Para agregar un archivo de datos**

- ▶ Elija en los menús:
  - Datos
  - Agregar...
- ▶ Seleccione una o más variables de segmentación que definan cómo deben agruparse los casos para crear datos agregados.
- ▶ Seleccione una o varias variables para incluir.
- ▶ Seleccione una función de agregación para cada variable agregada.

### **Almacenamiento de resultados agregados**

Puede añadir variables agregadas al conjunto de datos activo o crear un archivo de datos agregados nuevo.

- **Añadir variables agregadas al conjunto de datos activo.** Las nuevas variables basadas en las funciones de agregación se añaden al archivo de datos de trabajo. El propio archivo de datos no se agrega. Cada caso con los mismos valores de variables de segmentación recibe los mismos valores para las nuevas variables agregadas.
- **Crear un nuevo conjunto de datos que contenga únicamente las variables agregadas.** Guarda los datos agregados en un nuevo conjunto de datos durante la sesión actual. El conjunto de datos incluye las variables de segmentación que definen los casos agregados y todas las variables de agregación definidas por las funciones de agregación. El conjunto de datos activo no se ve afectado.
- **Escribir un nuevo archivo de datos que contenga sólo las variables agregadas.** Guarda los datos agregados en un archivo de datos externo. El archivo incluye las variables de segmentación que definen los casos agregados y todas las variables agregadas definidas por las funciones de agregación. El conjunto de datos activo no se ve afectado.

### **Opciones de ordenación para archivos de datos grandes**

En el caso de los archivos de datos muy grandes, puede resultar más eficiente agregar datos ordenados previamente.

**El archivo ya está ordenado según las variables de segmentación.** Si los datos ya han sido ordenados por los valores de las variables de segmentación, esta opción permite al procedimiento una ejecución más rápida y utilizar menos memoria. Utilice esta opción con precaución.

- Los datos se deben ordenar por valores de variables de segmentación en el mismo orden que las variables de segmentación especificadas para el procedimiento Agregar datos.
- Si va a añadir variables al conjunto de datos activo, seleccione sólo esta opción si los datos se han ordenado mediante valores ascendentes de las variables de segmentación.

**Ordenar archivo antes de agregarlo.** En situaciones muy extrañas y con archivos de datos voluminosos, puede ser necesario ordenar el archivo de datos por los valores de las variables de segmentación antes de realizar la agregación. No se recomienda esta opción a menos que se presenten problemas de memoria y/o rendimiento.

## Agregar datos: Función de agregación

Este cuadro de diálogo permite especificar la función que se utilizará para calcular los valores de los datos agregados para las variables seleccionadas en la lista Agregar variables, en el cuadro de diálogo Agregar datos. Las funciones de agregación incluyen:

- Funciones de resumen para variables numéricas, incluyendo la media, la mediana, la desviación típica y la suma
- Número de casos, incluyendo los no ponderados, los ponderados, los no perdidos y los perdidos
- Porcentaje o fracción de los valores por encima o por debajo de un valor especificado
- Porcentaje o fracción de los valores dentro o fuera de un rango especificado

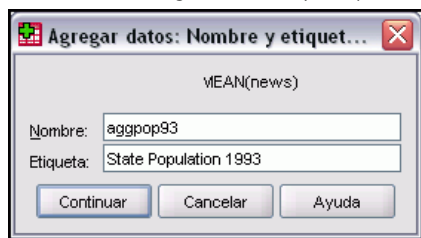
Figura 9-8  
Cuadro de diálogo Función de agregación

## Agregar datos: Nombre y etiqueta de variable

Agregar datos asigna nombres de variable por defecto a las variables agregadas al nuevo archivo de datos. Este cuadro de diálogo le permite cambiar el nombre de variable de la variable seleccionada en la lista Agregar variables y proporcionar una etiqueta de variable descriptiva. Si desea obtener más información, consulte [Nombres de variable](#) en Capítulo 5 en p. 82.



Figura 9-9  
Cuadro de diálogo Nombre y etiqueta de variable

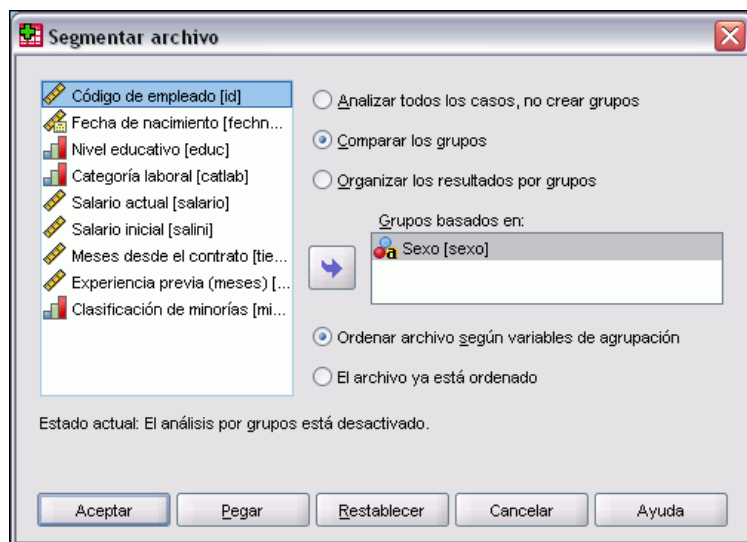


## Segmentar archivo

Segmentar archivo divide el archivo de datos en distintos grupos para el análisis basándose en los valores de una o más variables de agrupación. Si selecciona varias variables de agrupación, los casos se agruparán por variable dentro de las categorías de la variable anterior de la lista Grupos basados en. Por ejemplo, si selecciona *sexo* como la primera variable de agrupación y *minoría* como la segunda, los casos se agruparán por minorías dentro de cada categoría de sexo.

- Es posible especificar hasta ocho variables de agrupación.
- Cada ocho bytes de una variable de cadena larga (variables de cadena que superan los ocho bytes) cuenta como una variable hasta llegar al límite de ocho variables de agrupación.
- Los casos deben ordenarse según los valores de las variables de agrupación, en el mismo orden en el que aparecen las variables en la lista Grupos basados en. Si el archivo de datos todavía no está ordenado, seleccione Ordenar archivo según variables de agrupación.

Figura 9-10  
Cuadro de diálogo Segmentar archivo



**Comparar los grupos.** Los grupos de segmentación del archivo se presentan juntos para poder compararlos. Para las tablas pivote se crea una sola tabla y cada variable de segmentación del archivo puede desplazarse entre las dimensiones de la tabla. En el caso de los gráficos se crea un gráfico diferente para cada grupo y se muestran juntos en el Visor.

**Organizar los resultados por grupos.** Los resultados de cada procedimiento se muestran por separado para cada grupo de segmentación del archivo.

***Para segmentar un archivo de datos para el análisis***

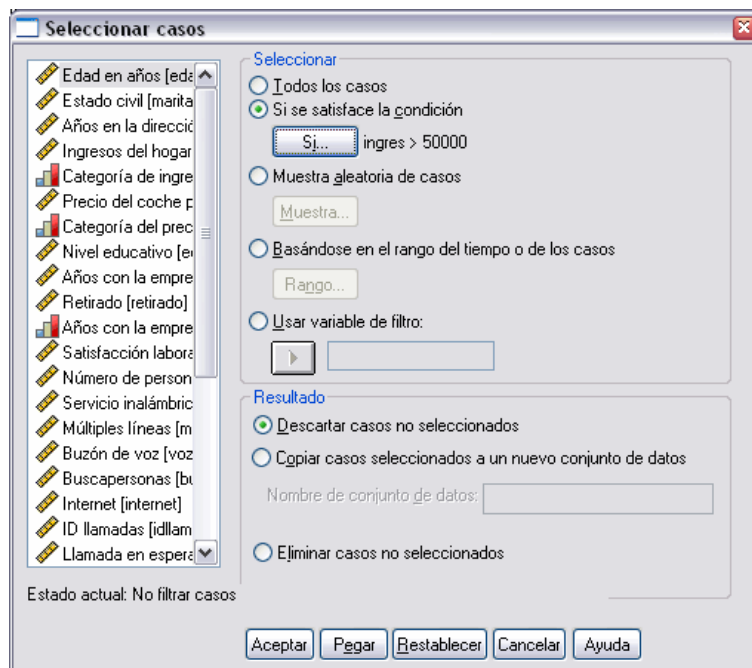
- ▶ Elija en los menús:  
Datos  
Segmentar archivo...
- ▶ Seleccione Comparar los grupos u Organizar los resultados por grupos.
- ▶ Seleccione una o más variables de agrupación.

## ***Seleccionar casos***

Seleccionar casos proporciona varios métodos para seleccionar un subgrupo de casos basándose en criterios que incluyen variables y expresiones complejas. También se puede seleccionar una muestra aleatoria de casos. Los criterios usados para definir un subgrupo pueden incluir:

- Valores y rangos de las variables
- Rangos de fechas y horas
- Números de caso (filas)
- Expresiones aritméticas
- Expresiones lógicas
- Funciones

Figura 9-11  
Cuadro de diálogo Seleccionar casos



**Todos los casos.** Desactiva el filtrado y utiliza todos los casos.

**Si se satisface la condición.** Utiliza una expresión condicional para seleccionar los casos. Si el resultado de la expresión condicional es verdadero, el caso se selecciona. Si el resultado es falso o perdido, entonces el caso no se selecciona.

**Muestra aleatoria de casos.** Selecciona una muestra aleatoria basándose en un porcentaje aproximado o en un número exacto de casos.

**Basándose en el rango del tiempo o de los casos.** Selecciona los casos basándose en un rango de los números de caso o en un rango de las fechas/horas.

**Usar variable de filtro.** Utiliza como variable para el filtrado la variable numérica seleccionada del archivo de datos. Se seleccionan los casos con cualquier valor distinto del 0 o del valor perdido para la variable seleccionada.

### **Resultados**

Esta sección controla el tratamiento de casos no seleccionados. Puede elegir una de las siguientes alternativas para tratar los casos no seleccionados:

- **Descartar casos no seleccionados.** Los casos no seleccionados no se incluyen en el análisis, pero se conservan en el conjunto de datos. Podrá utilizar los casos no seleccionados más adelante en la sesión, si desactiva el filtrado. Si selecciona una muestra aleatoria o si selecciona los casos mediante una expresión condicional, se generará una variable con el nombre *filter\_\$* que tendrá el valor 1 para los casos seleccionados y el valor 0 para los casos no seleccionados.
- **Copiar casos seleccionados a un nuevo conjunto de datos.** Los casos seleccionados se copiarán a un nuevo conjunto de datos, lo que mantendrá inalterado el conjunto de datos original. Los casos no seleccionados no se incluirán en el nuevo conjunto de datos y se mantendrán en su estado original en el conjunto de datos original.
- **Eliminar casos no seleccionados.** Los casos no seleccionados se eliminarán del conjunto de datos. Sólo se pueden recuperar los casos eliminados saliendo del archivo sin guardar ningún cambio y abriéndolo de nuevo. La eliminación de los casos será permanente si se guardan los cambios en el archivo de datos.

*Nota:* Si elimina los casos no seleccionados y guarda el archivo, no será posible recuperar estos casos.

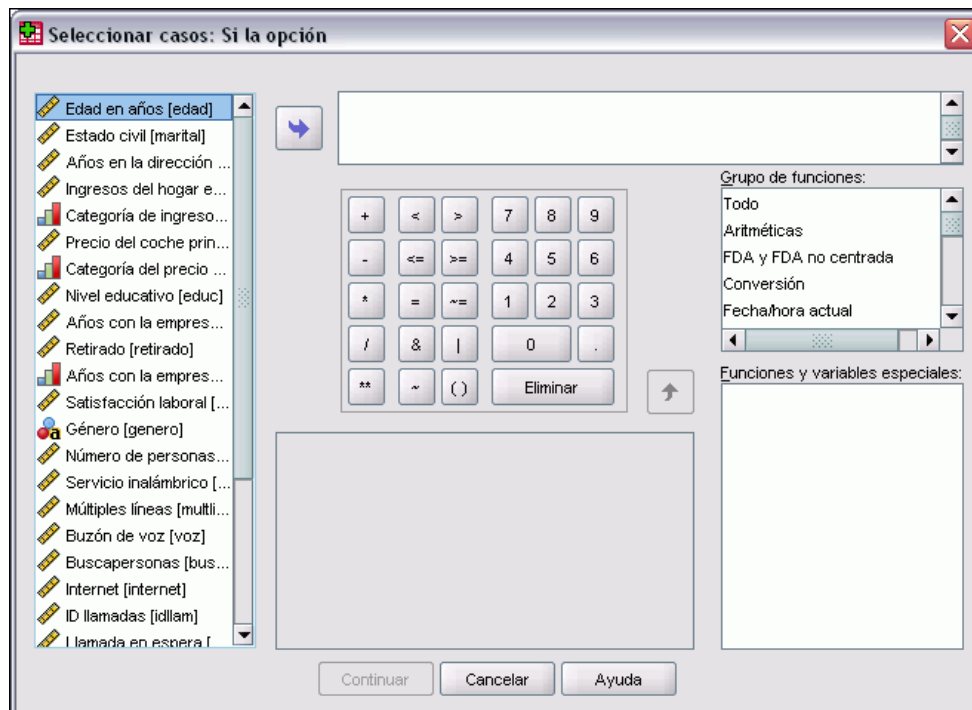
### **Para seleccionar un subconjunto de casos**

- ▶ Elija en los menús:
  - Datos
  - Seleccionar casos...
- ▶ Seleccione uno de los métodos de selección de casos.
- ▶ Especifique los criterios para la selección de casos.

## Seleccionar casos: Si la opción

Este cuadro de diálogo permite seleccionar subconjuntos de casos utilizando expresiones condicionales. Una expresión condicional devuelve un valor *verdadero*, *falso* o *perdido* para cada caso.

Figura 9-12  
Cuadro de diálogo Seleccionar casos: Si

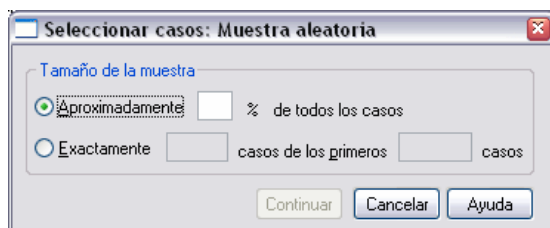


- Si el resultado de una expresión condicional es *verdadero*, se incluirá el caso en el subconjunto seleccionado.
- Si el resultado de una expresión condicional es *falso* o *perdido*, no se incluirá el caso en el subconjunto seleccionado.
- La mayoría de las expresiones condicionales utilizan al menos uno de los seis operadores de relación (<, >, <=, >=, =, and ~=) de la calculadora.
- Las expresiones condicionales pueden incluir nombres de variable, constantes, operadores aritméticos, funciones numéricas (y de otros tipos), variables lógicas y operadores de relación.

## Seleccionar casos: Muestra aleatoria

Este cuadro de diálogo permite seleccionar una muestra aleatoria basada en un porcentaje aproximado o en un número exacto de casos. El muestreo se realiza sin sustitución, de manera que el mismo caso no se puede seleccionar más de una vez.

Figura 9-13  
Cuadro de diálogo Seleccionar casos: Muestra aleatoria



**Aproximadamente.** Genera una muestra aleatoria con el porcentaje aproximado de casos indicado. Dado que esta rutina toma una decisión pseudo-aleatoria para cada caso, el porcentaje de casos seleccionados sólo se puede aproximar al especificado. Cuantos más casos contenga el archivo de datos, más se acercará el porcentaje de casos seleccionados al porcentaje especificado.

**Exactamente.** Un número de casos especificado por el usuario. También se debe especificar el número de casos a partir de los cuales se generará la muestra. Este segundo número debe ser menor o igual que el número total de casos presentes en el archivo de datos. Si lo excede, la muestra contendrá un número menor de casos proporcional al número solicitado.

## Seleccionar casos: Amplitud

Este cuadro de diálogo selecciona los casos basándose en un rango de números de caso o en un rango de fechas u horas.

- Los rangos de casos se basan en el número de fila que se muestra en el Editor de datos.
- Los rangos de fechas y horas sólo están disponibles para **datos de serie temporal** con variables de fecha definidas (menú Datos, Definir fechas).

Figura 9-14  
Cuadro de diálogo Seleccionar casos: Rango para rangos de casos (sin variables de fecha definidas)

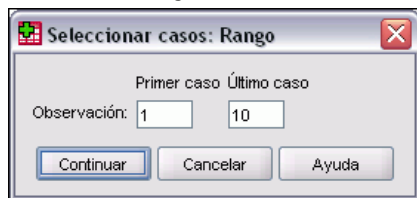


Figura 9-15  
Cuadro de diálogo Seleccionar casos: Rango para datos de la serie temporal con variables de fecha definidas



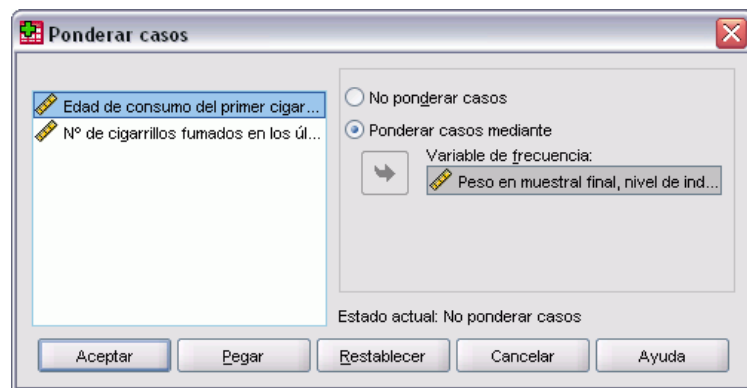
## Ponderar casos

Ponderar casos proporciona a los casos diferentes ponderaciones (mediante una réplica simulada) para el análisis estadístico.

- Los valores de la variable de ponderación deben indicar el número de observaciones representadas por casos únicos en el archivo de datos.
- Los casos con valores perdidos, negativos o cero para la variable de ponderación se excluyen del análisis.
- Los valores fraccionarios son válidos; se usan exactamente donde adquieren sentido y, con mayor probabilidad, donde se tabulan los casos.

Figura 9-16

Cuadro de diálogo Ponderar casos



Si aplica una variable de ponderación, ésta seguirá vigente hasta que se seleccione otra o se desactive la ponderación. Si guarda un archivo de datos ponderado, la información de ponderación se guardará con el archivo. Puede desactivar la ponderación en cualquier momento, incluso después de haber guardado el archivo de forma ponderada.

**Ponderaciones en las tablas de contingencia.** El procedimiento Tablas de contingencia cuenta con diversas opciones para el tratamiento de ponderaciones de los casos. Si desea obtener más información, consulte [Tablas de contingencia: Mostrar en las casillas](#) en Capítulo 16 en p. 286.

**Ponderaciones en los diagramas de dispersión y los histogramas.** Los diagramas de dispersión y los histogramas tienen una opción para activar y desactivar las ponderaciones de los casos, pero dicha opción no afecta a los casos que tienen un valor negativo, un valor 0 o un valor perdido para la variable de ponderación. Estos casos permanecen excluidos del gráfico incluso si se desactiva la ponderación desde el gráfico.

### Para ponderar casos

- ▶ Elija en los menús:  
Datos  
Ponderar casos...
- ▶ Seleccione Ponderar casos mediante.
- ▶ Seleccione una variable de frecuencia.

Los valores de la variable de frecuencia se utilizan como ponderaciones de los casos. Por ejemplo, un caso con un valor 3 para la variable de frecuencia representará tres casos en el archivo de datos ponderado.

## ***Reestructuración de los datos***

Utilice el Asistente de reestructuración de datos para reestructurar los datos de acuerdo con el procedimiento que desee utilizar. El asistente sustituye el archivo actual con un archivo nuevo reestructurado. El asistente puede:

- Reestructurar variables seleccionadas en casos
- Reestructurar casos seleccionados en variables
- Transponer todos los datos

### ***Para reestructurar datos***

- ▶ Elija en los menús:
  - Datos
  - Reestructurar...
- ▶ Seleccione el tipo de reestructuración que desea realizar.
- ▶ Seleccione los datos que se van a reestructurar.

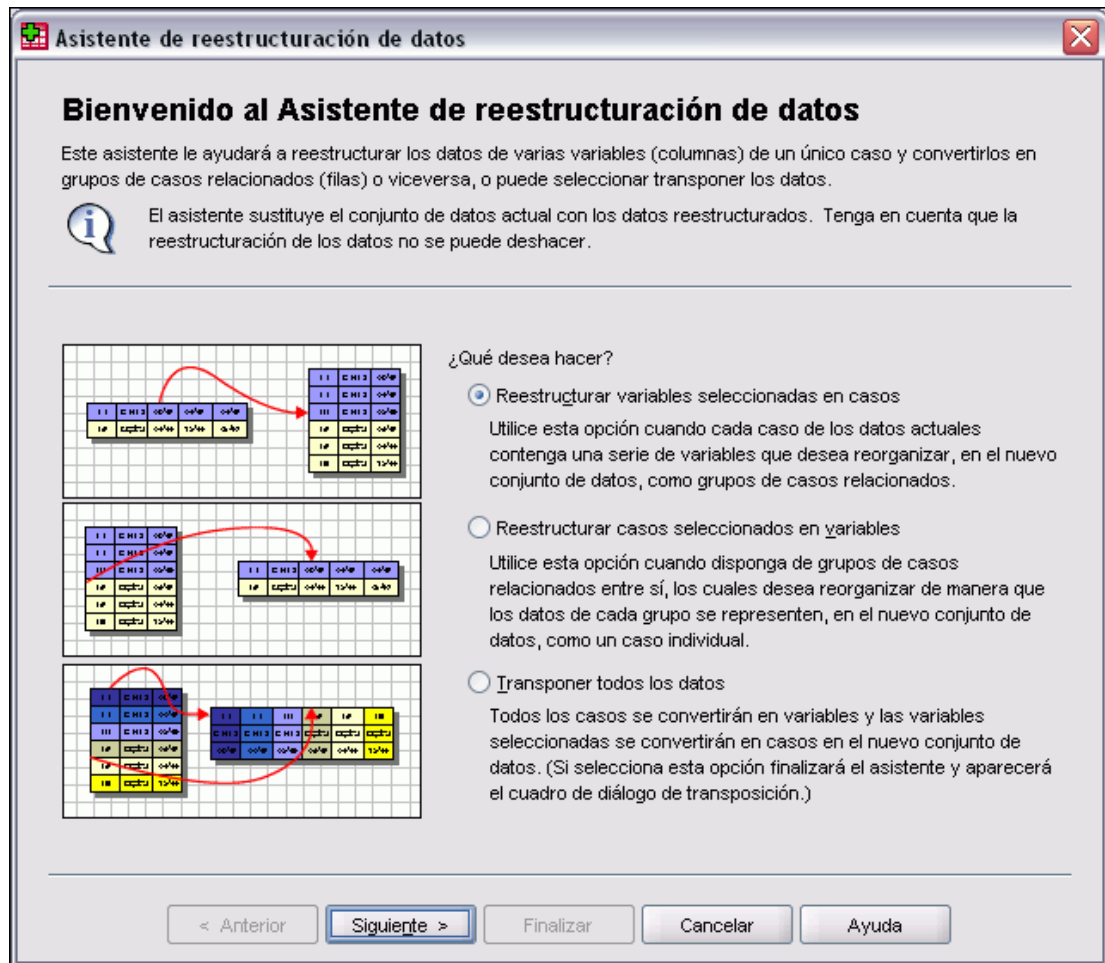
Si lo desea, puede:

- Crear variables de identificación, que permitirán un seguimiento de un valor del nuevo archivo a partir de un valor del archivo original
- Ordenar los datos antes de la reestructuración
- Definir opciones para el nuevo archivo
- Pegar la sintaxis de comandos en una ventana de sintaxis

### ***Asistente de reestructuración de datos: Seleccionar tipo***

Utilice el Asistente de reestructuración de datos para reestructurar los datos. En el primer cuadro de diálogo, seleccione el tipo de reestructuración que desea llevar a cabo.

Figura 9-17  
Asistente de reestructuración de datos



- **Reestructurar variables seleccionadas en casos.** Seleccione esta opción cuando disponga, en los datos, de grupos de columnas relacionadas y desee que aparezcan en el nuevo archivo de datos como grupos de filas. Si elige esta opción, el asistente mostrará los pasos para **Variables a casos**.
- **Reestructurar casos seleccionados en variables.** Seleccione esta opción cuando disponga, en los datos, de grupos de filas relacionadas y desee que aparezcan en el nuevo archivo de datos como grupos de columnas. Si elige esta opción, el asistente mostrará los pasos para **Casos a variables**.
- **Transponer todos los datos.** Seleccione esta opción cuando desee transponer los datos. Todas las filas se convertirán en columnas y todas las columnas en filas, en el nuevo archivo de datos. Esta opción cierra el Asistente de reestructuración de datos y abre el cuadro de diálogo Transponer datos.



### Opciones de reestructuración de los datos

Una **variable** contiene información que se desea analizar, por ejemplo, una medida o una puntuación. Un **caso** es una observación, por ejemplo, un individuo. En una estructura de datos *simple*, cada variable es una única columna de datos y cada caso es una única fila. De manera que, por ejemplo, si estuviera midiendo las puntuaciones de un examen realizado a todos los alumnos de una clase, todos los valores de las notas aparecerían en una única columna y habría una fila para cada alumno.

Cuando se analizan datos, a menudo se está analizando cómo varía una variable en función de cierta condición. Dicha condición puede ser un tratamiento experimental específico, un grupo demográfico, un momento en el tiempo u otra cosa. En el análisis de datos, a las condiciones de interés a menudo se las denomina **factores**. Al analizar factores, se dispone de una estructura de datos *compleja*. Es posible que haya información acerca de una variable en más de una columna de datos (por ejemplo, una columna para cada nivel de un factor), o que haya información acerca de un caso en más de una fila (por ejemplo, una fila para cada nivel de un factor). El Asistente de reestructuración de datos le ayuda a reestructurar archivos con una estructura de datos compleja.

La estructura del archivo actual y la estructura que se desea en el nuevo archivo determinan las elecciones que se deben seleccionar en el asistente.

**¿Cómo están organizados los datos en el archivo actual?** Es posible que los datos actuales estén organizados de manera que los factores estén registrados en una variable *diferente* (como grupos de casos) o *con* la variable (como grupos de variables).

- **Grupos de casos.** ¿El archivo actual tiene registradas las variables y las condiciones en columnas diferentes? Por ejemplo:

var	factor
8	1
9	1
3	2
1	2

En este ejemplo, las dos primeras filas son un **grupo de casos** porque están relacionadas. Contienen datos para el mismo nivel del factor. En el análisis de datos de SPSS, cuando los datos están estructurados de esta manera, se hace referencia al factor como **variable de agrupación**.

- **Grupos de columnas.** ¿El archivo actual tiene registradas las variables y las condiciones en la misma columna? Por ejemplo:

var_1	var_2
8	3
9	1

En este ejemplo, las dos primeras columnas son un **grupo de variables** porque están relacionadas. Contienen datos para la misma variable, *var\_1* para el nivel 1 del factor y *var\_2* para el nivel 2 del factor. En el análisis de datos de SPSS, si los datos se estructuran de esta manera, el factor se suele denominar **de medidas repetidas**.

**¿Cómo deben organizarse los datos en el archivo nuevo?** Normalmente, la organización estará determinada por el procedimiento que se vaya a utilizar para analizar los datos.

- **Procedimientos que requieren grupos de casos.** Los datos deberán estructurarse en grupos de casos para realizar los análisis que requieran una variable de agrupación. Algunos ejemplos son: *univariante*, *multivariante* y *componentes de la varianza* de los Modelos lineales generales; Modelos mixtos; Cubos OLAP; y *muestras independientes* de las Pruebas T o Pruebas no paramétricas. Si la estructura de datos actual es de grupos de variables y desea realizar estos análisis, seleccione Reestructurar variables seleccionadas en casos.
- **Procedimientos que requieren grupos de variables.** Los datos se deberán estructurar en grupos de variables para analizar medidas repetidas. Algunos ejemplos son: *medidas repetidas* de los Modelos lineales generales, análisis de *covariables dependientes del tiempo* del Análisis de regresión de Cox, *muestras relacionadas* de las Pruebas T o *muestras relacionadas* de las Pruebas no paramétricas. Si la estructura de datos actual es de grupos de casos y desea realizar estos análisis, seleccione Reestructurar casos seleccionados en variables.

### Ejemplo de variables a casos

En este ejemplo, las puntuaciones de las pruebas están registradas en columnas diferentes para cada factor, A y B.

Figura 9-18

Datos actuales para reestructurar variables a casos

	punt_a	punt_b
1	1014.00	864.00
2	684.00	636.00
3	810.00	638.00

Se desea realizar una prueba *t* para muestras independientes. Se dispone de un grupo de columnas compuesto por *puntuación\_a* y *puntuación\_b*, pero no se dispone de la **variable de agrupación** que requiere el procedimiento. Seleccione Reestructurar variables seleccionadas en casos en el Asistente de reestructuración de datos, reestructure un grupo de variables en una nueva variable denominada *puntuación* y cree un índice denominado *grupo*. El nuevo archivo de datos se muestra en la siguiente imagen.

Figura 9-19

Datos nuevos y reestructurados para variables a casos

	grupo	punt
1	PUNT_A	1014.00
2	PUNT_B	864.00
3	PUNT_A	684.00
4	PUNT_B	636.00
5	PUNT_A	810.00
6	PUNT_B	638.00

Cuando se ejecute la prueba *t* para muestras independientes, podrá utilizar *grupo* como variable de agrupación.

### Ejemplo de casos a variables

En este ejemplo, las puntuaciones de las pruebas están registradas dos veces para cada sujeto, antes y después de un tratamiento.

Figura 9-20  
Datos actuales para reestructurar casos a variables

	id	punt	tiempo
1	1	1014.00	bef
2	1	864.00	aft
3	2	684.00	bef
4	2	636.00	aft

Se desea realizar una prueba *t* para muestras relacionadas. La estructura de datos es de grupos de casos, pero no se dispone de las **medidas repetidas** para las variables relacionadas que requiere el procedimiento. Seleccione Reestructurar casos seleccionados en variables en el Asistente de reestructuración de datos, utilice *id* para identificar los grupos de filas en los datos actuales y utilice *tiempo* para crear el grupo de variables en el nuevo archivo.

Figura 9-21  
Datos nuevos y reestructurados para casos a variables

	id	ant	des
1	1	864.00	1014.00
2	2	636.00	684.00

Cuando se ejecute la prueba *t* de muestras relacionadas, podrá utilizar *ant* y *des* como el par de variables.

### Asistente de reestructuración de datos (variables a casos): Número de grupos de variables

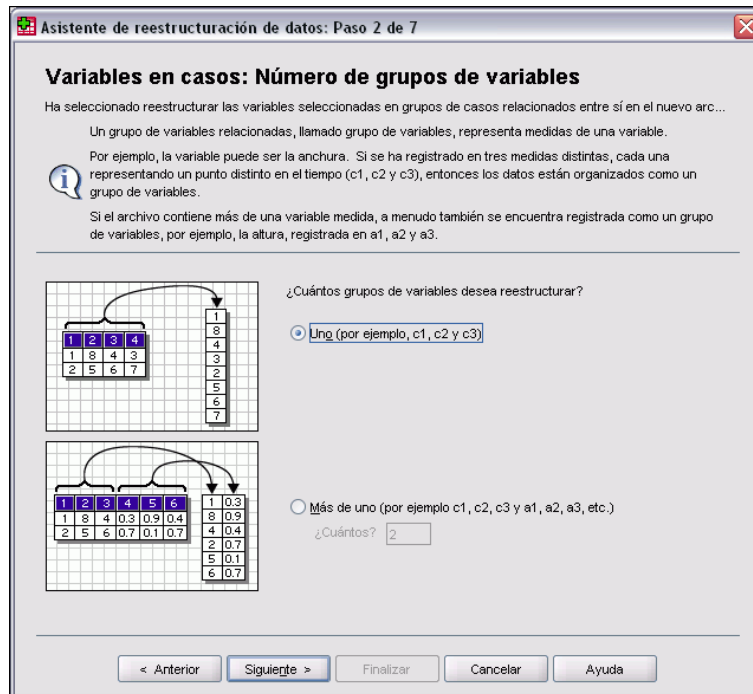
*Nota:* el asistente presenta este paso si se ha seleccionado reestructurar grupos de variables en filas.

En este paso, se debe elegir el número de grupos de variables del archivo actual que se desea reestructurar en el nuevo archivo.

**¿Cuántos grupos de variables hay en el archivo actual?** Piense cuántos grupos de variables existen en los datos actuales. Un grupo de columnas relacionadas, llamado **grupo de variables**, registra medidas repetidas de la misma variable en distintas columnas. Por ejemplo, si en los datos actuales hay tres columnas, *c1*, *c2* y *c3*, que registran el **contorno**, entonces hay un grupo de variables. Si además hay otras tres columnas, *a1*, *a2* y *a3*, que registran la **altura**, entonces hay dos grupos de variables.

**¿Cuántos grupos de variables debe haber en el archivo nuevo?** Considere cuántos grupos de variables desea que estén representados en el nuevo archivo de datos, teniendo en cuenta que no es necesario reestructurar todos los grupos de variables en el nuevo archivo.

Figura 9-22  
Asistente de reestructuración de datos: Número de grupos de variables, paso 2



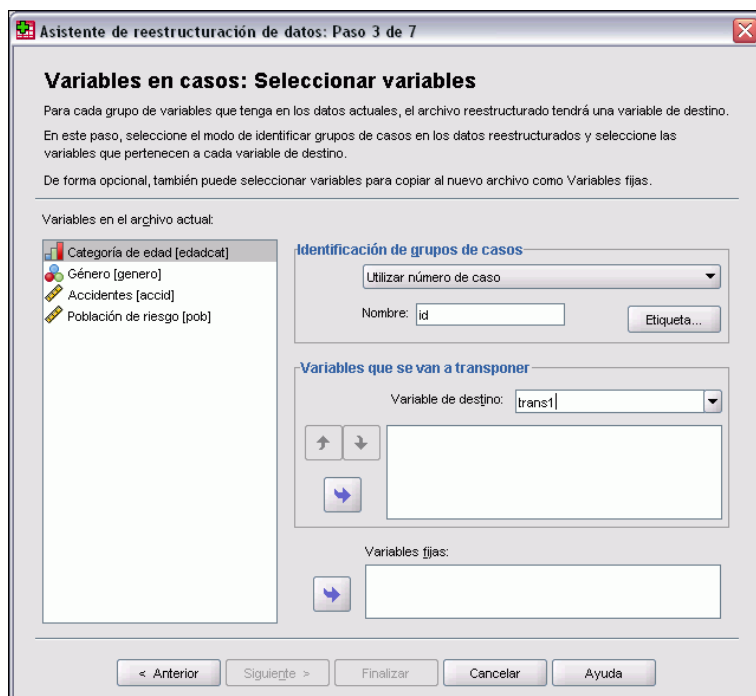
- **Uno.** El asistente creará una única variable reestructurada en el nuevo archivo a partir de un grupo de variables del archivo actual.
- **Más de uno.** El asistente creará varias variables reestructuradas en el nuevo archivo. El número que se especifique afectará al siguiente paso, en el que el asistente creará de forma automática el número especificado de nuevas variables.

### ***Asistente de reestructuración de datos (variables a casos): Seleccionar Variables***

*Nota:* el asistente presenta este paso si se ha seleccionado reestructurar grupos de variables en filas.

En este paso, se debe proporcionar información sobre cómo se van a utilizar las variables del archivo actual en el nuevo archivo. También se puede crear una variable que identifique las filas en el nuevo archivo.

Figura 9-23  
Asistente de reestructuración de datos: Seleccionar variables, paso 3



**¿Cómo se deben identificar las nuevas filas?** En el nuevo archivo de datos, puede crear una variable que identifique la fila del archivo de datos actual que ha sido utilizada para crear un grupo de filas nuevo. El identificador puede ser un número de caso secuencial o los valores de una variable. Utilice los controles disponibles en el apartado Identificación de grupos de casos para definir la variable de identificación utilizada en el nuevo archivo. Pulse en la casilla para cambiar el nombre de variable por defecto y para dotar a la variable de identificación de una etiqueta de variable descriptiva.

**¿Qué se debe reestructurar en el nuevo archivo?** En el paso anterior, se informó al asistente del número de grupos de variables que se deseaba reestructurar. El asistente creó una nueva variable para cada grupo. Los valores para el grupo de variables aparecerán en dicha variable en el nuevo archivo. Utilice los controles en Variables que se van a transponer para definir la variable reestructurada en el nuevo archivo.

**Para especificar una variable reestructurada**

- ▶ Ponga las variables que componen el grupo de variables que desea transformar en la lista Variables que se van a transponer. Todas las variables del grupo deberán ser del mismo tipo (numéricas o de cadena).

Se puede incluir la misma variable más de una vez en el grupo de variables (las variables se copian de la lista origen de variables en lugar de moverlas); los valores se repetirán en el nuevo archivo.

**Para especificar varias variables reestructuradas**

- ▶ Seleccione la primera variable de destino que desea definir de la lista desplegable Variable de destino.
- ▶ Ponga las variables que componen el grupo de variables que desea transformar en la lista Variables que se van a transponer. Todas las variables del grupo deberán ser del mismo tipo (numéricas o de cadena). Puede incluir la misma variable más de una vez en el grupo de variables. (Las variables se copian de la lista origen de variables en lugar de moverlas, y los valores se repetirán en el nuevo archivo.)
- ▶ Seleccione la siguiente variable de destino que desea definir y repita el proceso de selección de variables para todas las variables de destino disponibles.
  - Aunque puede incluir la misma variable más de una vez en el mismo grupo de variables de destino, no puede incluir la misma variable en más de un grupo de variables de destino.
  - Cada lista de grupos de variables de destino debe contener el mismo número de variables. (Las variables que aparecen más de una vez se incluyen en el recuento).
  - El número de grupos de variables de destino está determinado por el número de grupos de variables especificados en el paso anterior. Aquí puede cambiar los nombres de las variables por defecto, pero deberá volver al paso anterior para cambiar el número de grupos de variables que se van a reestructurar.
  - Debe definir los grupos de variables (seleccionando variables de la lista de origen) para todas las variables de destino disponibles antes de poder pasar al siguiente paso.

**¿Qué se debe copiar en el nuevo archivo?** En el nuevo archivo se pueden copiar variables que no se han reestructurado. Sus valores se propagarán en las nuevas filas. Desplace las variables que desea copiar en el nuevo archivo en la lista Variables fijas.

**Asistente de reestructuración de datos (variables a casos): Crear variables de Variables**

*Nota:* el asistente presenta este paso si se ha seleccionado reestructurar grupos de variables en filas.

En este paso, se debe decidir si se crean variables de índice. Un índice es una nueva variable que identifica de forma secuencial un grupo de filas en función de la variable original a partir de la cual se creó la nueva fila.

Figura 9-24  
Asistente de reestructuración de datos: Crear variables de índice, paso 4

**Variables en casos: Crear variables de índice**

En los datos actuales, los valores de un grupo de variables se encuentran en un caso individual recogido en varias variables. Por ejemplo, un caso individual contiene los valores de c1, c2 y c3.

En los nuevos datos, los valores de un grupo de variables se encontrarán en varios casos recogidos en una variable individual. Por ejemplo, habrá tres casos para c1, c2 y c3, respectivamente.

Un índice es una nueva variable que identifica el grupo de nuevos casos creados a partir del caso original. Por ejemplo, un índice llamado "c" tendría los valores 1, 2 y 3.

¿Cuántas variables de índice desea crear?

Prueba  
Utilice esta opción cuando un grupo de variables registre los efectos de un único factor, tratamiento o condición.

Más de una ¿Cuántos?   
Utilice esta opción cuando un grupo de variables registre los efectos de más de un factor, tratamiento o condición.

Ninguna  
Utilice esta opción si la información de índice se almacena en uno de los conjuntos de variables que se va a transponer.

< Anterior   Siguiente >   Finalizar   Cancelar   Ayuda

**¿Cuántas variables de índice debe haber en el archivo nuevo?** Las variables de índice se pueden utilizar como variables de agrupación en los procedimientos. En la mayoría de los casos, es suficiente una única variable de índice; no obstante, si los grupos de variables del archivo actual reflejan varios niveles de factor, puede ser conveniente utilizar varios índices.

- **Uno.** El asistente creará una única variable de índice.
- **Más de uno.** El asistente creará varios índices y deberá introducir el número de índices que desea crear. El número especificado afectará al siguiente paso, en el que el asistente crea de forma automática el número especificado de índices.
- **Ninguno.** Seleccione esta opción si no desea crear variables de índice en el nuevo archivo.

### Ejemplo de un índice para variables a casos

En los datos actuales, hay un grupo de variables, denominado *contorno*, y un factor, el *tiempo*. El contorno se ha medido en tres ocasiones y se ha registrado en *c1*, *c2* y *c3*.

Figura 9-25  
Datos actuales para un índice

	sujeto	w1	w2	w3
1	1	6.70	4.30	5.70
2	2	7.10	5.90	5.60

Se va a reestructurar el grupo de variables en una única variable, *contorno*, y se va a crear un único índice numérico. Los nuevos datos se muestran en la siguiente tabla.

Figura 9-26  
Datos nuevos y reestructurados con un índice

	sujeto	índice	anchura
1	1	1	6.70
2	1	2	4.30
3	1	3	5.70
4	2	1	7.10
5	2	2	5.90
6	2	3	5.60

El *Índice* comienza por 1 y se incrementa por cada variable del grupo. Vuelve a comenzar cada vez que se encuentra una fila en el archivo original. Ahora se puede utilizar *índice* en procedimientos que requieran una variable de agrupación.

### Ejemplo de dos índices para variables a casos

Cuando un grupo de variables registra más de un factor, se puede crear más de un índice; no obstante, se deben organizar los datos actuales de forma que los niveles del primer factor sean un índice primario dentro del cual varían los niveles de los siguientes factores. En los datos actuales, hay un grupo de variables, denominado *contorno*, y dos factores, *A* y *B*. Los datos se organizan de manera que los niveles del factor *B* varían dentro de los niveles del factor *A*.

Figura 9-27  
Datos actuales para dos índices

	sujeto	w_a1b1	w_a1b2	w_a2b1	w_a2b2
1	1	5.50	6.40	5.80	5.90
2	2	7.40	7.10	5.60	6.70

Se va a reestructurar el grupo de variables en una única variable, *contorno*, y se van a crear dos índices. Los nuevos datos se muestran en la siguiente tabla.

Figura 9-28  
Datos nuevos reestructurados con los dos índices

	sujeto	índice_a	índice_b	anchura
1	1	1	1	5.50
2	1	1	2	6.40
3	1	2	1	5.80
4	1	2	2	5.90
5	2	1	1	7.40
6	2	1	2	7.10
7	2	2	1	5.60
8	2	2	2	6.70

### Asistente de reestructuración de datos (variables a casos): Crear una variable de índices

*Nota:* el asistente presenta este paso si se ha seleccionado reestructurar grupos de variables en filas y crear una variable de índices.



En este paso, se debe decidir los valores que se desean para la variable de índice. Los valores pueden ser números secuenciales o los nombres de las variables en un grupo de variables original. También puede especificar un nombre y una etiqueta para la nueva variable de índice.

Figura 9-29

Asistente de reestructuración de datos: Crear una variable de índice, paso 5

**Variables en casos: Crear una variable de índices**

Ha seleccionado crear una variable de índice. Los valores de la variable pueden ser números secuenciales o los nombres de las variables en un grupo.

En la tabla se puede especificar el nombre y la etiqueta para la variable de índice.

¿Qué tipo de valores de índice?

Números secuenciales  
Valores de índice: 1, 2

Los nombres de las variables  
Valores de índice: genero, edadcat

Editar el nombre y la etiqueta de la variable de índice:

	Nombre	Etiqueta	Niveles	Valores de índice
1	índice1		2	1, 2

< Anterior   Siguiente >   Finalizar   Cancelar   Ayuda

Si desea obtener más información, consulte [Ejemplo de un índice para variables a casos](#) en p. 207.

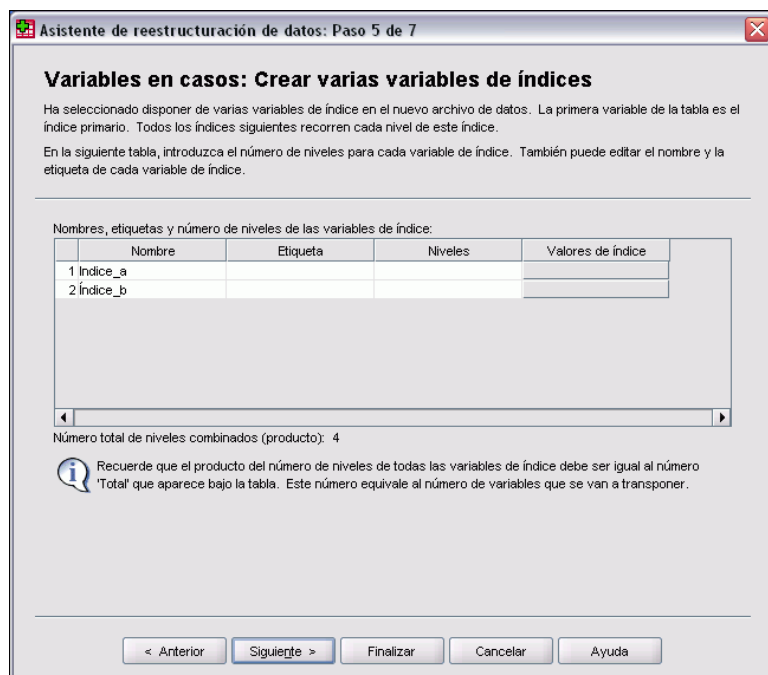
- **Números secuenciales.** El asistente asignará de forma automática números secuenciales como valores índice.
- **Los nombres de las variables.** El asistente utilizará los nombres del grupo de variables seleccionado como valores índice. Seleccione un grupo de variables de la lista.
- **Nombres y etiquetas.** Pulse en una casilla para cambiar el nombre de variable por defecto y proporcionar una etiqueta de variable descriptiva para la variable de índice.

### ***Asistente de reestructuración de datos (variables a casos): Crear varias variables de índices***

*Nota:* el asistente presenta este paso si se ha seleccionado reestructurar grupos de variables en filas y crear varias variables de índice.

En este paso, se debe especificar el número de niveles para cada variable de índice. También puede especificar un nombre y una etiqueta para la nueva variable de índice.

Figura 9-30  
Asistente de reestructuración de datos: Crear varias variables de índice, paso 5



Si desea obtener más información, consulte [Ejemplo de dos índices para variables a casos](#) en p. 208.

**¿Cuántos niveles hay registrados en el archivo actual?** Calcule los niveles de factor que hay registrados en los datos actuales. Un **nivel** define un grupo de casos que experimentan las mismas condiciones. Si hay varios factores, los datos actuales se deben organizar de manera que los niveles del primer factor sean un índice primario dentro del cual varían los niveles de los siguientes factores.

**¿Cuántos niveles debe haber en el archivo nuevo?** Introduzca el número de niveles para cada índice. Los valores para varias variables de índice son siempre números secuenciales. Los valores comienzan en 1 y se incrementan con cada nivel. El primer índice se incrementa más despacio y el último más deprisa.

**Número total de niveles combinados.** No se puede crear más niveles de los que existen en los datos actuales. Como los datos reestructurados contendrán una fila por cada combinación de tratamientos, el asistente realizará una comprobación del número de niveles que se crean. Comparará el producto de los niveles creados con el número de variables del grupo de variables. Deben coincidir.

**Nombres y etiquetas.** Pulse en una casilla para cambiar el nombre de variable por defecto y proporcionar una etiqueta de variable descriptiva para las variables de índice.

### ***Asistente de reestructuración de datos (variables a casos): Opciones***

*Nota:* el asistente presenta este paso si se ha seleccionado reestructurar grupos de variables en filas.

En este paso, se deben especificar las opciones para el nuevo archivo reestructurado.

Figura 9-31  
Asistente de reestructuración de datos: Opciones, paso 6

Asistente de reestructuración de datos: Paso 6 de 7

**Variables en casos: Opciones**

En este paso puede definir las opciones que se aplicarán al archivo de datos reestructurado.

**Tratamiento de variables no seleccionadas**

Colocar variables del nuevo archivo de datos

Conservar y tratar como variables fijas

**Valores vacíos o perdidos por el sistema en todas las variables transpuestas**

Crear un caso en el nuevo archivo

Descartar los datos

**Variable de recuento de casos**

Contar el número de casos nuevos creados por el caso de los datos actuales

Nombre:

Etiqueta:

< Anterior   Siguiente >   Finalizar   Cancelar   Ayuda

**¿Desea eliminar las variables no seleccionadas?** En el paso de selección de variables (paso 3), se seleccionaron los grupos de variables que se iban a reestructurar, las variables que se iban a copiar y una variable de identificación de los datos actuales. Los datos de las variables seleccionadas aparecerán en el nuevo archivo. Si hay más variables en los datos actuales, puede elegir descartarlas o conservarlas.

**¿Desea conservar los datos perdidos?** El asistente comprueba cada nueva fila potencial en busca de valores nulos. Un **valor nulo** es un valor en blanco o perdido por el sistema. Se puede elegir entre conservar o descartar las filas que contienen sólo valores nulos.

**¿Desea crear una variable de recuento?** El asistente puede crear una **variable de recuento** en el nuevo archivo. Dicha variable contiene el número de nuevas filas generadas por una fila de los datos actuales. Una variable de recuento puede ser de gran utilidad si decide descartar del nuevo archivo los valores nulos, ya que esto conlleva la generación de un número distinto de nuevas filas por una fila dada de los datos actuales. Pulse en una casilla para cambiar el nombre de variable por defecto y proporcionar una etiqueta de variable descriptiva para la variable de recuento.

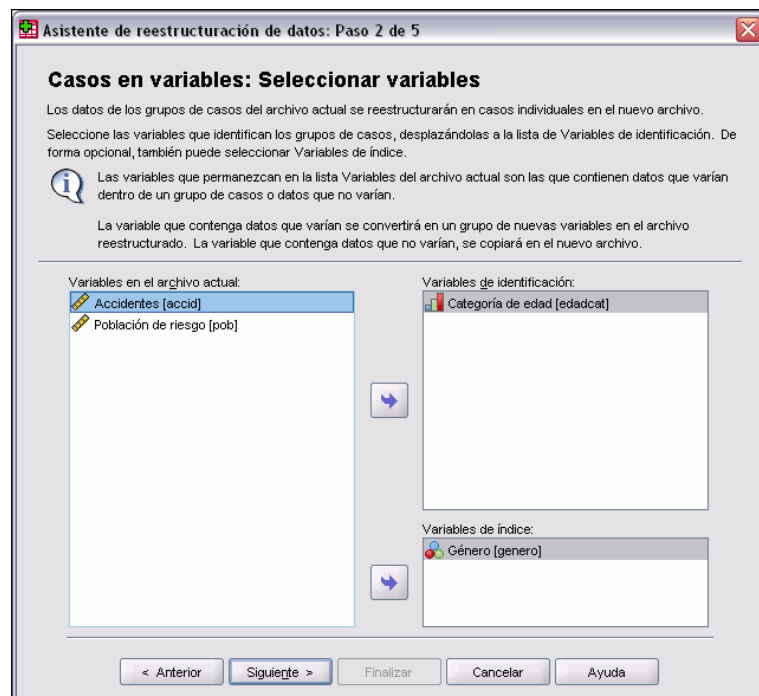
### **Asistente de reestructuración de datos (casos a variables): Seleccionar Variables**

*Nota:* el asistente presenta este paso si se ha seleccionado reestructurar grupos de casos en columnas.

En este paso, se debe proporcionar información sobre cómo se van a utilizar las variables del archivo actual en el nuevo archivo.

Figura 9-32

Asistente de reestructuración de datos: Seleccionar variables, paso 2



**¿Qué identifica los grupos de casos en los datos actuales?** Un **grupo de casos** es un grupo de filas relacionadas porque miden la misma unidad de observación, por ejemplo, un individuo o una institución. El asistente necesita conocer cuáles son las variables del archivo actual que identifican los grupos de casos para que se pueda consolidar cada grupo en una única fila del nuevo archivo. Desplace las variables que identifican grupos de casos en el archivo actual a la lista de Variables de identificación. Las variables que se utilizan para segmentar el archivo de datos actual se utilizan de forma automática para identificar los grupos de casos. Cada vez que se encuentra una nueva combinación de valores de identificación, el asistente creará una nueva fila, de manera que los casos del archivo actual deberán ordenarse en función de los valores de las variables de identificación, en el mismo orden en el que aparecen las variables en la lista Variables de identificación. Si el archivo de datos actual no está aún ordenado, podrá hacerlo en el siguiente paso.

**¿Cómo deben crearse los nuevos grupos de variables en el archivo nuevo?** En los datos originales, una variable aparece en una única columna. En el nuevo archivo de datos, dicha variable aparecerá en varias columnas. Las **variables de índice** son variables existentes en los datos actuales que el asistente deberá utilizar para crear las nuevas columnas. Los datos reestructurados contendrán una nueva variable por cada valor único contenido en dichas columnas. Desplace a la lista Variables de índice las variables que se deben utilizar para formar los nuevos grupos de variables. Cuando el asistente ofrezca opciones, también puede elegir ordenar las nuevas columnas por el índice.

**¿Qué sucede con las demás columnas?** El asistente decide de forma automática lo que hay que hacer con las variables que quedan en la lista Archivo actual. Comprueba cada variable para ver si los valores de los datos varían dentro de un grupo de casos. Si hay alguna variación, el asistente reestructurará los valores en un grupo de variables en el nuevo archivo. Si no la hay, el asistente copiará los valores en el nuevo archivo.

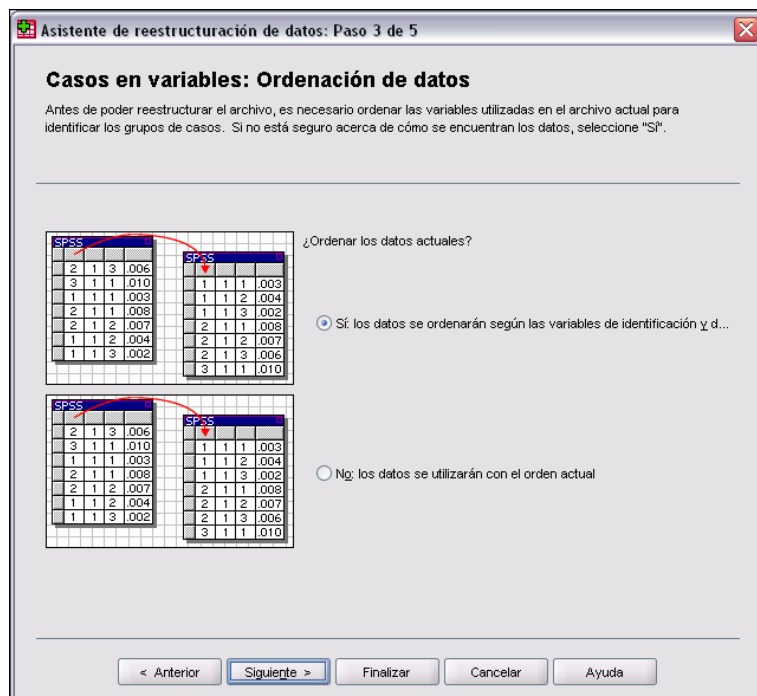
## Asistente de reestructuración de datos (casos a variables): Ordenar Datos

*Nota:* el asistente presenta este paso si se ha seleccionado reestructurar grupos de casos en columnas.

En este paso, debe decidir si se ordena el archivo actual antes de reestructurarlo. Cada vez que el asistente se encuentra una nueva combinación de valores de identificación, se crea una nueva fila, por lo tanto, es importante que los datos estén ordenados por las variables que identifican los grupos de casos.

Figura 9-33

Asistente de reestructuración de datos: Ordenación de los datos, paso 3



**¿Cómo están ordenadas las filas en el archivo actual?** Tenga en cuenta la ordenación de los datos actuales y cuáles son las variables que se están utilizando para identificar grupos de casos (especificadas en el paso anterior).

- **Sí.** El asistente ordenará de forma automática los datos actuales en función de la variable de identificación, con el mismo orden en el que aparecen las variables en la lista Variables de identificación en el paso anterior. Seleccione esta opción cuando los datos no estén ordenados en función de las variables de identificación o cuando no esté seguro. Esta opción requiere

una lectura adicional de los datos, pero garantiza que las filas estén correctamente ordenadas antes de la reestructuración.

- **No.** El asistente no ordenará los datos actuales. Seleccione esta opción cuando esté seguro de que los datos actuales están ordenados en función de las variables que identifican los grupos de casos.

### **Asistente de reestructuración de datos (casos a variables): Opciones**

*Nota:* el asistente presenta este paso si se ha seleccionado reestructurar grupos de casos en columnas.

En este paso, se deben especificar las opciones para el nuevo archivo reestructurado.

Figura 9-34  
Asistente de reestructuración de datos: Opciones, paso 4

#### **¿Cómo deben ordenarse los nuevos grupos de variables en el archivo nuevo?**

- **Por variable.** El asistente agrupa juntas las nuevas variables creadas a partir de una variable original.
- **Por índice.** El asistente agrupa las variables en función de los valores de las variables de índice.

**Ejemplo.** Las variables que se van a reestructurar son *w* y *h*, y el índice es *mes*:

w                    h                    mes

La agrupación por variable dará como resultado:

w.ene            w.feb            h.ene

La agrupación por índice dará como resultado:

w.ene            h.ene            w.feb

**¿Desea crear una variable de recuento?** El asistente puede crear una variable de recuento en el nuevo archivo. Dicha variable contendrá el número de filas de los datos actuales que se utilizaron para crear una fila en el nuevo archivo de datos.

**¿Desea crear variables indicadoras?** El asistente puede utilizar las variables de índice para crear **variables indicadoras** en el nuevo archivo de datos. Creará una nueva variable por cada valor único de la variable de índice. Las variables indicadoras indican la presencia o ausencia de un valor para un caso. Una variable indicadora toma el valor 1 si el caso tiene un valor; en caso contrario, vale 0.

**Ejemplo.** La variable de índice es *producto*. Registra los productos que ha comprado un cliente. Los datos originales son:

cliente	producto
1	pollo
1	huevos
2	huevos
3	pollo

La creación de una variable indicadora da como resultado una nueva variable para cada valor único de *producto*. Los datos reestructurados son:

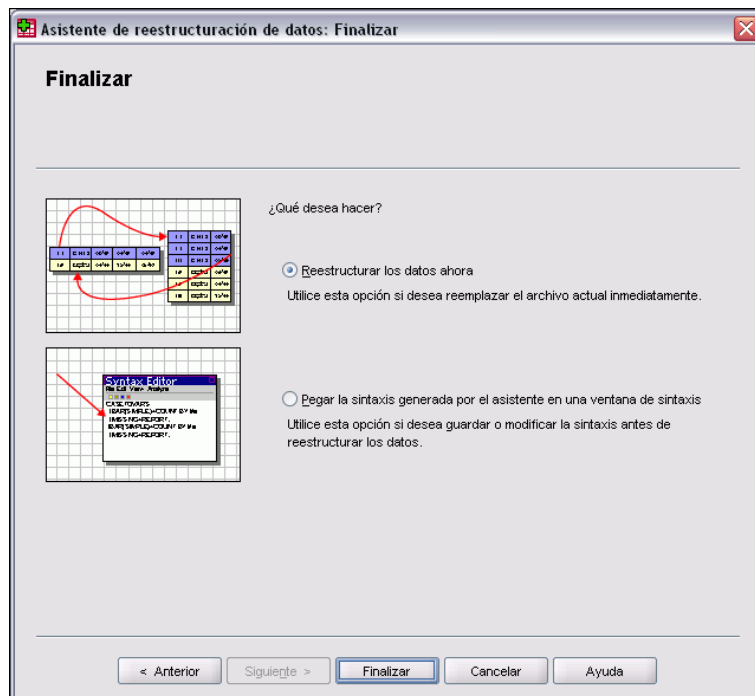
cliente	indpollo	indhuevos
1	1	1
2	0	1
3	1	0

En este ejemplo, se pueden utilizar los datos reestructurados para obtener recuentos de frecuencias de los productos que compran los clientes.

### ***Asistente de reestructuración de datos: Finalizar***

Este es el paso final del Asistente de reestructuración de datos. Debe decidir qué hacer con las especificaciones.

Figura 9-35  
Asistente de reestructuración de datos: Finalizar



- **Reestructurar los datos ahora.** El asistente creará el nuevo archivo de datos reestructurado. Seleccione esta opción si desea reemplazar el archivo actual inmediatamente.

*Nota:* si los datos originales están ponderados, los nuevos datos también lo estarán, a menos que la variable utilizada como ponderación se reestructure o se elimine del nuevo archivo.

- **Pegar la sintaxis.** El asistente pegará la sintaxis que ha generado en una ventana de sintaxis. Seleccione esta opción si no está preparado para reemplazar el archivo actual, si desea modificar la sintaxis o si desea guardarla para utilizarla en el futuro.



# Trabajo con resultados

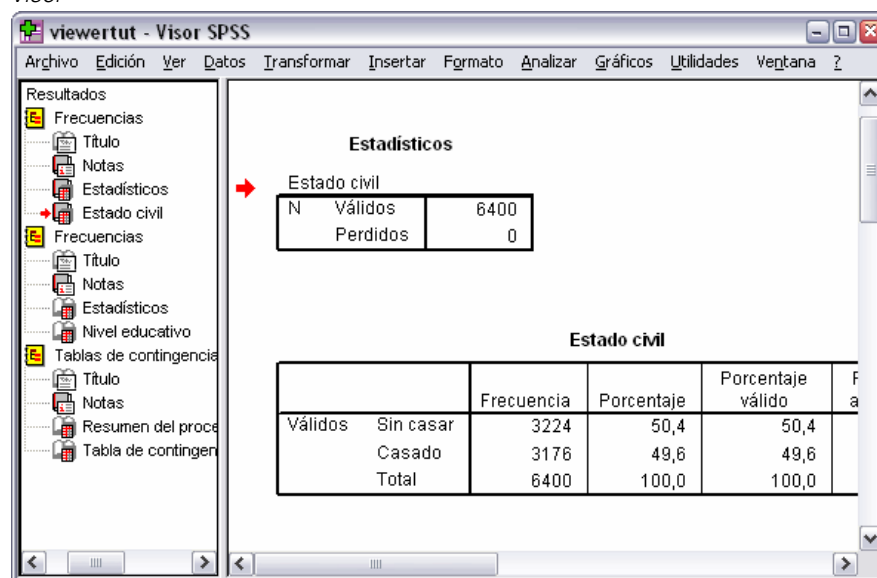
Cuando ejecute un procedimiento, los resultados se mostrarán en una ventana llamada Visor. Desde esta ventana puede desplazarse con facilidad a los resultados que desee ver. También puede modificar los resultados y crear un documento que contenga exactamente los resultados que desee.

## Visor

Los resultados se muestran en el Visor. Puede utilizar el Visor para:

- Examinar los resultados
- Mostrar u ocultar tablas y gráficos seleccionados
- Cambiar el orden de presentación de los resultados moviendo los elementos seleccionados
- Mover elementos entre el Visor y otras aplicaciones

Figura 10-1  
Visor



El Visor se divide en dos paneles:

- El panel izquierdo contiene una vista de titulares de los contenidos.
- El panel derecho contiene tablas estadísticas, gráficos y resultados de texto.

Puede pulsar en un elemento de los titulares para dirigirse directamente a la tabla o al gráfico correspondiente. Puede pulsar y arrastrar el borde derecho del panel de titulares para cambiar la anchura del mismo.

### **Mostrar y ocultar resultados**

En el Visor, puede mostrar y ocultar de forma selectiva las tablas o los resultados individuales de todo un procedimiento. Este proceso resulta de utilidad cuando desea reducir la cantidad de resultados visibles en el panel de contenido.

#### **Para ocultar tablas y gráficos**

- ▶ En el panel de titulares del Visor, pulse dos veces en el icono de libro del elemento.  
*o*
- ▶ Pulse en el elemento para seleccionarlo.
- ▶ Elija en los menús:  
Ver  
Ocult.  
*o*
- ▶ Pulse en el icono de libro cerrado (Ocultar) de la barra de herramientas de titulares.  
El icono de libro abierto (Mostrar) se convierte en el icono activo, indicando que el elemento no está oculto.

#### **Para ocultar los resultados de un procedimiento**

- ▶ Pulse en el cuadro situado a la izquierda del nombre del procedimiento en el panel de titulares.  
Se ocultarán todos los resultados del procedimiento y se contraerá la presentación de titulares.

### **Desplazamiento, eliminación y copia de resultados**

Los resultados se pueden reorganizar copiando, moviendo o eliminando un elemento o un grupo de elementos.

#### **Para desplazar resultados en el Visor**

- ▶ Seleccione los elementos en el panel de titulares o de contenido.
- ▶ Arrastre y coloque los elementos seleccionados en una ubicación diferente.

#### **Para eliminar resultados en el Visor**

- ▶ Seleccione los elementos en el panel de titulares o de contenido.

- ▶ Pulse la tecla Supr.

*o*

- ▶ Elija en los menús:  
Edición  
Borrar

### ***Cambio de la alineación inicial***

Por defecto, todos los resultados están alineados inicialmente a la izquierda. Para cambiar la alineación inicial de los nuevos elementos de los resultados:

- ▶ Elija en los menús:  
Edición  
Opciones
- ▶ Pulse en la pestaña Visor.
- ▶ En el grupo Estado inicial de los resultados, seleccione el tipo de elemento (por ejemplo, tabla pivote, gráfico o resultados de texto).
- ▶ Seleccione la opción de alineación que desee.

### ***Cambio de la alineación de los elementos de resultados***

- ▶ En el panel de titulares o de contenido, seleccione los elementos que desea alinear.
- ▶ Elija en los menús:  
Formato  
Alineación izquierda

*o*

Formato  
Centrado

*o*

Formato  
Alinear a la derecha

### ***Resultados del Visor***

El panel de titulares proporciona una tabla de contenido del documento del Visor. Utilice este panel para navegar por los resultados y controlar su presentación. La mayoría de las acciones en dicho panel tienen su efecto correspondiente en el panel de contenido.

- Si se selecciona un elemento en el panel de titulares, también se mostrará el elemento correspondiente en el panel de contenido.

- Si se mueve un elemento en el panel de titulares, también se moverá el elemento correspondiente en el panel de contenido.
- Si se contrae la vista de titulares, se ocultarán los resultados de todos los elementos en los niveles contraídos.

**Control de la presentación de los titulares.** Para controlar la presentación de titulares, puede:

- Expandir y contraer la presentación de titulares
- Cambiar el nivel de los titulares para los elementos seleccionados
- Cambiar el tamaño de los elementos en la presentación de titulares
- Cambiar la fuente utilizada en la presentación de titulares

### ***Para expandir y contraer la vista de titulares***

- ▶ Pulse en el cuadro situado a la izquierda del elemento de los titulares que desee contraer o expandir.

*o*

- ▶ Pulse en el elemento de los titulares.

- ▶ Elija en los menús:

Ver  
Contraer

*o*

Ver  
Expandir

### ***Para cambiar el nivel de titulares***

- ▶ Pulse en el elemento del panel de titulares.
- ▶ Pulse en la flecha izquierda de la barra de herramientas de titulares para ascender el elemento (mueva el elemento hacia la izquierda).

*o*

Pulse en la flecha derecha de la barra de herramientas de titulares para degradar el elemento (mueva el elemento hacia la derecha).

*o*

- ▶ Elija en los menús:

Edición  
Resaltado  
Ascender

*o*

Edición  
Resaltado  
Descender

El cambio del nivel de titulares es particularmente útil después de mover elementos en el nivel de titulares. El desplazamiento de elementos puede cambiar el nivel de titulares de los elementos y puede utilizar los botones de flecha izquierda y derecha de la barra de herramientas de los titulares para restaurar el nivel de titulares original.

### ***Para cambiar el tamaño de elementos de los titulares***

- ▶ Elija en los menús:
  - Ver
  - Tamaño de los titulares
- ▶ Seleccione el tamaño de los titulares (Pequeño, Mediano o Grande).

### ***Para cambiar la fuente de los titulares***

- ▶ Elija en los menús:
  - Ver
  - Fuente de los titulares
- ▶ Seleccione una fuente.

## ***Adición de elementos al Visor***

En el Visor puede añadir elementos tales como títulos, nuevo texto, gráficos o material de otras aplicaciones.

### ***Para añadir un título o texto***

Pueden añadirse al Visor elementos de texto que no estén conectados a una tabla o a un gráfico.

- ▶ Pulse en la tabla, en el gráfico o en el otro objeto que precederá al título o al texto.
- ▶ Elija en los menús:
  - Insertar
  - Nuevo título
- o*
  - Insertar
  - Nuevo texto
- ▶ Pulse dos veces en el nuevo objeto.
- ▶ Escriba el texto.

### ***Para añadir un archivo de texto***

- ▶ En el panel de titulares o en el panel de contenido del Visor, pulse en la tabla, en el gráfico o en otro objeto que vaya a preceder al texto.

- ▶ Elija en los menús:
  - Insertar
  - Archivo de texto...
- ▶ Seleccione un archivo de texto.
  - Para editar el texto, pulse en él dos veces.

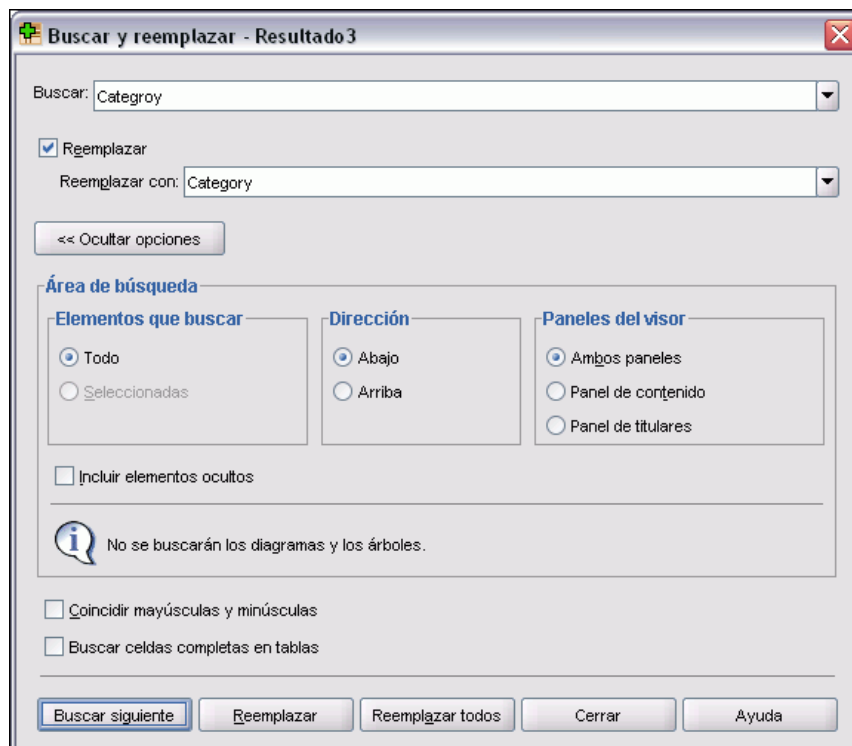
### ***Pegado de objetos en el Visor***

Es posible pegar objetos de otras aplicaciones en el Visor. Puede utilizar Pegar debajo o Pegado especial. Cualquiera de estos tipos de pegado coloca el nuevo objeto después del objeto actualmente seleccionado en el Visor. Utilice Pegado especial cuando desee seleccionar el formato del objeto pegado.

### ***Búsqueda y sustitución de información en el Visor***

- ▶ Para buscar o reemplazar información en el Visor, elija en los menús:
    - Edición
    - Buscar
- o*
- Edición
  - Reemplazar

Figura 10-2  
Cuadro de diálogo *Buscar y reemplazar*



Puede utilizar la función Buscar y reemplazar para:

- Buscar en todo un documento o únicamente en los elementos seleccionados.
- Buscar hacia abajo o hacia arriba, desde la ubicación actual.
- Buscar en ambos paneles o restringir la búsqueda al panel de contenido o de titulares.
- Buscar elementos ocultos, incluidos los elementos ocultos del panel de contenido (por ejemplo, las tablas Notas, que están ocultas por defecto) y las filas y columnas ocultas de las tablas pivote.
- Restringir los criterios de búsqueda a coincidencias que distingan entre mayúsculas y minúsculas.
- Restringir los criterios de búsqueda de las tablas pivote a coincidencias del contenido completo de las casillas.

#### ***Elementos ocultos y capas de la tabla pivote***

- Las capas situadas por debajo de la capa visible actual de una tabla pivote multidimensional no se consideran ocultas y se incluirán en el área de búsqueda aunque no incluyan los elementos ocultos en la búsqueda.
- Los elementos ocultos incluyen elementos ocultos del panel de contenido (elementos con iconos de libro cerrado en el panel de titulares o incluidos dentro de bloques plegados del panel) y filas y columnas de las tablas pivote ocultas por defecto (por ejemplo, las filas y columnas vacías) o manualmente mediante la edición de la tabla y posterior selección de que se desea ocultar determinadas filas y columnas. Los elementos ocultos sólo se incluirán en la búsqueda si selecciona de forma explícita Incluir elementos ocultos.
- En ambos casos, se mostrarán los elementos ocultos o no visibles que contengan el texto o valor de búsqueda, pero a continuación volverán a su estado original.

## ***Copia de resultados a otras aplicaciones***

Los objetos de resultados pueden copiarse y pegarse en otras aplicaciones, como puede ser un procesador de textos o una hoja de cálculo. Puede pegar los resultados de diversas formas. Según la aplicación de destino, pueden estar disponibles todos o algunos de los siguientes formatos:

**Imagen (metarchivo).** Puede pegar las tablas pivote, los resultados de texto y los gráficos como imágenes de metarchivo. El formato de la imagen puede cambiarse de tamaño en la otra aplicación y a veces determinados aspectos de la edición se pueden realizar con las características que proporcione la otra aplicación. Las tablas pivote pegadas como imágenes conservan todas las características de bordes y fuente. Este formato sólo está disponible en los sistemas operativos Windows.

**RTF (formato de texto enriquecido).** Las tablas pivote pueden pegarse en otras aplicaciones con formato RTF. En la mayoría de las aplicaciones, esto significa que la tabla pivote se pega como una tabla que puede editarse posteriormente en la otra aplicación.

*Nota:* es posible que Microsoft Word no muestre correctamente las tablas muy anchas.

**Mapa de bits.** Las tablas pivote y los gráficos se pueden pegar en otras aplicaciones como mapas de bits.

**BIFF.** El contenido de una tabla se puede pegar en una hoja de cálculo y conservar la precisión numérica.

**Texto.** El contenido de una tabla se puede copiar y pegar como texto. Este proceso puede ser útil en aplicaciones como el correo electrónico, donde sólo se puede aceptar o transmitir texto.

Si la aplicación de destino admite varios de los formatos disponibles, es posible que tenga un elemento de menú Pegado especial que le permita seleccionar el formato o que muestre de forma automática una lista de los formatos disponibles.

### ***Para copiar y pegar elementos de resultados en otra aplicación***

- ▶ Seleccione los objetos en el panel de titulares o de contenido del Visor.
- ▶ Seleccione en los menús del Visor:
  - Edición
  - Copiar
- ▶ Elija en los menús de la aplicación de destino:
  - Edición
  - Pegar
- o*
  - Edición
  - Pegado especial...

**Pegar.** Los resultados se copian al Portapapeles en varios formatos. Cada aplicación determinará el mejor formato para la operación de pegado.

**Pegado especial.** Los resultados se pegan al Portapapeles en varios formatos. Pegado especial le permite seleccionar el formato que desea de la lista de formatos disponibles en la aplicación de destino.

### ***Exportar resultados***

La opción Exportar resultados guarda los resultados del Visor en formato HTML, texto, Word/RTF, Excel, PowerPoint (requiere PowerPoint 97 o posterior) y PDF. Los gráficos se pueden exportar también en varios formatos de gráficos distintos.

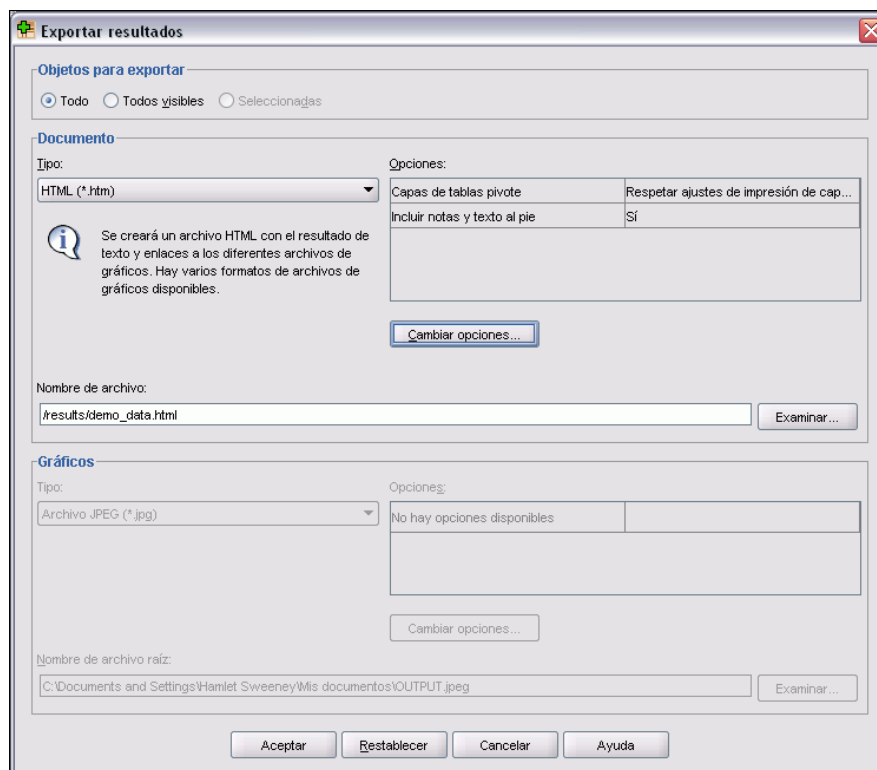
*Nota:* la exportación a PowerPoint sólo está disponible en los sistemas operativos Windows y no está disponible con la versión para estudiantes.

#### ***Para exportar resultados***

- ▶ Active la ventana del Visor (pulse en cualquier punto de la ventana).
- ▶ Elija en los menús:
  - Archivo
  - Exportar...
- ▶ Introduzca un nombre de archivo (o prefijo para los gráficos) y seleccione un formato de exportación.



Figura 10-3  
Cuadro de diálogo *Exportar resultados*



**Objetos para exportar.** Permite exportar todos los objetos del Visor, todos los objetos visibles o sólo los objetos seleccionados.

**Tipo de documento.** Las opciones disponibles son:

- **Word/RTF (\*.doc).** Las tablas pivote se exportan como tablas de Word, con todos los atributos de formato intactos (por ejemplo, bordes de casilla, estilos de fuente y colores de fondo). Los resultados de texto se exportan en formato RTF. En los sistemas operativos Windows, los gráficos se incluyen en el documento en formato EMF (metarchivo mejorado). En los demás sistemas operativos, los gráficos se incluyen en formato PNG. Los resultados de texto siempre se muestran con un tipo de letra de paso fijo (monoespaciadas) y se exportan con los mismos atributos de fuente. Es necesario utilizar una fuente de paso fijo para que los resultados de texto separado por espacios queden correctamente alineados; por defecto, la mayoría de los navegadores de Web utilizan una fuente de paso fijo para el texto con formato previo.

*Nota:* es posible que Microsoft Word no muestre correctamente las tablas muy anchas.

- **Excel (\*.xls).** Las filas, columnas y casillas de tablas pivote se exportan como filas, columnas y casillas de Excel, con todos los atributos de formato intactos, por ejemplo, bordes de casilla, estilos de fuente y colores de fondo, etc. Los resultados de texto se exportan con todos los atributos de fuente intactos. Cada línea del resultado de texto constituye una fila del archivo de Excel y se incluye todo su contenido en una sola casilla. No se incluyen los gráficos.

- **HTML (\*.htm).** Las tablas pivote se exportan como tablas HTML. Los resultados de texto se exportan como formato previo de HTML. Los gráficos se incrustan mediante referencias y deben exportarse en un formato adecuado para incluirlos en los documentos HTML (por ejemplo, PNG y JPEG).
- **Formato de documento portátil (\*.pdf).** Todos los resultados se exportan como aparecen en la presentación preliminar, con todos los atributos de formato intactos.
- **Archivo de PowerPoint (\*.ppt).** Las tablas pivote se exportan como tablas de Word y se incrustan en diapositivas diferentes en el archivo de PowerPoint (una diapositiva por cada tabla pivote). Se conservan todos los atributos de formato de la tabla pivote, por ejemplo, bordes de casillas, estilos de fuente y colores de fondo. Los gráficos se exportan en formato TIFF. No se incluyen los resultados de texto.

*Nota:* la exportación a PowerPoint sólo está disponible en los sistemas operativos Windows y no está disponible con la versión para estudiantes.

- **Texto (\*.txt).** Entre los formatos de resultados de texto se incluyen texto sin formato, UTF-8, and UTF-16. Las tablas pivote se pueden exportar en formato separado por tabuladores o por espacios. Todos los resultados de texto se exportan en formato separado por espacios. Para los gráficos, se inserta en el archivo de texto una línea por cada gráfico, indicando el nombre de archivo del gráfico exportado.
- **Ninguno (sólo gráficos).** Los formatos de exportación disponibles incluyen: EPS, JPEG, TIFF, PNG y BMP. En los sistemas operativos Windows, también está disponible el formato EMF (metarchivo mejorado).

**Sistema de gestión de resultados.** Puede exportar automáticamente además todos los resultados o tipos de resultados especificados por el usuario como texto, HTML, XML y archivos de datos en formato SPSS. Si desea obtener más información, consulte [Sistema de gestión de resultados](#) en Capítulo 48 en p. 511.

## ***Opciones de HTML, Word/RTF y Excel***

Las siguientes opciones están disponibles para la exportación de tablas pivote a Word/RTF, Excel y HTML:

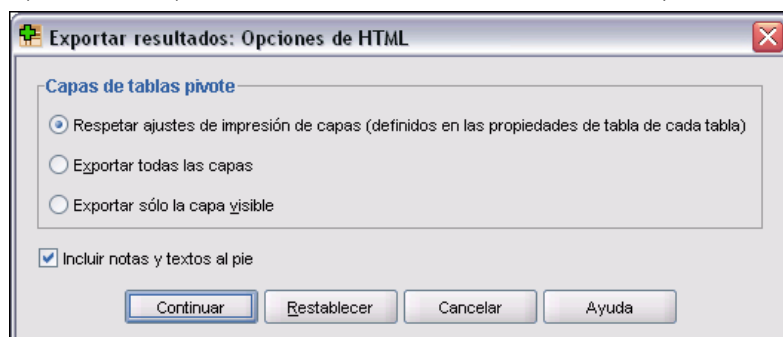
- **Capas en tablas pivote.** Por defecto, la inclusión o exclusión de las capas de una tabla pivote está controlada por las propiedades de la tabla de cada tabla pivote. Puede anular este ajuste e incluir todas las capas o excluir todas excepto la capa visible en ese momento. Si desea obtener más información, consulte [Propiedades de tabla: Imprimiendo](#) en Capítulo 11 en p. 251.
- **Incluir notas y textos al pie.** Controla la inclusión o exclusión de todas las notas y textos al pie de la tabla pivote.

*Nota:* para HTML, también es posible controlar el formato de archivo de imagen de los gráficos exportados. Si desea obtener más información, consulte [Opciones para la exportación de gráficos](#) en p. 230.

### ***Para establecer las opciones de exportación de HTML, Word/RTF y Excel***

- ▶ Seleccione HTML, Word/RTF o Excel como formato de exportación.
- ▶ Pulse en Cambiar opciones.

Figura 10-4  
Opciones de exportación de resultados en HTML, Word/RTF y Excel



## Opciones de PowerPoint

Las siguientes opciones están disponibles para PowerPoint:

- **Capas en tablas pivote.** Por defecto, la inclusión o exclusión de las capas de una tabla pivote está controlada por las propiedades de la tabla de cada tabla pivote. Puede anular este ajuste e incluir todas las capas o excluir todas excepto la capa visible en ese momento. Si desea obtener más información, consulte [Propiedades de tabla: Imprimiendo](#) en Capítulo 11 en p. 251.
- **Incluir notas y textos al pie.** Controla la inclusión o exclusión de todas las notas y textos al pie de la tabla pivote.
- **Usar entradas de titulares del Visor como títulos de diapositivas.** Incluye un título en cada diapositiva creada en la exportación. Cada diapositiva contiene un único elemento exportado del Visor. El título se genera a partir de la entrada del titular para el elemento en el panel de titulares del Visor.

### Para configurar las opciones de exportación de PowerPoint

- ▶ Seleccione PowerPoint como formato de exportación.
- ▶ Pulse en Cambiar opciones.

*Nota:* la exportación a PowerPoint sólo está disponible en los sistemas operativos Windows.

## Opciones de PDF

Las siguientes opciones están disponibles para PDF:

**Incrustar marcadores.** Esta opción incluye en el documento PDF los marcadores correspondientes a las entradas de titulares del Visor. Al igual que el panel de titulares del Visor, los marcadores pueden facilitar mucho la navegación por los documentos que tienen un gran número de objetos de resultados.

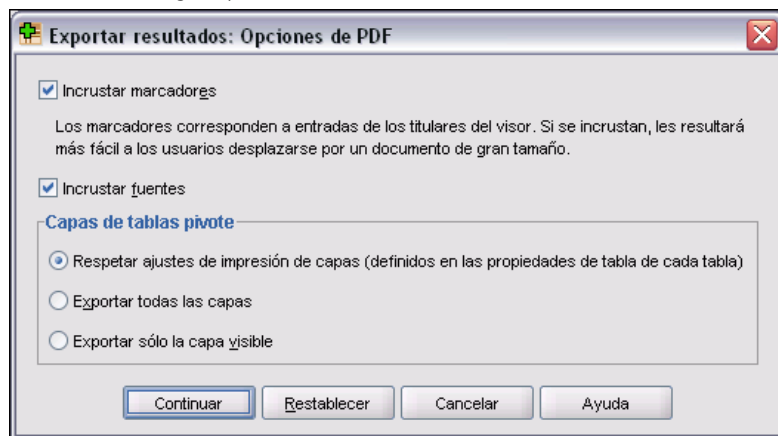
**Incrustar fuentes.** Al incrustar fuentes se garantiza que el documento PDF presentará el mismo aspecto en cualquier ordenador. De lo contrario, si algunas fuentes utilizadas en el documento no están disponibles en el ordenador donde se visualiza o imprime el documento PDF, la sustitución de fuentes puede resultar en una calidad menor.

**Capas en tablas pivote.** Por defecto, la inclusión o exclusión de las capas de una tabla pivote está controlada por las propiedades de la tabla de cada tabla pivote. Puede anular este ajuste e incluir todas las capas o excluir todas excepto la capa visible en ese momento. Si desea obtener más información, consulte [Propiedades de tabla: Imprimiendo](#) en Capítulo 11 en p. 251.

### **Para configurar las opciones de exportación de PDF**

- ▶ Seleccione Formato de documento portátil como formato de exportación.
- ▶ Pulse en Cambiar opciones.

Figura 10-5  
Cuadro de diálogo Opciones de PDF



### **Otros ajustes que afectan al resultado del PDF**

**Configuración de página / Atributos de página.** El tamaño de la página, la orientación, los márgenes, el contenido y la presentación de los encabezados y pies de página, así como el tamaño del gráfico impreso en los documentos PDF están controlados por las opciones de configuración de página y las opciones de atributos de página.

**Propiedades de tabla/Aspectos de tabla.** El escalamiento de tablas largas o anchas y la impresión de capas de tablas se controla mediante las propiedades de tabla. Estas propiedades también se pueden guardar en Aspectos de tabla. Si desea obtener más información, consulte [Propiedades de tabla: Imprimiendo](#) en Capítulo 11 en p. 251.

**Impresora por defecto/actual.** La resolución (PPP) del documento PDF es la configuración de la resolución actual para la impresora por defecto o que esté seleccionada en ese momento (y que puede cambiarse mediante Preparar página). La resolución máxima es de 1200 PPP. Si el valor de la impresora es superior, la resolución del documento PDF será de 1200 PPP.

*Nota:* los documentos de alta resolución pueden generar resultados de baja calidad si se imprimen en impresoras con menor resolución.

## Opciones del texto

Las siguientes opciones están disponibles para la exportación de texto:

**Formato de tablas pivote.** Las tablas pivote se pueden exportar en formato separado por tabuladores o por espacios. Para el formato separado por espacios, también puede controlar:

- **Ancho de columna.** Autoajuste no muestra nunca el contenido de las columnas en varias líneas. Cada columna tiene el ancho correspondiente al valor o etiqueta más ancho que haya en dicha columna. Personalizado establece un ancho de columna máximo que se aplica a todas las columnas de la tabla. Los valores que superen dicho ancho se mostrarán en otra línea en la misma columna.
- **Carácter de borde de fila/columna.** Controla los caracteres utilizados para crear los bordes de fila y columna. Para suprimir la visualización de los bordes de fila y columna, especifique espacios en blanco para estos valores.

**Capas en tablas pivote.** Por defecto, la inclusión o exclusión de las capas de una tabla pivote está controlada por las propiedades de la tabla de cada tabla pivote. Puede anular este ajuste e incluir todas las capas o excluir todas excepto la capa visible en ese momento. Si desea obtener más información, consulte [Propiedades de tabla: Imprimiendo](#) en Capítulo 11 en p. 251.

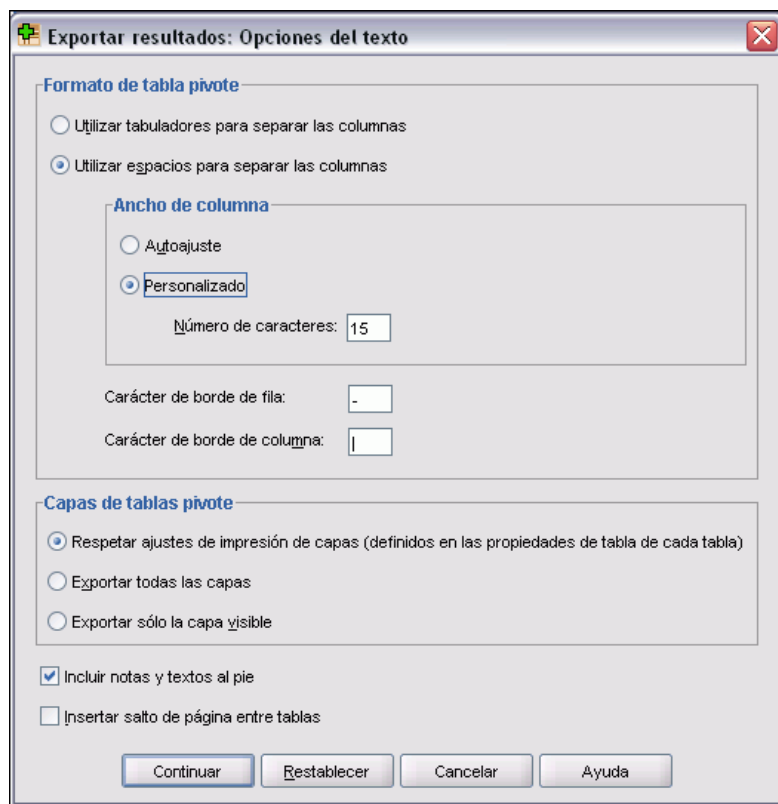
**Incluir notas y textos al pie.** Controla la inclusión o exclusión de todas las notas y textos al pie de la tabla pivote.

**Insertar salto de página entre tablas.** Inserta un avance de página o salto de página entre cada tabla. En el caso de las tablas pivote de varias capas, esta opción inserta un salto de página entre cada capa.

### **Para configurar las opciones de exportación de texto**

- ▶ Seleccione Texto como formato de exportación.
- ▶ Pulse en Cambiar opciones.

Figura 10-6  
Cuadro de diálogo *Opciones del texto*



## ***Opciones para la exportación de gráficos***

Para los documentos de texto o HTML y para exportar sólo gráficos, puede seleccionar el formato gráfico y, para cada formato gráfico, controlar varias opciones.

Para seleccionar el formato gráfico y las opciones de los gráficos exportados:

- ▶ Seleccione HTML, Texto o Ninguno (sólo gráficos) como tipo de documento.
- ▶ Seleccione el formato de archivo gráfico en la lista desplegable.
- ▶ Pulse en Cambiar opciones para cambiar las opciones para el formato de archivo gráfico seleccionado.

### ***Opciones de exportación de gráficos JPEG***

- **Tamaño de imagen.** Porcentaje del tamaño de gráfico original (hasta 200%).
- **Convertir a escala de grises.** Convierte los colores en tonos de gris.

### **Opciones de exportación de gráficos BMP**

- **Tamaño de imagen.** Porcentaje del tamaño de gráfico original (hasta 200%).
- **Comprimir imagen para reducir el tamaño de archivo.** Técnica de compresión sin pérdida que crea archivos más pequeños sin afectar a la calidad de la imagen.

### **Opciones de exportación de gráficos PNG**

**Tamaño de imagen.** Porcentaje del tamaño de gráfico original (hasta 200%).

**Profundidad de color.** Determina el número de colores del gráfico exportado. Un gráfico que se guarda con cualquier profundidad tendrá un mínimo del número de colores que se utilizan y un máximo del número de colores permitidos por la profundidad. Por ejemplo, si el gráfico incluye tres colores (rojo, blanco y negro) y se guarda como un gráfico de 16 colores, permanecerá como gráfico de tres colores.

- Si el número de colores del gráfico es superior al número de colores para dicha profundidad, los colores se interpolarán para reproducir los colores del gráfico.
- Profundidad de la pantalla actual es el número de colores mostrados actualmente en el monitor del ordenador.

### **Opciones de exportación de gráficos EMF y TIFF**

**Tamaño de imagen.** Porcentaje del tamaño de gráfico original (hasta 200%).

*Nota:* el formato EMF (metarchivo mejorado) sólo está disponible en los sistemas operativos Windows.

### **Opciones de exportación de gráficos EPS**

**Tamaño de imagen.** Puede especificar el tamaño como porcentaje del tamaño de imagen original (hasta 200%) o especificar un ancho de imagen en píxeles (el alto se determina en función del valor del ancho y la relación de aspecto). La imagen exportada siempre es proporcional a la original.

**Incluir vista preliminar TIFF.** Guarda una vista preliminar con la imagen EPS en formato TIFF para su visualización en aplicaciones que no pueden mostrar imágenes EPS en la pantalla.

**Fuentes.** Controla el tratamiento de fuentes en las imágenes EPS.

- **Utilizar referencias de fuentes.** Si las fuentes que se utilizan en el gráfico están disponibles en el dispositivo de resultados, se hará uso de ellas. En caso contrario, el dispositivo de resultados utiliza fuentes alternativas.
- **Sustituir fuentes con curvas.** Convierte las fuentes en datos de curvas PostScript. El texto ya no se puede editar como texto en las aplicaciones que pueden editar los gráficos EPS. Esta opción es útil si las fuentes que se utilizan en el gráfico no están disponibles en el dispositivo de los resultados.

## ***Impresión de documentos del Visor***

Hay dos opciones para imprimir el contenido de la ventana del Visor:

**Todos los resultados visibles.** Se imprimen sólo los elementos que se muestran actualmente en el panel de contenido. No se imprimen los elementos ocultos (los elementos con un icono de libro cerrado en el panel de titulares o los ocultos en las capas de titulares contraídas).

**Selección.** Se imprimen sólo los elementos que están seleccionados actualmente en los paneles de titulares y de contenido.

### ***Para imprimir resultados y gráficos***

- ▶ Active la ventana del Visor (pulse en cualquier punto de la ventana).
- ▶ Elija en los menús:
  - Archivo
  - Imprimir...
- ▶ Seleccione los ajustes de impresión que desee.
- ▶ Pulse en Aceptar para imprimir.

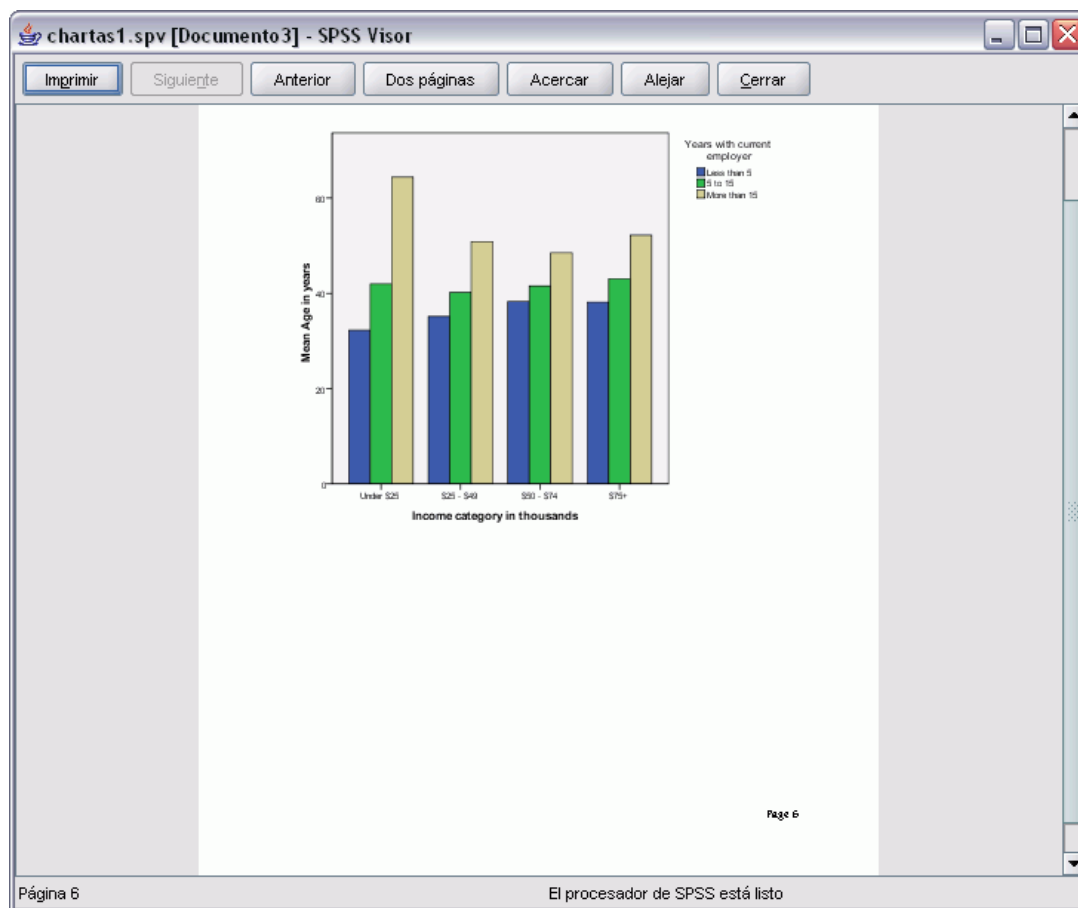
### ***Presentación preliminar***

Presentación preliminar muestra lo que se imprimirá en cada página de los documentos del Visor. Es una buena idea comprobar la presentación preliminar antes de imprimir un documento del Visor, ya que muestra elementos que quizá no puedan verse en el panel de contenido; entre ellos:

- Los saltos de página
- Las capas ocultas de las tablas pivote
- Las rupturas de las tablas anchas
- Los encabezados y pies que están impresos en cada página



Figura 10-7  
Presentación preliminar



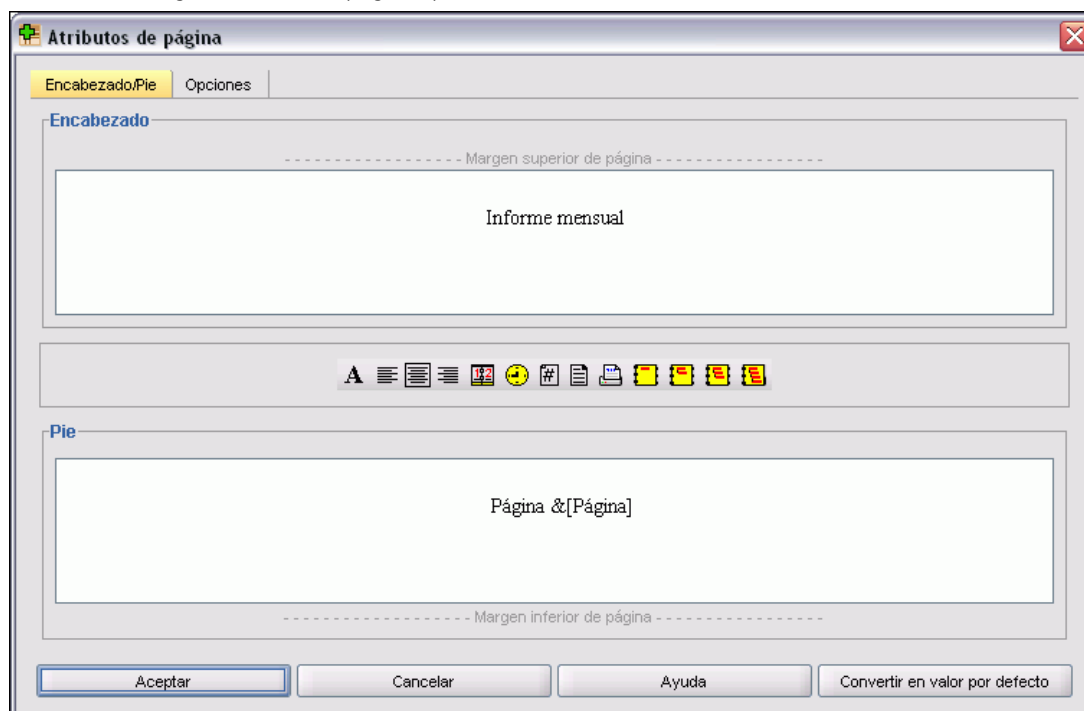
Si se han seleccionado resultados en el Visor, la presentación preliminar sólo mostrará estos resultados. Si desea ver una presentación preliminar de todos los resultados, asegúrese de que no haya nada seleccionado en el Visor.

### ***Atributos de página: Encabezados y pies***

Los encabezados y los pies constituyen la información que está impresa en la parte superior e inferior de cada página. Puede introducir cualquier texto que desee utilizar como encabezados y pies. También puede utilizar la barra de herramientas, situada en medio del cuadro de diálogo, para insertar:

- La fecha y la hora
- Los números de páginas
- El nombre del archivo del Visor
- Las etiquetas de los encabezados de los titulares
- Los títulos y subtítulos de página.

Figura 10-8  
Cuadro de diálogo Atributos de página: pestaña Encabezado/Pie



- Convertir en valor por defecto emplea la configuración especificada como configuración por defecto para los nuevos documentos del Visor.
- Las etiquetas de las cabeceras de los titulares indican el primer, el segundo, el tercer y/o el cuarto nivel de cabecera del titular para el primer elemento en cada página.
- Los títulos y los subtítulos de página imprimen los títulos y subtítulos de página actuales. Se pueden crear con la opción Nuevo título de página del menú Insertar del Visor o con los comandos TITLE y SUBTITLE. Si no ha especificado ningún título ni subtítulo de página, este ajuste no se tendrá en cuenta.

*Nota:* las características de las fuentes de los nuevos títulos y subtítulos de página se controlan en la pestaña Visor del cuadro de diálogo Opciones (a la que se accede al seleccionar Opciones en el menú Edición). También se pueden cambiar las características de los títulos y subtítulos de página existentes editándolos en el Visor.

Para ver cómo aparecerán los encabezados y pies en la página impresa, seleccione Presentación preliminar en el menú Archivo.

### **Para insertar encabezados y pies de página**

- ▶ Active la ventana del Visor (pulse en cualquier punto de la ventana).
- ▶ Elija en los menús:
  - Archivo
  - Atributos de página...

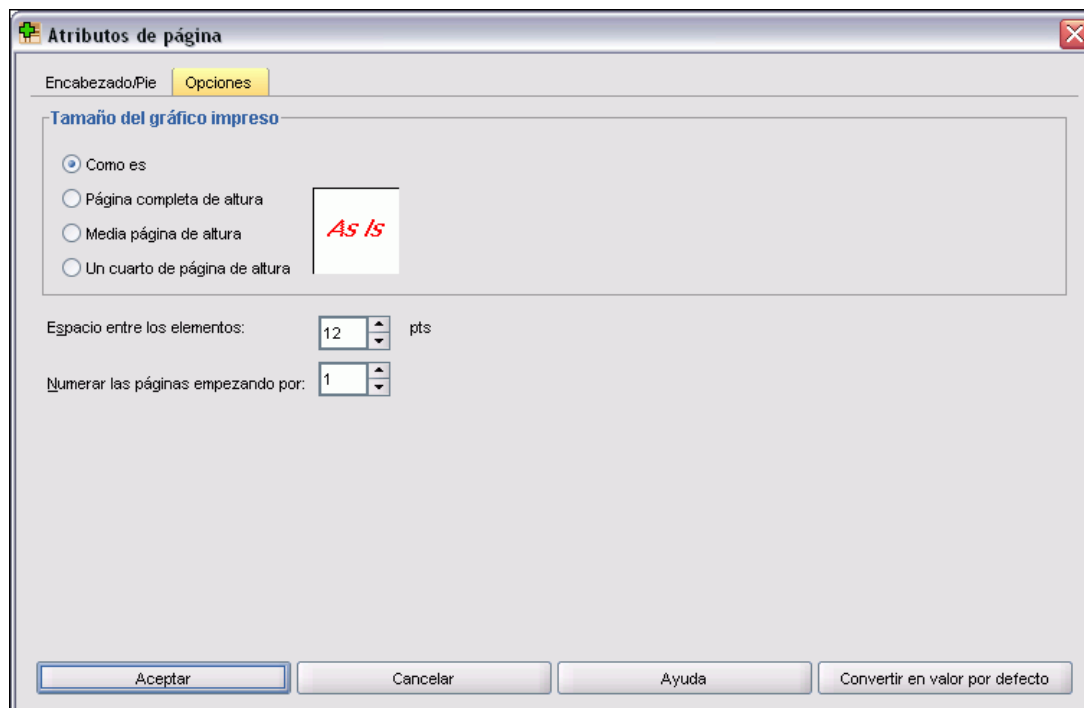
- ▶ Pulse en la pestaña Encabezado/Pie.
- ▶ Introduzca el encabezado y/o el pie que desee que aparezca en cada página.

### **Atributos de página: Opciones**

Este cuadro de diálogo controla el tamaño de los gráficos impresos, el espacio entre los elementos de resultado impresos y la numeración de las páginas.

- **Tamaño del gráfico impreso.** Controla el tamaño del gráfico impreso relativo al tamaño de la página definido. La relación de aspecto de los gráficos (proporción anchura-altura) no se ve afectada por el tamaño del gráfico impreso. El tamaño global impreso de un gráfico está limitado tanto por su altura como por su anchura. Cuando los bordes externos de un gráfico alcancen los bordes izquierdo y derecho de la página, el tamaño del gráfico no se podrá aumentar más para completar la altura de página adicional.
- **Espacio entre los elementos.** Controla el espacio entre los elementos impresos. Cada tabla pivote, gráfico y objeto de texto es un elemento diferente. Este ajuste no afecta a la presentación de los elementos en el Visor.
- **Numerar las páginas empezando por.** Numera las páginas secuencialmente, empezando por el número especificado.
- **Convertir en valor por defecto.** Esta opción emplea la configuración especificada como configuración por defecto para los nuevos documentos del Visor.

Figura 10-9  
Cuadro de diálogo Atributos de página, pestaña Opciones



***Para cambiar el tamaño del gráfico impreso, numeración de páginas y espacio entre elementos impresos***

- ▶ Active la ventana del Visor (pulse en cualquier punto de la ventana).
- ▶ Elija en los menús:
  - Archivo
  - Atributos de página...
- ▶ Pulse en la pestaña Opciones.
- ▶ Cambie los ajustes y pulse en Aceptar.

***Almacenamiento de resultados***

El contenido del Visor puede guardarse en un documento del Visor. El documento incluirá los dos paneles de la ventana del Visor (titulares y contenidos).

***Para guardar un documento del Visor***

- ▶ Seleccione en los menús de la ventana del Visor:
  - Archivo
  - Guardar
- ▶ Escriba el nombre del documento y pulse en Guardar.

Para guardar los resultados en otros formatos externos (por ejemplo, HTML o texto), utilice el comando Exportar del menú Archivo.

# Tablas pivote

Muchos de los resultados se presentan en tablas que se pueden pivotar de forma interactiva. Es decir, puede reorganizar las filas, columnas y capas.

## Gestión de una tabla pivote

Entre las opciones para manipular una tabla pivote se incluyen:

- Transposición de filas y columnas
- Desplazamiento de filas y columnas
- Creación de capas multidimensionales
- Agrupación y separación de filas y columnas
- Visualización y ocultación de filas, columnas y otra información
- Rotación de etiquetas de fila y de columna
- Búsqueda de definiciones de términos

## Activación de una tabla pivote

Antes de manipular o modificar una tabla pivote es necesario **activar** la tabla. Para activar una tabla:

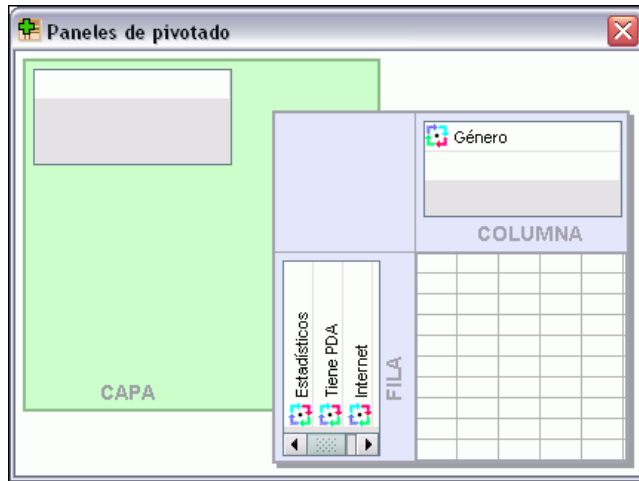
- ▶ Pulse dos veces en la tabla.
- o*
- ▶ Pulse con el botón derecho del ratón en la tabla y seleccione Editar el contenido.
  - ▶ En el submenú, elija En el visor o En la ventana independiente.
    - Por defecto, al activar la tabla pulsando dos veces en ella, se activan todas las tablas en la ventana del Visor, excepto las que son muy grandes. Si desea obtener más información, consulte [Tabla pivote: Opciones](#) en Capítulo 45 en p. 490.
    - Si desea tener más de una tabla pivote activada al mismo tiempo, deberá activarlas en ventanas diferentes.

## Pivotado de una tabla

- ▶ Active la tabla pivote.

- ▶ Elija en los menús:
  - Pivotar
  - Paneles de pivotado

Figura 11-1  
Paneles de pivotado



Una tabla tiene tres dimensiones: filas, columnas y capas. Una dimensión puede contener varios elementos (o ninguno). Puede cambiar la organización de la tabla moviendo los elementos entre las dimensiones o dentro de éstas. Para mover un elemento sólo tiene que arrastrarlo y colocarlo en el sitio que desea.

### ***Cambio del orden de presentación de los elementos dentro de una dimensión***

Para cambiar el orden de presentación de los elementos dentro de una dimensión de la tabla (fila, columna o capa):

- ▶ Si los paneles de pivotado no está activados, elija en el menú de la tabla pivote:
  - Pivotar
  - Paneles de pivotado
- ▶ Arrastre y coloque los elementos dentro de la dimensión en el panel de pivotado.

### ***Movimiento de filas y columnas dentro de un elemento de una dimensión***

- ▶ En la propia tabla (no en los paneles de pivotado), pulse en la etiqueta de la fila o columna que desea mover.
- ▶ Arrastre la etiqueta hasta la nueva posición.
- ▶ En el menú contextual, seleccione Insertar antes o Intercambiar.

*Nota:* Asegúrese de que Arrastrar para copiar en el menú Edición *no* está activado (marcado). Si está activada, anule su selección.

## Para transponer filas y columnas

Si lo único que desea es transponer filas y columnas, existe una alternativa sencilla al uso de los paneles de pivotado:

- ▶ Elija en los menús:
  - Pivotar
  - Transponer filas y columnas

Esta operación tiene el mismo efecto que si se arrastraran todos los elementos de fila hasta la dimensión de las columnas y viceversa.

## Agrupación de filas o columnas

- ▶ Seleccione las etiquetas de las filas o las columnas que desee agrupar; para ello, pulse y arrastre, o bien mantenga pulsada la tecla Mayús para seleccionar varias etiquetas.
- ▶ Elija en los menús:
  - Edición
  - Grupo

Se insertará automáticamente una etiqueta de grupo. Pulse dos veces en la etiqueta de grupo para editar el texto de la etiqueta.

Figura 11-2  
Etiquetas y grupos de filas y columnas

		Etiquette de groupe de colonnes		Total
		Female	Male	
Etiquette de Groupe de Lignes	Clerical	206	157	363
	Custodial		27	27
	Manager	10	74	84

*Nota:* para añadir filas o columnas a un grupo existente, debe desagrupar en primer lugar los elementos que hay actualmente en el grupo. A continuación, puede crear un nuevo grupo que incluya los elementos adicionales.

## Desagrupación de filas o columnas

- ▶ Pulse en cualquier punto de la etiqueta de grupo de las filas o columnas que desea desagrupar.
- ▶ Elija en los menús:
  - Edición
  - Desagrupar

La desagrupación eliminará automáticamente la etiqueta de grupo.

## Rotación de etiquetas de filas o columnas

Puede rotar las etiquetas entre la presentación horizontal y vertical para las etiquetas de columna más al interior y las etiquetas de columna más al exterior

- ▶ Elija en los menús:  
Formato  
Rotar etiquetas de columna interior

o

- Formato  
Rotar etiquetas de fila exterior

Figura 11-3

*Etiquetas de columna rotadas*

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Menos de \$25	1174	18,3	18,3	18,3
	\$25 - \$49	2388	37,3	37,3	55,7
	\$50 - \$74	1120	17,5	17,5	73,2
	\$75+	1718	26,8	26,8	100,0
	Total	6400	100,0	100,0	

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Menos de \$25	1174	18,3	18,3	18,3
	\$25 - \$49	2388	37,3	37,3	55,7
	\$50 - \$74	1120	17,5	17,5	73,2
	\$75+	1718	26,8	26,8	100,0
	Total	6400	100,0	100,0	

Sólo se pueden rotar las etiquetas de columna más al interior y las etiquetas de fila más al exterior.

## ***Trabajo con capas***

Es posible mostrar una tabla de dos dimensiones distinta para cada categoría o combinación de categorías. La tabla se puede considerar como una apilación de capas en la que sólo puede verse la capa superior.

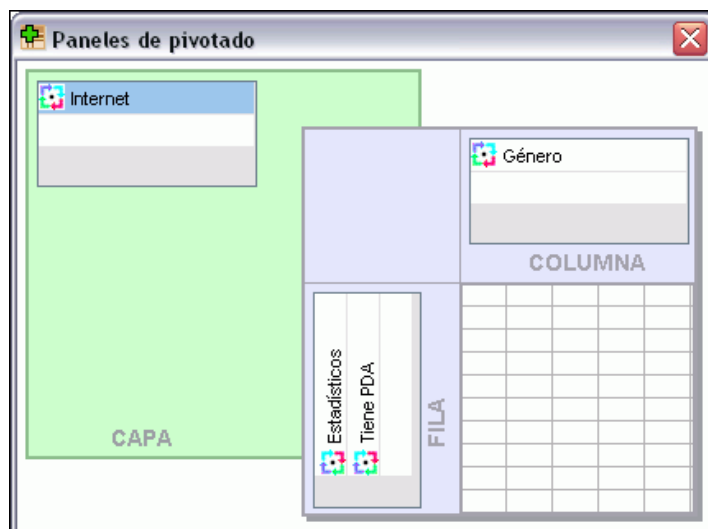
### ***Creación y presentación de capas***

Para crear capas:

- ▶ Active la tabla pivote.
- ▶ Si los paneles de pivotado no está activados, elija en el menú de la tabla pivote:  
Pivotar  
Paneles de pivotado
- ▶ Arrastre un elemento de la dimensión de filas o columnas a la dimensión de capas.



Figura 11-4  
Desplazamiento de categorías a las capas



Al desplazar elementos a la dimensión de capas se crea una tabla multidimensional, pero sólo se muestra un “sector” bidimensional. La tabla visible es la tabla de la capa superior. Por ejemplo, si una variable categórica sí/no está en la dimensión de capa, la tabla multidimensional tendrá dos capas: una para la categoría *sí* y otra para la categoría *no*.

Figura 11-5  
Categorías en capas distintas

Internet=Sí			Género		
			Hombre	Mujer	Total
Tiene PDA	Sí	Recuento	1962	1897	3859

Internet=No			Género		
			Hombre	Mujer	Total
Tiene PDA	Sí	Recuento	512	513	1025
		Frecuencia esperada	512,5	512,5	1025,0
		% de Tiene PDA	50,0%	50,0%	100,0%
	No	Recuento	306	305	611
		Frecuencia esperada	305,5	305,5	611,0
		% de Tiene PDA	50,1%	49,9%	100,0%
Total		Recuento	818	818	1636
		Frecuencia esperada	818,0	818,0	1636,0
		% de Tiene PDA	50,0%	50,0%	100,0%

### **Cambio de la capa mostrada**

- Seleccione una categoría en la lista desplegable de capas (en la propia tabla pivote, no en el panel de pivotado).

Figura 11-6  
Selección de capas en las listas desplegables

Internet		Género			
		Hombre	Mujer	Total	
Tiene PDA	Sí	Recuento	512	513	1025
		Frecuencia esperada	512,5	512,5	1025,0
		% de Tiene PDA	50,0%	50,0%	100,0%
	No	Recuento	306	305	611
		Frecuencia esperada	305,5	305,5	611,0
		% de Tiene PDA	50,1%	49,9%	100,0%
Total		Recuento	818	818	1636
		Frecuencia esperada	818,0	818,0	1636,0
		% de Tiene PDA	50,0%	50,0%	100,0%

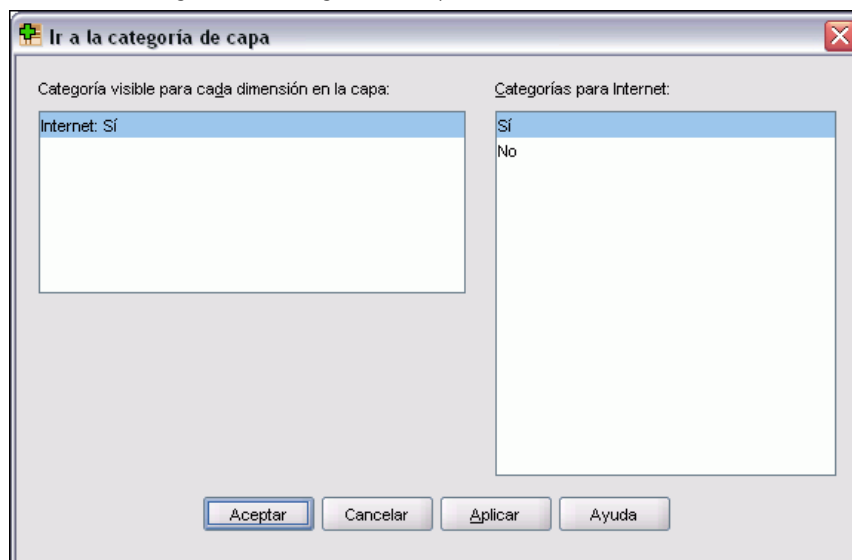
### Ir a la categoría de capa

Ir a la categoría de capa permite cambiar las capas de una tabla pivote. Este cuadro de diálogo es especialmente útil cuando existen muchas capas o la capa seleccionada tiene muchas categorías.

- Elija en los menús:

Pivotar  
Ir a capa...

Figura 11-7  
Cuadro de diálogo Ir a la categoría de capa



- En la lista Categoría visible, seleccione una dimensión de capa. La lista Categorías mostrará todas las categorías de la dimensión seleccionada.
- En la lista Categorías, seleccione la categoría que desea y, a continuación, haga clic en Aceptar o Aplicar.

## **Visualización y ocultación de elementos**

Se pueden ocultar muchos tipos de casillas, incluidas:

- Las etiquetas de dimensión
- Las categorías, incluidas la casilla de la etiqueta y las casillas de datos de una fila o columna
- Las etiquetas de las categorías (sin ocultar las casillas de datos)
- Las notas al pie, los títulos y los pies

### **Ocultación de filas y columnas de una tabla**

- ▶ Pulse con el botón derecho del ratón en la etiqueta de categoría de la fila o la columna que desee ocultar.
- ▶ Elija en el menú contextual:  
Seleccione  
Casillas de datos y etiquetas
- ▶ Vuelva a pulsar con el botón derecho del ratón en la etiqueta de categoría y seleccione Ocultar categoría en el menú contextual.

*o*

- ▶ En el menú Ver, seleccione Ocultar.

### **Visualización de las filas y columnas ocultas de una tabla**

- ▶ Pulse con el botón derecho del ratón en otra etiqueta de fila o columna de la misma dimensión que la fila o columna oculta.
- ▶ Elija en el menú contextual:  
Seleccione  
Casillas de datos y etiquetas
- ▶ Elija en los menús:  
Ver  
Mostrar todas las categorías en [nombre de la dimensión]

*o*

- ▶ Para mostrar todas las filas y columnas ocultas de una tabla pivote activada, seleccione en los menús:  
Ver  
Mostrar todo

Al hacerlo se mostrarán todas las filas y columnas ocultas de la tabla. Si está seleccionada la opción Ocultar filas y columnas vacías en Propiedades de tabla para esta tabla, las filas o columnas completamente vacías permanecerán ocultas.

### ***Ocultación y visualización de las etiquetas de dimensión***

- ▶ Seleccione la etiqueta de dimensión o cualquier etiqueta de categoría dentro de la dimensión.
- ▶ En el menú Ver o el menú contextual, seleccione Ocultar etiqueta de dimensión o Mostrar etiqueta de dimensión.

### ***Ocultación y presentación de los títulos de tabla***

Para ocultar un título:

- ▶ Seleccione el título.
- ▶ En el menú Ver, seleccione Ocultar.

Para mostrar los títulos ocultos:

- ▶ En el menú Ver, seleccione Mostrar todo.

### ***Aspectos de tabla***

Un aspecto de tabla es un conjunto de propiedades que definen el aspecto de una tabla. Puede seleccionar un aspecto previamente definido o crear uno propio.

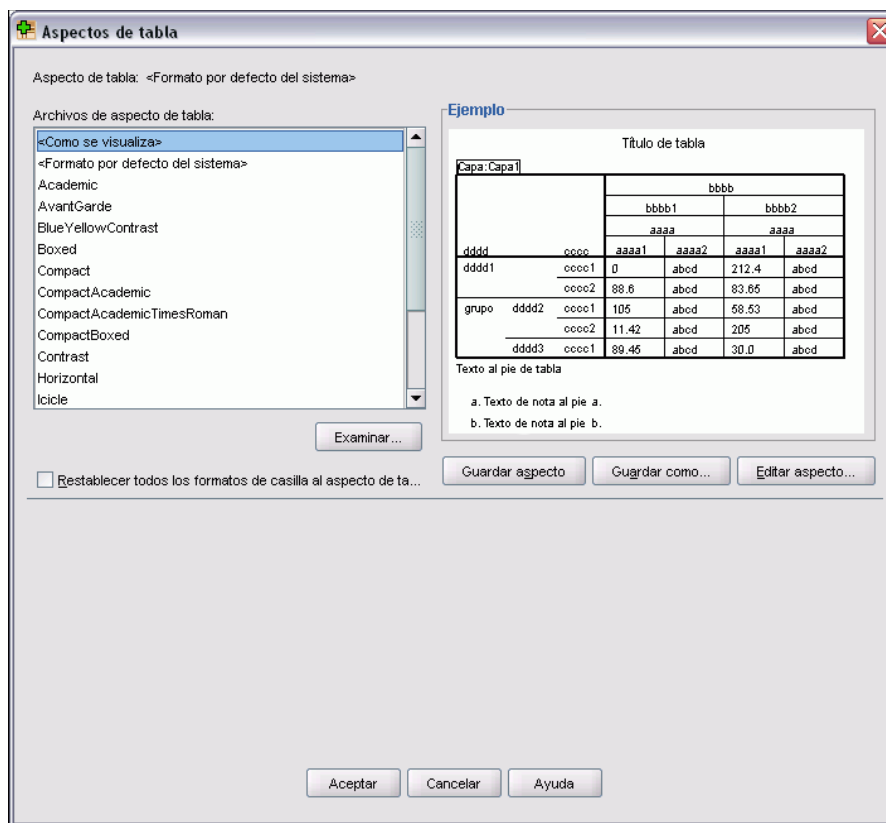
- Antes o después de aplicar un aspecto de tabla, se pueden cambiar los formatos de casillas individuales o de grupos de casillas utilizando las propiedades de casilla. Los formatos de casilla editados se conservan intactos, aunque se aplique un nuevo aspecto de tabla. Si desea obtener más información, consulte [Propiedades de casilla](#) en p. 252.
- Si lo desea, puede restablecer todas las casillas a los formatos de casilla definidos por el aspecto de tabla actual. De este modo se restablece cualquier casilla que se haya editado. Si la opción Como se visualiza está seleccionada en la lista de archivos de aspecto de tabla, las casillas editadas se restablecerán a las propiedades de tabla actuales.

*Nota:* los aspectos de tabla creados en versiones anteriores de SPSS no se pueden utilizar en la versión 16.0 o posterior.

### ***Para aplicar o guardar un aspecto de tabla***

- ▶ Active una tabla pivote.
- ▶ Elija en los menús:
  - Formato
  - Aspectos de tabla...

Figura 11-8  
Cuadro de diálogo Aspectos de tabla



- ▶ Seleccione un aspecto de tabla en la lista de archivos. Para seleccionar un archivo de otro directorio, pulse en Examinar.
- ▶ Pulse en Aceptar para aplicar el aspecto de tabla a la tabla pivote seleccionada.

### **Para editar o crear un aspecto de tabla**

- ▶ En el cuadro de diálogo Aspectos de tabla, seleccione un aspecto de tabla en la lista de archivos.
- ▶ Pulse en Editar aspecto.
- ▶ Edite las propiedades de la tabla seleccionando los atributos que desee y, a continuación, pulse en Aceptar.
- ▶ Pulse en Guardar aspecto para guardar el aspecto de tabla editado, o en Guardar como para guardarlo como un nuevo aspecto de tabla.

La edición de un aspecto de tabla sólo afecta a la tabla pivote seleccionada. El aspecto de tabla editado no se aplicará a ninguna otra tabla que utilice ese aspecto, a menos que se seleccione esa tabla y le vuelva a aplicar el aspecto.

## ***Propiedades de tabla***

Propiedades de tabla permite establecer las propiedades generales de una tabla, definir los estilos de casilla de varias partes de la tabla y guardar un conjunto de esas propiedades como un aspecto de tabla. Tiene la posibilidad de:

- Controlar las propiedades generales, como ocultar filas o columnas vacías y ajustar las propiedades de impresión.
- Controlar el formato y la posición de los marcadores de las notas al pie.
- Determinar formatos específicos para las casillas en el área de datos, para las etiquetas de fila y columna y para otras áreas de la tabla.
- Controlar la anchura y el color de las líneas que forman los bordes de cada área de la tabla.

### ***Para modificar las propiedades de las tablas pivote***

- ▶ Active la tabla pivote.
- ▶ Elija en los menús:
  - Formato
  - Propiedades de tabla...
- ▶ Elija una pestaña (General, Notas al pie, Formatos de casilla, Bordes o Impresión).
- ▶ Seleccione las opciones que desee.
- ▶ Pulse en Aceptar o Aplicar.

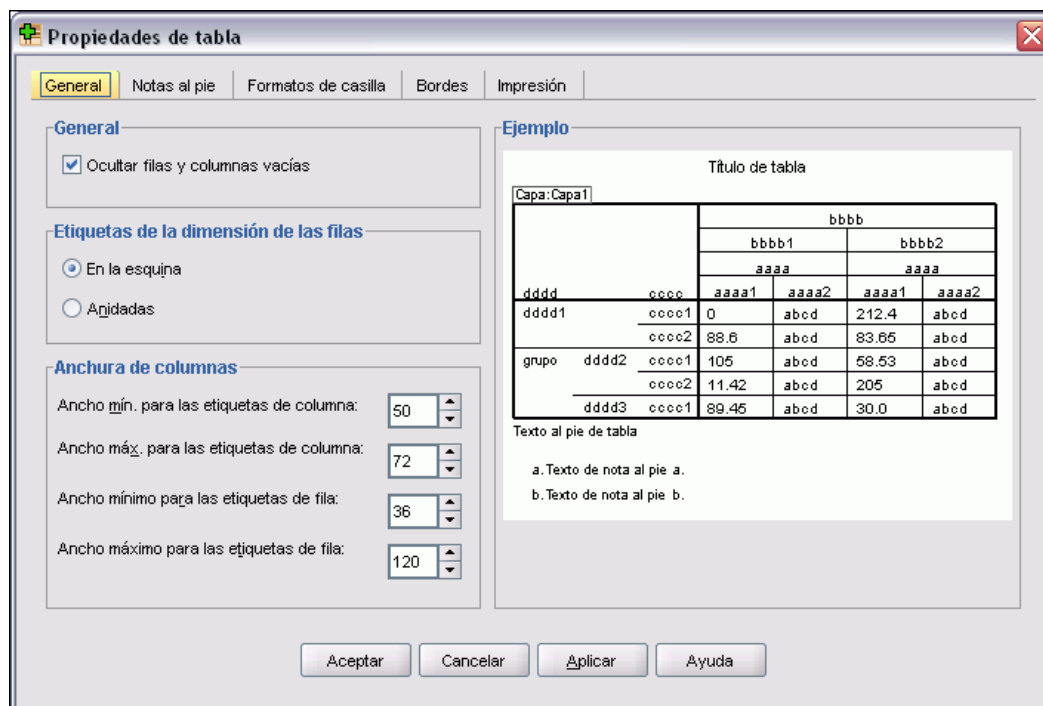
Las nuevas propiedades se aplicarán a la tabla pivote seleccionada. Para aplicar nuevas propiedades de tabla a un aspecto de tabla en lugar de hacerlo a la tabla seleccionada, edite el aspecto de tabla (menú Formato, Aspectos de tabla).

### ***Propiedades de tabla: General***

Algunas propiedades son aplicables a la tabla en su conjunto. Tiene la posibilidad de:

- Mostrar u ocultar filas y columnas vacías. (Una fila o una columna vacía no contiene nada en ninguna casilla de datos.)
- Controlar la colocación de las etiquetas de fila, que pueden estar anidadas o en la esquina superior izquierda.
- Controlar la anchura máxima y mínima de la columna (expresada en puntos).

Figura 11-9  
Pestaña General del cuadro de diálogo Propiedades de tabla



Para modificar las propiedades generales de las tablas:

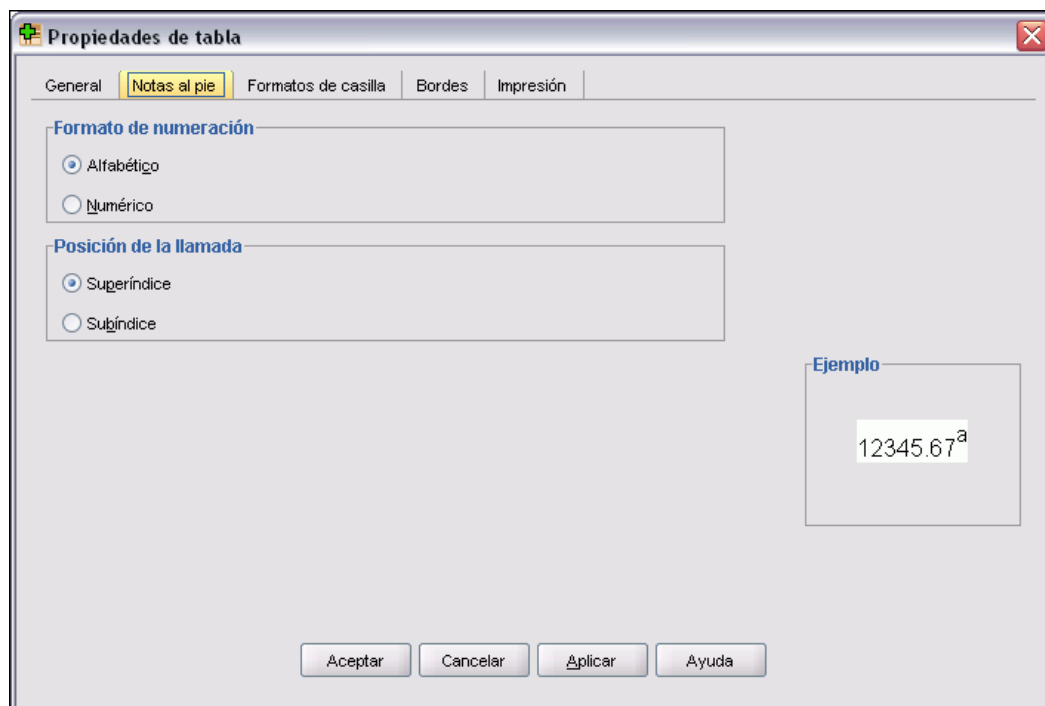
- ▶ Pulse en la pestaña General.
- ▶ Seleccione las opciones que desee.
- ▶ Pulse en Aceptar o Aplicar.

### **Propiedades de tabla: Notas al pie**

Las propiedades de los marcadores de las notas al pie incluyen el estilo y la posición en relación al texto.

- El estilo de los marcadores de las notas al pie puede ser números (1, 2, 3...) o letras (a, b, c...).
- Los marcadores de las notas al pie se pueden anexar al texto como superíndices o como subíndices.

Figura 11-10  
Pestaña Notas al pie del cuadro de diálogo Propiedades de tabla



Para modificar las propiedades de los marcadores de las notas al pie:

- ▶ Pulse en la pestaña Notas al pie.
- ▶ Elija el formato de numeración de notas al pie.
- ▶ Seleccione la posición del marcador.
- ▶ Pulse en Aceptar o Aplicar.

### ***Propiedades de tabla: Formatos de casilla***

Para el formato, una tabla se divide en áreas: título, capas, etiquetas de esquina, etiquetas de fila, etiquetas de columna, datos, texto al pie y notas al pie. Para cada área de una tabla se pueden modificar los formatos de casilla asociados. Los formatos de casilla incluyen características del texto (como fuente, tamaño, color y estilo), alineación horizontal y vertical, colores de fondo y márgenes internos de las casillas.



Figura 11-11  
Áreas de una tabla

Título	
Capas	
Etiqueta de Esquina	Etiqueta de Columnas
Etiqueta de Filas	Datos

Texto al pie

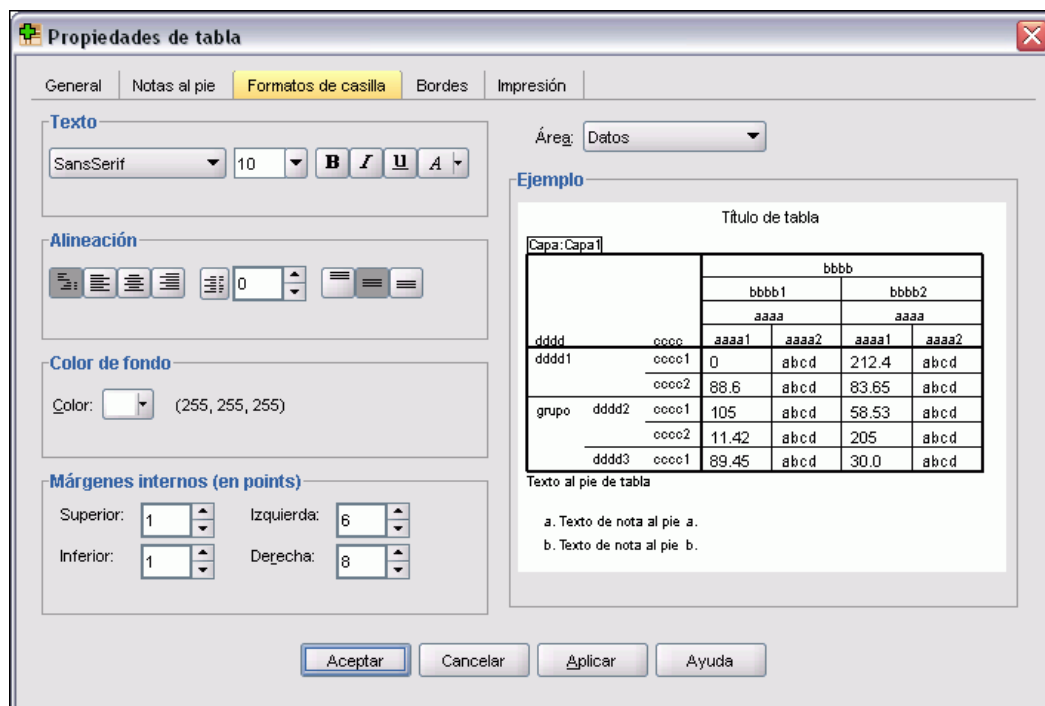
Notas al pie

Los formatos de casilla se aplican a las áreas (categorías de información); no son características de las casillas individuales. No son características de casillas individuales. Esta distinción es un aspecto importante a la hora de pivotar una tabla.

Por ejemplo:

- Si especifica una fuente en negrita como formato de casilla de las etiquetas de columna, éstas aparecerán en negrita independientemente de la información que se muestre actualmente en la dimensión de columna. Si mueve un elemento de la dimensión de columna a otra dimensión, no conservará la característica de negrita de las etiquetas de columna.
- Si pone en negrita las etiquetas de columna simplemente resaltando las casillas de una tabla pivote activada y pulsando en el botón Negrita de la barra de herramientas, el contenido de esas casillas permanecerá en negrita independientemente de la dimensión a las que las mueva, y las etiquetas de columna no conservarán la característica de negrita para otros elementos movidos a la dimensión de columna.

Figura 11-12  
Pestaña Formatos de casilla del cuadro de diálogo Propiedades de tabla



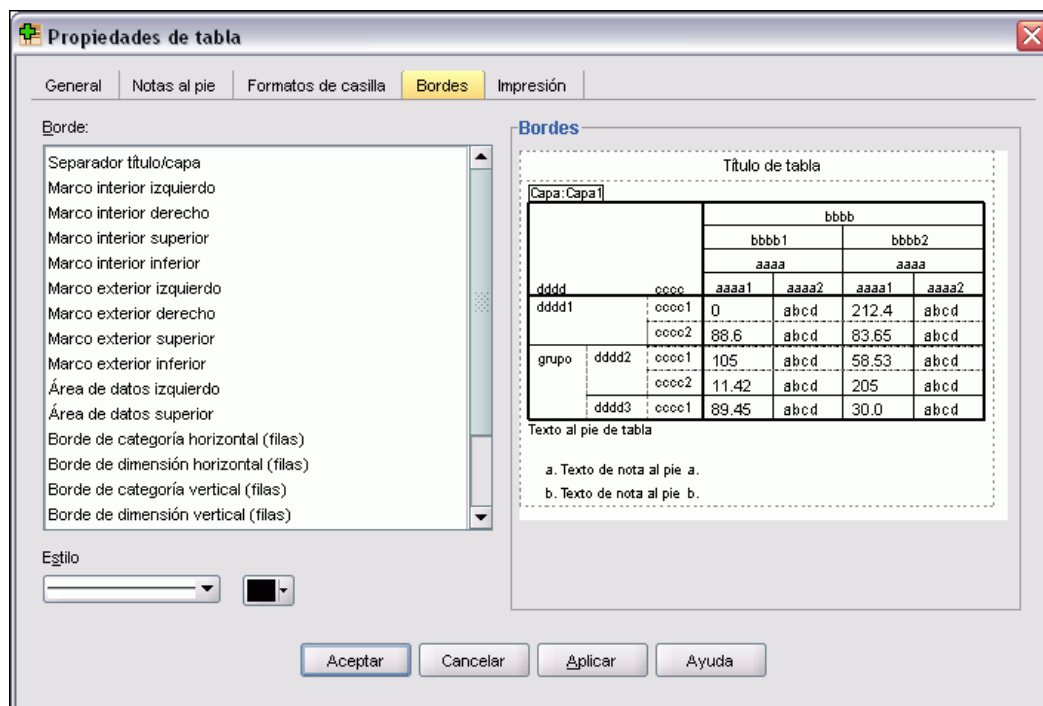
Para cambiar los formatos de casilla:

- ▶ Seleccione la pestaña Formatos de casilla.
- ▶ Elija un área de la lista desplegable o pulse en un área de la muestra.
- ▶ Seleccione las características para el área. Sus selecciones se reflejarán en la muestra.
- ▶ Pulse en Aceptar o Aplicar.

### **Propiedades de tabla: Bordes**

Para cada ubicación del borde de una tabla, puede seleccionar un estilo de línea y un color. Si selecciona Ninguno como estilo, no habrá ninguna línea en la ubicación seleccionada.

Figura 11-13  
Pestaña Bordes del cuadro de diálogo Propiedades de tabla



Para cambiar los bordes de la tabla:

- ▶ Pulse en la pestaña Bordes.
- ▶ Seleccione una ubicación del borde pulsando en el nombre que aparece en la lista o en una línea del área de muestra.
- ▶ Seleccione un estilo de línea o Ninguno.
- ▶ Elija un color.
- ▶ Pulse en Aceptar o Aplicar.

### ***Propiedades de tabla: Imprimiendo***

Puede controlar las siguientes propiedades de la impresión de las tablas pivote:

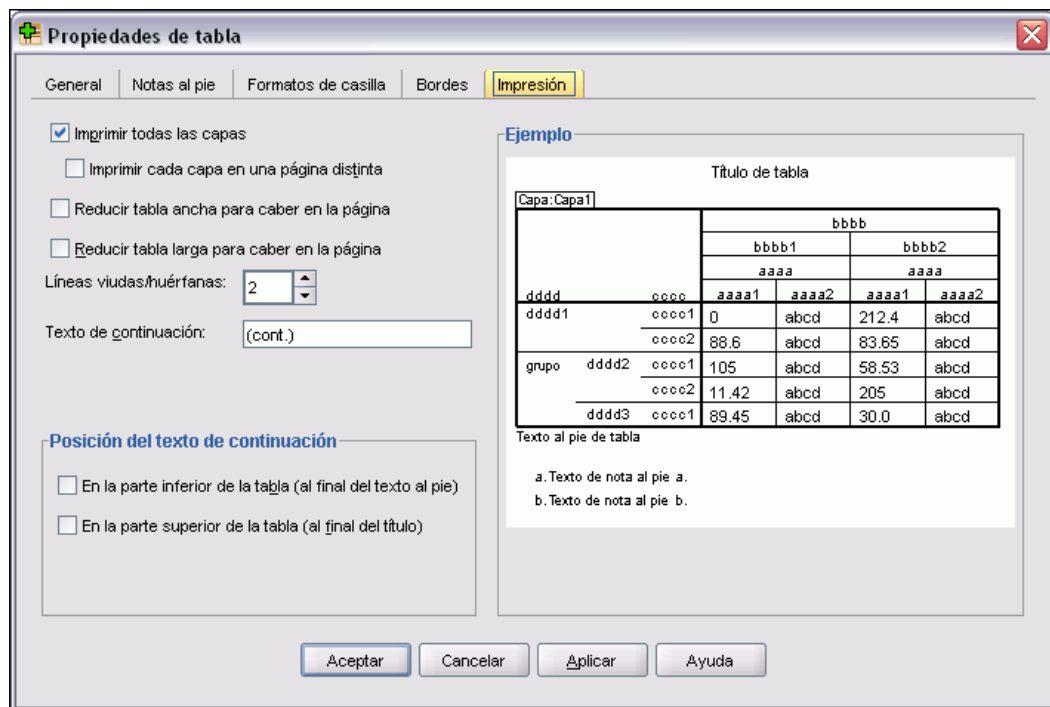
- Imprimir todas las capas o sólo la capa superior de la tabla e imprimir cada capa en una página distinta (esto sólo afecta a la impresión, no a la presentación de las capas en el Visor).
- Reducir una tabla horizontal o verticalmente para que se ajuste a la página a la hora de imprimirla.
- Controlar las líneas viudas y huérfanas mediante el control del número mínimo de filas y columnas que podrá incluir cualquier sección impresa de una tabla si ésta es demasiado ancha o larga para el tamaño de página definido.

*Nota:* Si una tabla es demasiado larga para que quepa en la página actual porque hay otros resultados en la página por encima de ella pero cabe en la longitud de página definida, la tabla se imprimirá automáticamente en una nueva página, independientemente del valor de líneas viudas/huérfanas establecido.

- Incluir el texto de continuación para las tablas que no quepan en una sola página. Este texto puede mostrarse en la parte inferior o superior de cada página. Si no selecciona ninguna de estas opciones, no aparecerá el texto de continuación.

Figura 11-14

Cuadro de diálogo *Propiedades de tabla: pestaña Impresión*



Para controlar las propiedades de impresión de la tabla pivote:

- ▶ Pulse en la pestaña *Impresión*.
- ▶ Seleccione las opciones de impresión que desee.
- ▶ Pulse en *Aceptar* o *Aplicar*.

## Propiedades de casilla

Las opciones de *Propiedades de casilla* sólo se aplican a una casilla individual seleccionada. Puede cambiar la fuente, el formato del valor, la alineación, los márgenes y los colores. Las propiedades de casilla anulan las propiedades de tabla; por tanto, si modifica estas últimas, no cambiarán las propiedades de casilla aplicadas individualmente.

Para modificar las propiedades de una casilla:

- ▶ Active una tabla y seleccione una o varias de sus casillas.

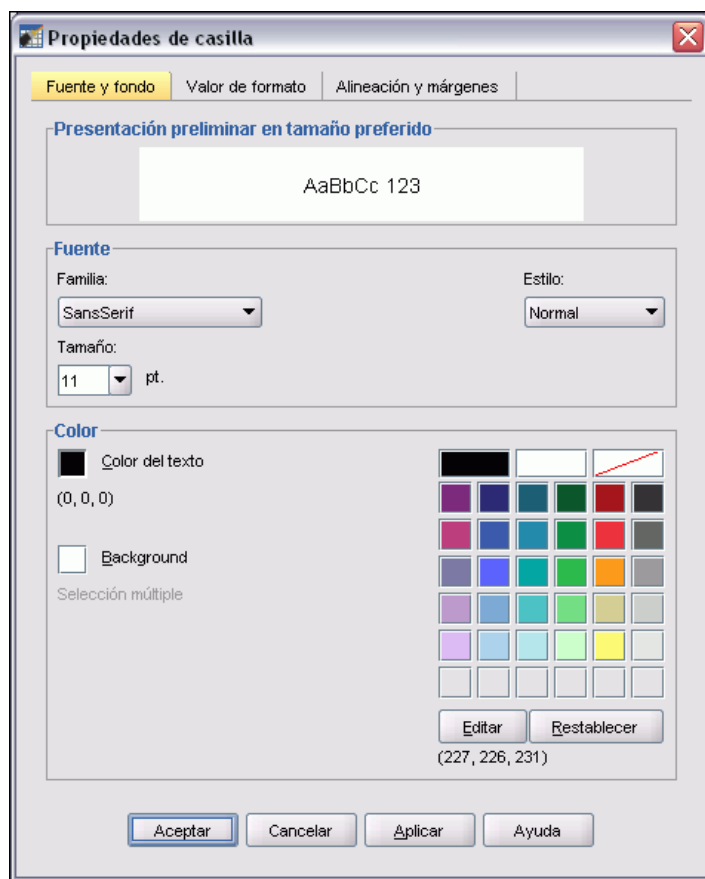
- En el menú Formato o en el menú contextual, seleccione Propiedades de casilla.

## Fuente y fondo

La pestaña Fuente y fondo controla el estilo y color de la fuente, y el color de fondo de las casillas seleccionadas de la tabla.

Figura 11-15

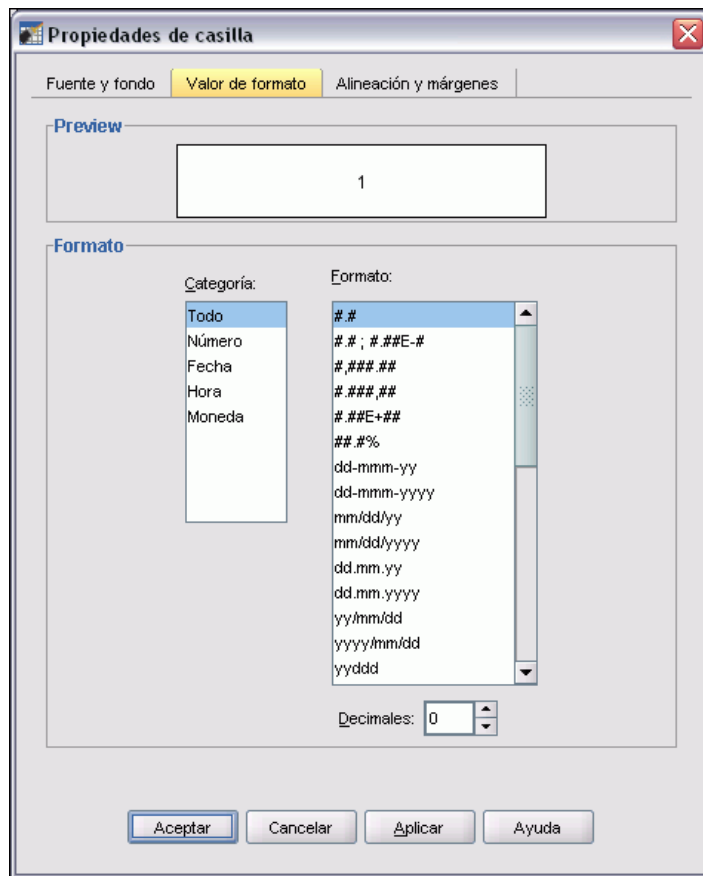
Cuadro de diálogo Propiedades de casilla: pestaña Fuente y fondo



## Formato de valor

La pestaña Formato de valor controla los formatos del valor para las casillas seleccionadas. Puede seleccionar formatos para números, fechas, horas o monedas, así como ajustar el número de dígitos decimales mostrados.

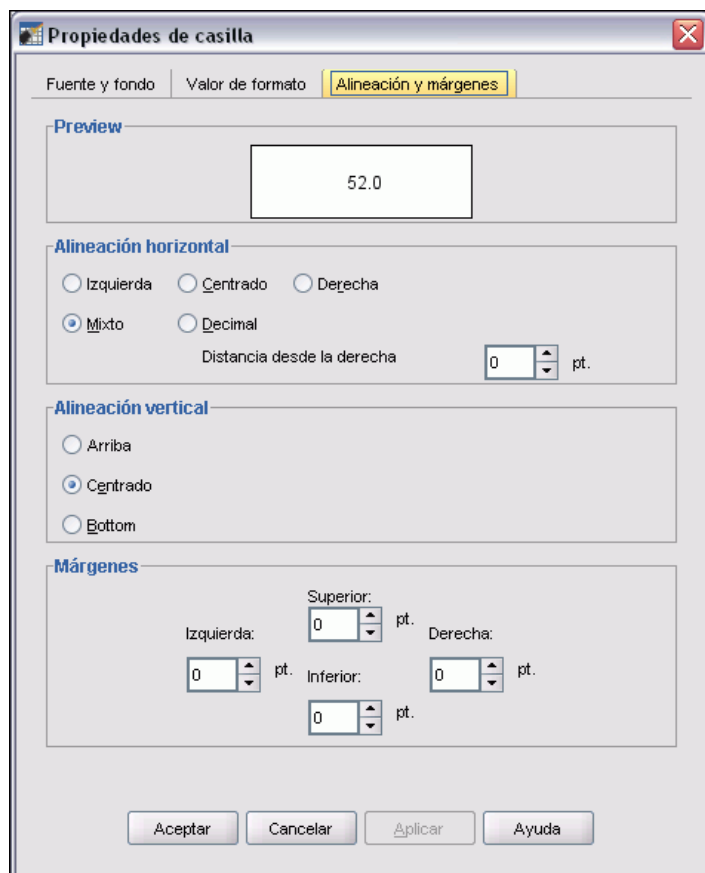
Figura 11-16  
Cuadro de diálogo *Propiedades de casilla*: pestaña *Formato de valor*



## ***Alineación y márgenes***

La pestaña *Alineación y márgenes* controla la alineación horizontal y vertical de los valores y los márgenes superior, inferior, izquierdo y derecho para las casillas seleccionadas. La alineación horizontal *Mixta* alinea el contenido de cada celda según su tipo. Por ejemplo, las fechas se alinean a la derecha y los valores de texto a la izquierda.

Figura 11-17  
Cuadro de diálogo *Propiedades de casilla*: pestaña *Alineación y márgenes*



## Texto y notas al pie

Puede añadir a la tabla texto y notas al pie. También puede ocultar el texto y las notas al pie, modificar los marcadores de las notas al pie y reenumerar las notas al pie.

Algunos de los atributos de las notas al pie están controlados por las propiedades de la tabla. Si desea obtener más información, consulte [Propiedades de tabla: Notas al pie](#) en p. 247.

## Adición de texto y notas al pie

Para añadir texto a la tabla:

- ▶ En el menú Insertar, seleccione Texto al pie.

Se puede anexar una nota al pie a cualquier elemento de la tabla. Para añadir una nota al pie:

- ▶ Pulse en un título, casilla o texto al pie dentro de una tabla pivote activada.
- ▶ En el menú Insertar, seleccione Nota al pie.

### ***Para mostrar u ocultar el texto al pie***

Para ocultar texto:

- ▶ Seleccione el texto.
- ▶ En el menú Ver, seleccione Ocultar.

Para mostrar los textos ocultos:

- ▶ En el menú Ver, seleccione Mostrar todo.

### ***Para ocultar o mostrar una nota al pie de una tabla***

Para ocultar una nota al pie:

- ▶ Seleccione la nota al pie.
- ▶ En el menú Ver, seleccione Ocultar o seleccione Ocultar nota al pie en el menú contextual.

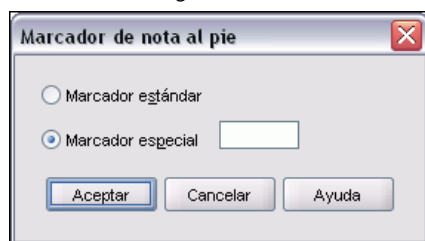
Para mostrar las notas al pie ocultas:

- ▶ En el menú Ver, seleccione Mostrar todas las notas al pie.

### ***Marcador de nota al pie***

Marcador de nota al pie cambia el carácter o los caracteres que pueden usarse para marcar una nota al pie.

Figura 11-18  
Cuadro de diálogo Marcador de nota al pie



Para cambiar los marcadores de las notas al pie:

- ▶ Seleccione una nota al pie.
- ▶ En el menú Formato, seleccione Marcador de nota al pie.
- ▶ Escriba uno o dos caracteres.



### ***Para volver a numerar las notas al pie***

Cuando se ha pivotado una tabla intercambiando las filas, las columnas y las capas, las notas al pie pueden quedar desordenadas. Para volver a numerarlas:

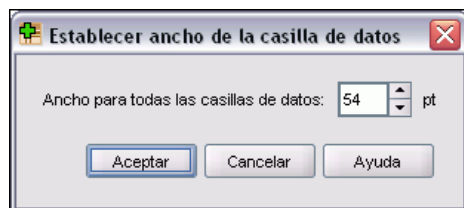
- ▶ En el menú Formato, seleccione Renumerar notas al pie.

### ***Anchos de casillas de datos***

El cuadro de diálogo Establecer ancho de la casilla de datos se utiliza para determinar la misma anchura para todas las casillas de datos.

Figura 11-19

*Cuadro de diálogo Establecer ancho de la casilla de datos*



Para establecer el ancho de todas las casillas de datos:

- ▶ Elija en los menús:  
Formato  
Ancho de casillas de datos...
- ▶ Introduzca un valor para el ancho de las casillas.

### ***Cambio del ancho de columna***

- ▶ Pulse en el borde de la columna y arrástrelo.

### ***Presentación de los bordes ocultos de una tabla pivote***

En las tablas que tienen pocos bordes visibles es posible mostrar los bordes ocultos. Esto puede simplificar la realización de algunas tareas, como el cambio de los anchos de columna.

- ▶ En el menú Ver, seleccione Cuadrículas.

Figura 11-20  
Cuadrículas mostradas para los bordes ocultos

Tabla de contingencia Owns PDA * Gender * Internet						
Internet				Gender		
				Female	Male	Total
No	Owens PDA	No	Recuento	1897	1962	3859
			Frecuencia esperada	1903,4	1955,6	3859,0
			% de Owens PDA	49,2%	50,8%	100,0%
		Yes	Recuento	327	323	650

Tabla de contingencia Owns PDA * Gender * Internet						
Internet				Gender		
				Female	Male	Total
No	Owens PDA	No	Recuento	1897	1962	3859
			Frecuencia esperada	1903,4	1955,6	3859,0
			% de Owens PDA	49,2%	50,8%	100,0%
		Yes	Recuento	327	323	650
			Frecuencia esperada	320,6	329,4	650,0
			% de Owens PDA	50,3%	49,7%	100,0%
		Total	Recuento	2224	2285	4509
			Frecuencia esperada	2224,0	2285,0	4509,0
			% de Owens PDA	49,3%	50,7%	100,0%
Yes	Owens PDA	No	Recuento	513	512	1025
			Frecuencia esperada	512,5	512,5	1025,0
			% de Owens PDA	50,0%	50,0%	100,0%
		Yes	Recuento	305	306	611
			Frecuencia esperada	305,5	305,5	611,0
			% de Owens PDA	49,9%	50,1%	100,0%
		Total	Recuento	818	818	1636
			Frecuencia esperada	818,0	818,0	1636,0
			% de Owens PDA	50,0%	50,0%	100,0%

## Selección de filas y columnas de una tabla pivote

En las tablas pivote existen algunas restricciones en el modo de seleccionar filas y columnas enteras, y el resalte visual que indica la fila o la columna seleccionada puede abarcar áreas no contiguas de la tabla. Para seleccionar una fila o columna entera:

- ▶ Pulse en una etiqueta de fila o de columna.
- ▶ Elija en los menús:
  - Edición
  - Seleccione
  - Casillas de datos y etiquetas
- o
- ▶ Pulse con el botón derecho del ratón en la etiqueta de categoría de la fila o columna.
- ▶ Elija en el menú contextual:
  - Seleccione
  - Casillas de datos y etiquetas

o

- ▶ Pulse Ctrl+Alt y, a continuación, la etiqueta de la fila o columna.

## ***Impresión de tablas pivote***

Existen muchos factores que pueden afectar al aspecto que presentan las tablas pivote impresas. Estos factores pueden controlarse cambiando los atributos de las tablas pivote.

- En las tablas pivote multidimensionales (tablas con capas), puede imprimir todas las capas o sólo la capa superior (la visible). Si desea obtener más información, consulte [Propiedades de tabla: Imprimiendo](#) en p. 251.
- En las tablas pivote largas o anchas, puede cambiar automáticamente el tamaño de la tabla para que quepa en la página o controlar la posición de las rupturas de tabla y de los saltos de página. Si desea obtener más información, consulte [Propiedades de tabla: Imprimiendo](#) en p. 251.
- Para las tablas que son demasiado anchas o largas para una sola página, puede controlar la ubicación de las rupturas de tabla entre las páginas.

Utilice Presentación preliminar, en el menú Archivo, para ver cómo aparecerán las tablas pivote en la página impresa.

## ***Control de las rupturas de tabla para tablas anchas y largas***

Las tablas pivote que son o demasiado anchas o largas para imprimirlas en el tamaño de página definido se dividen automáticamente y se imprimen en varias secciones. (En las tablas anchas, se imprimen varias secciones en la misma página si hay suficiente espacio.) Tiene la posibilidad de:

- Controlar la posición de las filas y las columnas por donde se dividen las tablas de gran tamaño.
- Especificar las filas y las columnas que deben permanecer unidas cuando se dividen las tablas.
- Ajustar las tablas de gran tamaño para que quepan en el tamaño de página definido.

### ***Para especificar las rupturas de filas y columnas para las tablas pivote impresas***

- ▶ Pulse en la etiqueta de la columna situada a la izquierda del lugar donde desee insertar la ruptura, o en la etiqueta de la fila sobre la que desee insertarla.
- ▶ Elija en los menús:
  - Formato
  - Romper aquí

### ***Para especificar las filas o las columnas que se han de mantener unidas***

- ▶ Seleccione las etiquetas de las filas o columnas que desee mantener unidas (Pulse y arrastre, o bien mantenga pulsada la tecla Mayús para seleccionar varias etiquetas de fila o columna.)
- ▶ Elija en los menús:
  - Formato
  - Mantener juntos

## ***Creación de un gráfico a partir de una tabla pivote***

- ▶ Pulse dos veces en la tabla pivote para activarla.
- ▶ Seleccione las filas, columnas o celdas que desee mostrar en el gráfico.
- ▶ Pulse con el botón derecho del ratón en cualquier punto del área seleccionada.
- ▶ Seleccione Crear gráfico en el menú contextual y seleccione el tipo de gráfico.

## ***Trabajar con sintaxis de comandos***

Un lenguaje de comandos eficaz permite guardar y automatizar muchas tareas habituales. El lenguaje de comandos también proporciona algunas funcionalidades no incluidas en los menús y cuadros de diálogo.

Puede acceder a la mayoría de los comandos desde los menús y cuadros de diálogo. No obstante, algunos comandos y opciones sólo están disponibles mediante el uso del lenguaje de comandos. El lenguaje de comandos también permite guardar los trabajos en un archivo de sintaxis, con lo que podrá repetir los análisis en otro momento o ejecutarlos en un trabajo automatizado con un trabajo de producción.

Un archivo de sintaxis es simplemente un archivo de texto que contiene comandos. Aunque es posible abrir una ventana de sintaxis y escribir comandos, suele ser más sencillo permitir que el programa le ayude a construir un archivo de sintaxis mediante uno de los siguientes métodos:

- Pegando la sintaxis de comandos desde los cuadros de diálogo
- Copiando la sintaxis desde las anotaciones de los resultados
- Copiando la sintaxis desde el archivo diario

La información de referencia detallada sobre la sintaxis de comandos está disponible en dos formatos: integrada en el sistema de ayuda global y como un archivo PDF independiente de referencia de sintaxis de comandos, denominado *Command Syntax Reference*, igualmente disponible en el menú Ayuda.

Si pulsa la tecla F1, podrá disponer de la ayuda contextual relativa del comando actual en una ventana de sintaxis.

### ***Reglas de la sintaxis***

Al ejecutar comandos desde una ventana de sintaxis de comandos en el transcurso de una sesión, lo hará en modo **interactivo**.

Las siguientes reglas se aplican a las especificaciones de los comandos en el modo interactivo:

- Cada comando debe comenzar en una línea nueva. Los comandos pueden comenzar en cualquier columna de una línea de comandos y continuar en tantas líneas como se precise. Existe una excepción con el comando `END DATA`, que debe comenzar en la primera columna de la primera línea que sigue al final de los datos.
- Cada comando debe terminar con un punto como terminador del comando. Sin embargo, es mejor omitir el terminador en `BEGIN DATA`, para que los datos interactivos se traten como una especificación continua.

- El terminador del comando debe ser el último carácter de un comando que no esté en blanco.
- En ausencia de un punto como terminador del comando, las líneas en blanco se interpretan como un terminador del comando.

*Nota:* para que exista compatibilidad con otros modos de ejecución de comandos (incluidos los archivos de comandos que se ejecutan con los comandos `INSERT` o `INCLUDE` en una sesión interactiva), la sintaxis de línea de comandos no debe exceder los 256 bytes.

- La mayoría de los subcomandos están separados por barras inclinadas (/). La barra inclinada que precede al primer subcomando de un comando, generalmente es opcional.
- Los nombres de variable deben escribirse completos.
- El texto incluido entre apóstrofes o comillas debe ir contenido en una sola línea.
- Debe utilizarse un punto (.) para indicar decimales, independientemente de la configuración regional.
- Los nombres de variable que terminen en un punto pueden causar errores en los comandos creados por los cuadros de diálogo. No es posible crear nombres de variable de este tipo en los cuadros de diálogo y en general deben evitarse.

La sintaxis de comandos no distingue las mayúsculas de las minúsculas y permite el uso de abreviaturas de tres o cuatro letras en la mayoría de las especificaciones de los comandos. Puede usar tantas líneas como desee para especificar un único comando. Puede añadir espacios o líneas de separación en casi cualquier punto donde se permita un único espacio en blanco, como alrededor de las barras inclinadas, los paréntesis, los operadores aritméticos o entre los nombres de variable. Por ejemplo:

```
FRECUENCIES VARIABLES=CATLAB SEXO /PERCENTILES=25 50 75 /BARCHART.
```

y

```
freq var=catlab sexo /percent=25 50 75 /bar.
```

son alternativas aceptables que generan los mismos resultados.

### **Archivos *INCLUDE***

Para los archivos de comandos ejecutados mediante el comando `INCLUDE`, se aplican las reglas de sintaxis del modo por lotes.

Las siguientes reglas se aplican a las especificaciones de los comandos en el modo por lotes:

- Todos los comandos del archivo de comandos deben comenzar en la columna 1. Puede utilizar los signos más (+) o menos (-) en la primera columna si desea sangrar la especificación del comando para facilitar la lectura del archivo de comandos.
- Si se utilizan varias líneas para un comando, la columna 1 de cada línea de continuación debe estar en blanco.
- Los terminadores de los comandos son opcionales.
- Las líneas no pueden exceder los 256 bytes; los caracteres adicionales quedarán truncados.

A menos que tenga archivos de comandos que ya utilizan el comando INCLUDE, debe utilizar el comando INSERT en su lugar dado que puede adaptar los archivos de comandos que se ajustan a los dos conjuntos de reglas. Si genera la sintaxis de comandos pegando las selecciones del cuadro de diálogo en una ventana de sintaxis, el formato de los comandos es apto para cualquier modo de operación. Consulte la *referencia de sintaxis de comandos* (disponible en formato PDF en el menú Ayuda) si desea obtener más información.

## Pegar sintaxis desde cuadros de diálogo

La manera más fácil de construir un archivo de sintaxis de comandos es realizando las selecciones en los cuadros de diálogo y pegar la sintaxis de las selecciones en una ventana de sintaxis. Si pega la sintaxis en cada paso de un análisis largo, podrá generar un archivo de trabajo que le permitirá repetir el análisis con posterioridad o ejecutar un trabajo automatizado con la Unidad de producción.

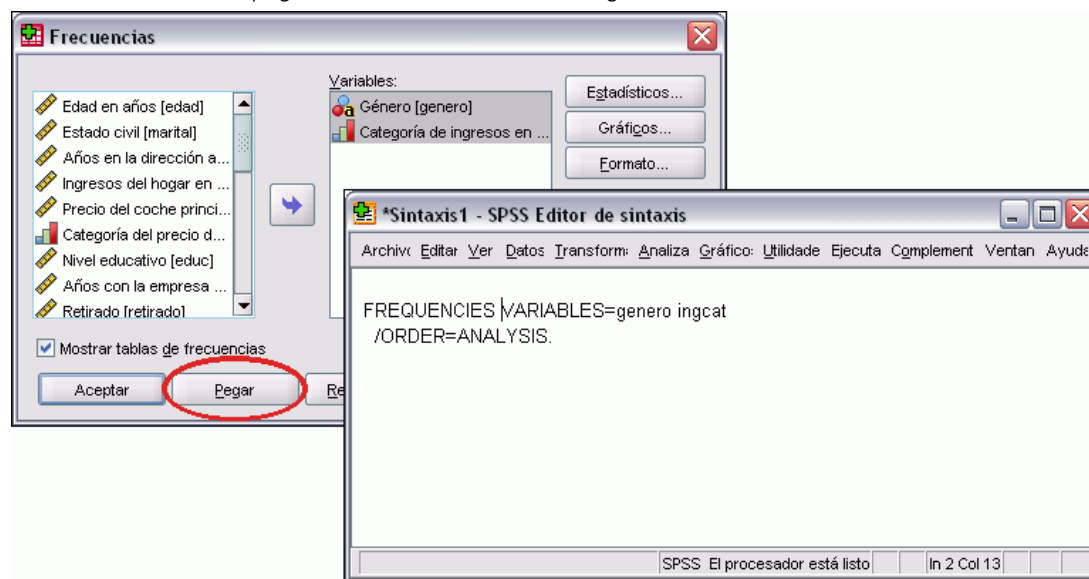
En la ventana de sintaxis, puede ejecutar la sintaxis pegada, editarla y guardarla en un archivo de sintaxis.

### Para pegar sintaxis desde cuadros de diálogo

- ▶ Abra el cuadro de diálogo y realice las selecciones que desee.
- ▶ Pulse en Pegar.

La sintaxis de comandos se pegará en la ventana de sintaxis designada. Si no tiene abierta una ventana de sintaxis, se abrirá automáticamente una nueva y se pegará la sintaxis en ella.

Figura 12-1  
Sintaxis de comandos pegada desde un cuadro de diálogo

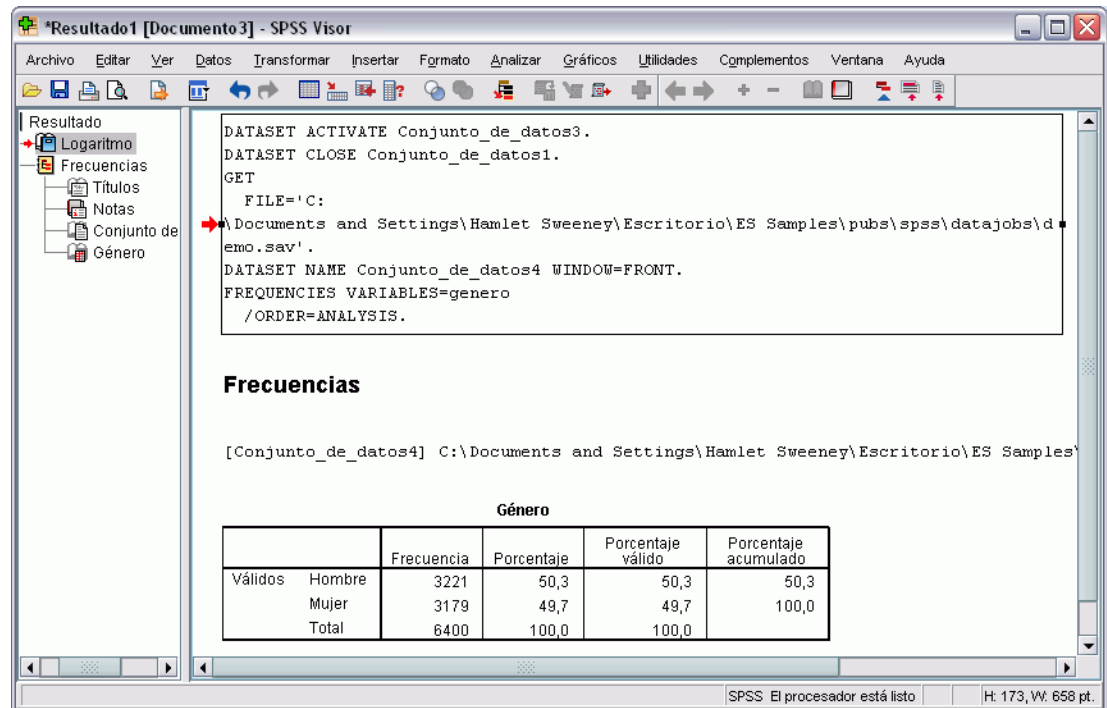


## Copia de la sintaxis desde las anotaciones de los resultados

Puede construir un archivo de sintaxis copiando la sintaxis de comandos de la anotación que aparece en el Visor. Para usar este método debe seleccionar **Mostrar comandos en anotaciones** en la configuración del Visor (menú Edición, Opciones, pestaña Visor) antes de ejecutar el análisis. Todos los comandos aparecerán en el Visor junto con los resultados del análisis.

En la ventana de sintaxis, puede ejecutar la sintaxis pegada, editarla y guardarla en un archivo de sintaxis.

Figura 12-2  
Sintaxis de comando en la anotación



### Para copiar la sintaxis desde las anotaciones de los resultados

- ▶ Antes de ejecutar el análisis, elija en los menús:
  - Edición
  - Opciones...
- ▶ En la pestaña Visor, seleccione **Mostrar comandos en anotaciones**.  
 Mientras ejecuta los análisis, los comandos de las selecciones del cuadro de diálogo se graban en la anotación.
- ▶ Abra un archivo de sintaxis previamente guardado o cree uno nuevo. Para crear un archivo de sintaxis, elija en los menús:
  - Archivo
  - Nuevo
  - Sintaxis



- ▶ En el Visor, pulse dos veces en un elemento de anotación para activarlo.
- ▶ Seleccione el texto que desee copiar.
- ▶ Seleccione en los menús del Visor:
  - Edición
  - Copiar
- ▶ En una ventana de sintaxis, elija en los menús:
  - Edición
  - Pegar

## ***Para ejecutar la sintaxis de comandos***

- ▶ Resalte los comandos que desee ejecutar en la ventana de sintaxis.
  - ▶ Pulse en el botón Ejecutar (el triángulo que apunta hacia la derecha) en la barra de herramientas del Editor de sintaxis.
- o*
- ▶ Seleccione una de las opciones del menú Ejecutar.
    - **Todo.** Ejecuta todos los comandos de la ventana de sintaxis.
    - **Selección.** Ejecuta los comandos seleccionados. Esto incluye los comandos parcialmente resaltados. Si no se ha seleccionado ninguno, ejecuta el comando donde se encuentra el cursor.
    - **Actual.** Ejecuta el comando donde se encuentra el cursor.
    - **Hasta el final.** Ejecuta todos los comandos incluidos desde la posición actual del cursor hasta el final del archivo de sintaxis de comandos.

El botón Ejecutar de la barra de herramientas del Editor de sintaxis ejecuta los comandos seleccionados o el comando donde se encuentra el cursor si no hay nada seleccionado.

Figura 12-3  
Botón Ejecutar



## ***Archivos de sintaxis Unicode***

En el modo Unicode, el formato por defecto para guardar archivos de sintaxis de comandos creados o modificados durante la sesión, también es Unicode (UTF-8). Los archivos de sintaxis de comandos con formato Unicode no se pueden leer en versiones de SPSS anteriores a la 16.0. Si desea obtener más información sobre el modo Unicode, consulte [General: Opciones en p. 479](#).

Para guardar un archivo de sintaxis en un formato compatible con las versiones anteriores:

- ▶ En los menús de la ventana de sintaxis, seleccione:
  - Archivo
  - Guardar como

- ▶ En el cuadro de diálogo Guardar como, seleccione Codificación local de la lista desplegable Codificación. La codificación local está determinada por la configuración regional actual.

## **Varios comandos Ejecutar**

La sintaxis pegada desde cuadros de diálogo o copiada desde el registro o el diario puede contener comandos EXECUTE. Cuando se ejecutan comandos desde una ventana de sintaxis, los comandos EXECUTE suelen ser innecesarios y ralentizar el rendimiento, en especial con archivos de datos de más tamaño, ya que cada comando EXECUTE lee todo el archivo de datos. Si desea obtener más información, consulte el comando EXECUTE en la referencia de sintaxis de comandos (*Command Syntax Reference*) (disponible en el menú Ayuda de cualquier ventana de SPSS ).

### **Funciones de retardo**

Una excepción importante son los comandos de transformación que contienen funciones de retardo. En una serie de comandos de transformación sin intervención de comandos EXECUTE, ni ningún otro comando que lea datos, las funciones de retardo se calculan después de las restantes transformaciones, con independencia del orden de los comandos. Por ejemplo:

```
COMPUTE lagvar=LAG(var1). COMPUTE var1=var1*2.
```

y

```
COMPUTE lagvar=LAG(var1). EXECUTE. COMPUTE var1=var1*2.
```

ofrece resultados muy diferentes para el valor de *lagvar* dado que el anterior utiliza el valor transformado de *var1* mientras que el último utiliza el valor original.

# Frecuencias

El procedimiento Frecuencias proporciona estadísticos y representaciones gráficas que resultan útiles para describir muchos tipos de variables. El procedimiento Frecuencias es un comienzo para empezar a consultar los datos.

Para los informes de frecuencias y los gráficos de barras, puede organizar los diferentes valores en orden ascendente o descendente u ordenar las categorías por sus frecuencias. Es posible suprimir el informe de frecuencias cuando una variable posee muchos valores diferentes. Puede etiquetar los gráficos con las frecuencias (la opción por defecto) o con los porcentajes.

**Ejemplo.** ¿Cuál es la distribución de los clientes de una empresa por tipo de industria? En los resultados podría observar que el 37,5% de sus clientes pertenece a agencias gubernamentales, el 24,9% a corporaciones, el 28,1% a instituciones académicas, y el 9,4% a la industria sanitaria. Con respecto a los datos continuos, cuantitativos, como los ingresos por ventas, podría comprobar que el promedio de ventas de productos es de 3.576 dólares con una desviación típica de 1.078 dólares.

**Estadísticos y gráficos.** Frecuencias, porcentajes, porcentajes acumulados, media, mediana, moda, suma, desviación típica, varianza, amplitud, valores mínimo y máximo, error típico de la media, asimetría y curtosis (ambos con sus errores típicos), cuartiles, percentiles especificados por el usuario, gráficos de barras, gráficos de sectores e histogramas.

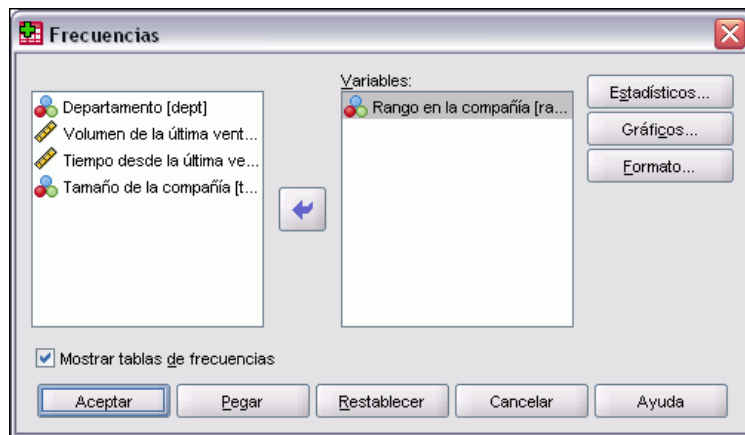
**Datos.** Utilice códigos numéricos o cadenas para codificar las variables categóricas (medidas de nivel nominal u ordinal).

**Supuestos.** Las tabulaciones y los porcentajes proporcionan una descripción útil para los datos de cualquier distribución, especialmente para las variables con categorías ordenadas o desordenadas. Muchos de los estadísticos de resumen optativos, tales como la media y la desviación típica, se basan en la teoría normal y son apropiados para las variables cuantitativas con distribuciones simétricas. Los estadísticos robustos, tales como la mediana, los cuartiles y los percentiles son apropiados para las variables cuantitativas que pueden o no cumplir el supuesto de normalidad.

## **Para obtener tablas de frecuencias**

- ▶ Elija en los menús:
  - Analizar
  - Estadísticos descriptivos
  - Frecuencias...

Figura 13-1  
Cuadro de diálogo principal Frecuencias



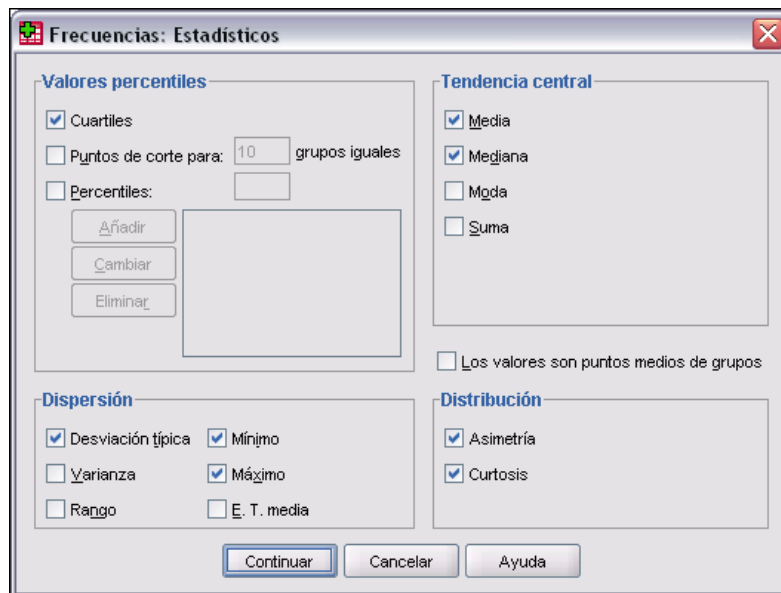
- ▶ Seleccione una o más variables categóricas o cuantitativas.

Si lo desea, puede:

- Pulsar en Estadísticos para obtener estadísticos descriptivos para las variables cuantitativas.
- Pulsar en Gráficos para obtener gráficos de barras, gráficos de sectores e histogramas.
- Pulsar en Formato para determinar el orden en el que se muestran los resultados.

## ***Frecuencias: Estadísticos***

Figura 13-2  
Cuadro de diálogo Frecuencias: Estadísticos



**Valores percentiles.** Los valores de una variable cuantitativa que dividen los datos ordenados en grupos, de forma que un porcentaje de los casos se encuentre por encima y otro porcentaje se encuentre por debajo. Los cuartiles (los percentiles 25, 50 y 75) dividen las observaciones en cuatro grupos de igual tamaño. Si desea un número igual de grupos que no sea cuatro, seleccione Puntos de corte para  $n$  grupos iguales. También puede especificar percentiles individuales (por ejemplo, el percentil 95, el valor por debajo del cual se encuentran el 95% de las observaciones).

**Tendencia central.** Los estadísticos que describen la localización de la distribución, incluyen: Media, Mediana, Moda y Suma de todos los valores.

- **Media.** Una medida de tendencia central. El promedio aritmético; la suma dividida por el número de casos.
- **Mediana.** Valor por encima y por debajo del cual se encuentran la mitad de los casos; el percentil 50. Si hay un número par de casos, la mediana es la media de los dos valores centrales, cuando los casos se ordenan en orden ascendente o descendente. La mediana es una medida de tendencia central que no es sensible a los valores atípicos (a diferencia de la media, que puede resultar afectada por unos pocos valores extremadamente altos o bajos).
- **Moda.** El valor que ocurre con mayor frecuencia. Si varios valores comparten la mayor frecuencia de aparición, cada una de ellas es una moda. El procedimiento de frecuencias devuelve sólo la más pequeña de esas modas múltiples.
- **Suma.** La suma o total de todos los valores, a lo largo de todos los casos que no tengan valores perdidos.

**Dispersión.** Los estadísticos que miden la cantidad de variación o de dispersión en los datos, incluyen: Desviación típica, Varianza, Rango, Mínimo, Máximo y Error típico de la media.

- **Desviación típica.** Medida de dispersión en torno a la media. En una distribución normal, el 68% de los casos se encuentra dentro de una desviación típica de la media y el 95% queda entre dos desviaciones típicas. Por ejemplo, si la media de edad es 45, con una desviación típica de 10, el 95% de los casos estaría entre 25 y 65 en una distribución normal.
- **Varianza.** Medida de dispersión en torno a la media, igual a la suma de los cuadrados de las desviaciones respecto a la media dividida por el número de casos menos 1. La varianza se mide en unas unidades que son el cuadrado de las de la propia variable.
- **Amplitud.** Diferencia entre los valores mayor y menor de una variable numérica; el máximo menos el mínimo. También se denomina el recorrido de la variable.
- **Mínimo.** Valor más pequeño de una variable numérica.
- **Máximo.** Valor más grande de una variable numérica.
- **E. T. media.** Medida de cuánto puede variar el valor de la media de una muestra a otra, extraídas éstas de la misma distribución. Puede utilizarse para comparar de forma aproximada la media observada con un valor hipotetizado (es decir, podremos concluir que dos valores son distintos si la razón de la diferencia respecto al error típico es menor que  $-2$  o mayor que  $+2$ ).

**Distribución.** Asimetría y curtosis son estadísticos que describen la forma y la simetría de la distribución. Estos estadísticos se muestran con sus errores típicos.

- **Asimetría.** Medida de la asimetría de una distribución. Las distribuciones normales son simétricas, por lo que tienen un valor de asimetría 0. Una distribución que tenga una asimetría significativa positiva tiene una cola derecha larga. Una distribución que tenga una asimetría negativa significativa tiene una cola izquierda larga. Como regla aproximada, un valor de

la asimetría mayor que el doble de su error típico se asume que indica una desviación de la simetría.

- **Curtosis.** Es una medida del grado en que las observaciones se agrupan en torno a un punto central. En una distribución normal, el valor del estadístico de curtosis es 0. Una curtosis positiva indica que las observaciones se concentran más y presentan colas más largas que las de una distribución normal. Una curtosis negativa indica que las observaciones se agrupan menos y presentan colas más cortas.

**Los valores son puntos medios de grupos.** Si los valores de los datos son puntos medios de grupos (por ejemplo, si las edades de todas las personas entre treinta y cuarenta años se codifican como 35), seleccione esta opción para estimar la mediana y los percentiles para los datos originales no agrupados.

## Frecuencias: Gráficos

Figura 13-3  
Cuadro de diálogo Frecuencias: Gráficos

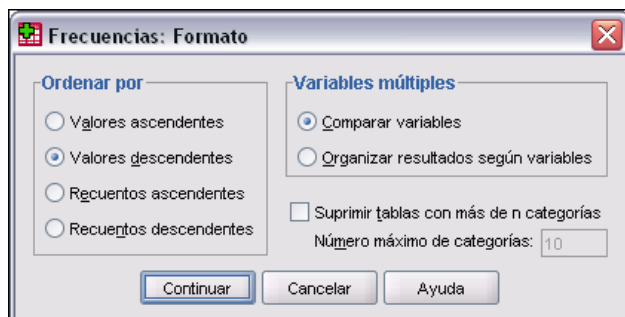


**Tipo de gráfico.** Los gráficos de sectores muestran la contribución de las partes a un todo. Cada sector de un gráfico de este tipo corresponde a un grupo, definido por una única variable de agrupación. Los gráficos de barras muestran la frecuencia de cada valor o categoría distinta como una barra diferente, permitiendo comparar las categorías de forma visual. Los histogramas también cuentan con barras, pero se representan a lo largo de una escala de intervalos iguales. La altura de cada barra es el recuento de los valores que están dentro del intervalo para una variable cuantitativa. Los histogramas muestran la forma, el centro y la dispersión de la distribución. Una curva normal superpuesta en un histograma ayuda a juzgar si los datos están normalmente distribuidos.

**Valores del gráfico.** Para los gráficos de barras, puede etiquetar el eje de escala con las frecuencias o los porcentajes.

## Frecuencias: Formato

Figura 13-4  
Cuadro de diálogo Frecuencias: Formato



**Ordenar por.** La tabla de frecuencias se puede organizar respecto a los valores actuales de los datos o respecto al recuento (frecuencia de aparición) de esos valores y la tabla puede organizarse en orden ascendente o descendente. Sin embargo, si solicita un histograma o percentiles, Frecuencias asumirá que la variable es cuantitativa y mostrará sus valores en orden ascendente.

**Múltiples variables.** Si desea generar tablas de estadísticos para múltiples variables, podrá mostrar todas las variables en una sola tabla (Comparar variables), o bien mostrar una tabla de estadísticos independiente para cada variable (Organizar resultados según variables).

**Suprimir tablas con más de n categorías.** Esta opción impide que se muestren tablas que contengan más valores que el número especificado.

# Descriptivos

El procedimiento Descriptivos muestra estadísticos de resumen univariados para varias variables en una única tabla y calcula valores tipificados (puntuaciones  $z$ ). Las variables se pueden ordenar por el tamaño de sus medias (en orden ascendente o descendente), alfabéticamente o por el orden en el que se seleccionen las variables (el valor por defecto).

Cuando se guardan las puntuaciones  $z$ , éstas se añaden a los datos del Editor de datos y quedan disponibles para los gráficos, el listado de los datos y los análisis. Cuando las variables se registran en unidades diferentes (por ejemplo, producto interior bruto per cápita y porcentaje de alfabetización), una transformación de puntuación  $z$  pondrá las variables en una escala común para poder compararlas visualmente con más facilidad.

**Ejemplo.** Si cada caso de los datos contiene los totales de ventas diarias de cada vendedor (por ejemplo, una entrada para Bob, una para Kim y una para Brian) recogidas cada día durante varios meses, el procedimiento Descriptivos puede calcular la media diaria de ventas para cada vendedor y ordenar los resultados del promedio de ventas de mayor a menor.

**Estadísticos.** Tamaño de muestra, media, mínimo, máximo, desviación típica, varianza, rango, suma, error típico de la media, curtosis y asimetría con sus errores típicos.

**Datos.** Utilice variables numéricas después de haberlas inspeccionado gráficamente para registrar errores, valores atípicos y anomalías de distribución. El procedimiento Descriptivos es muy eficaz para archivos grandes (de miles de casos).

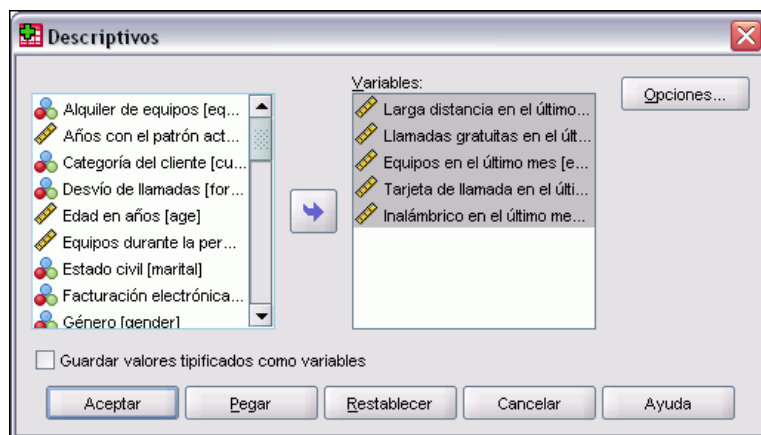
**Supuestos.** La mayoría de los estadísticos disponibles (incluyendo las puntuaciones  $z$ ) se basan en la teoría normal y son adecuados para variables cuantitativas (medidas a nivel de razón o de intervalo) con distribuciones simétricas. Se deben evitar las variables con categorías no ordenadas o distribuciones asimétricas. La distribución de puntuaciones  $z$  tiene la misma forma que la de los datos originales; por tanto, el cálculo de puntuaciones  $z$  no es una solución para los datos con problemas.

## ***Para obtener estadísticos descriptivos***

- ▶ Elija en los menús:
  - Analizar
  - Estadísticos descriptivos
  - Descriptivos...



Figura 14-1  
Cuadro de diálogo Descriptivos



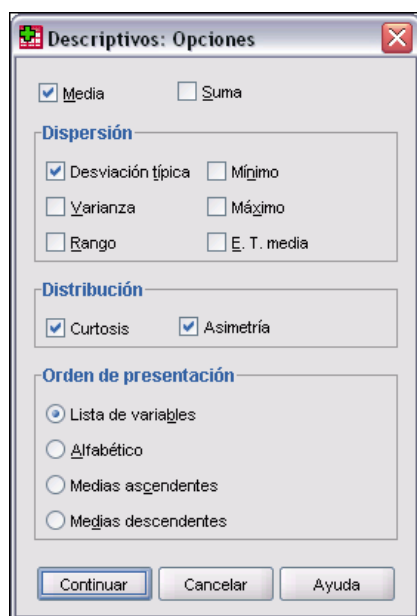
- Seleccione una o más variables.

Si lo desea, puede:

- Seleccionar Guardar valores tipificados como variables para guardar las puntuaciones  $z$  como nuevas variables.
- Pulsar en Opciones para seleccionar estadísticos opcionales y el orden de presentación.

## Descriptivos: Opciones

Figura 14-2  
Cuadro de diálogo Descriptivos: Opciones



**Media y suma.** Se muestra por defecto la media o promedio aritmético.

**Dispersión.** Los estadísticos que miden la dispersión o variación en los datos incluyen la desviación típica, la varianza, el rango, el mínimo, el máximo y el error típico de la media.

- **Desviación típica.** Medida de dispersión en torno a la media. En una distribución normal, el 68% de los casos se encuentra dentro de una desviación típica de la media y el 95% queda entre dos desviaciones típicas. Por ejemplo, si la media de edad es 45, con una desviación típica de 10, el 95% de los casos estaría entre 25 y 65 en una distribución normal.
- **Varianza.** Medida de dispersión en torno a la media, igual a la suma de los cuadrados de las desviaciones respecto a la media dividida por el número de casos menos 1. La varianza se mide en unas unidades que son el cuadrado de las de la propia variable.
- **Amplitud.** Diferencia entre los valores mayor y menor de una variable numérica; el máximo menos el mínimo. También se denomina el recorrido de la variable.
- **Mínimo.** Valor más pequeño de una variable numérica.
- **Máximo.** Valor más grande de una variable numérica.
- **E. T. media.** Medida de cuánto puede variar el valor de la media de una muestra a otra, extraídas éstas de la misma distribución. Puede utilizarse para comparar de forma aproximada la media observada con un valor hipotetizado (es decir, podremos concluir que dos valores son distintos si la razón de la diferencia respecto al error típico es menor que  $-2$  o mayor que  $+2$ ).

**Distribución.** La curtosis y la asimetría son los estadísticos que caracterizan la forma y simetría de la distribución. Estos estadísticos se muestran con sus errores típicos.

- **Curtosis.** Es una medida del grado en que las observaciones se agrupan en torno a un punto central. En una distribución normal, el valor del estadístico de curtosis es 0. Una curtosis positiva indica que las observaciones se concentran más y presentan colas más largas que las de una distribución normal. Una curtosis negativa indica que las observaciones se agrupan menos y presentan colas más cortas.
- **Asimetría.** Medida de la asimetría de una distribución. Las distribuciones normales son simétricas, por lo que tienen un valor de asimetría 0. Una distribución que tenga una asimetría significativa positiva tiene una cola derecha larga. Una distribución que tenga una asimetría negativa significativa tiene una cola izquierda larga. Como regla aproximada, un valor de la asimetría mayor que el doble de su error típico se asume que indica una desviación de la simetría.

**Orden de presentación.** Por defecto, las variables se muestran en el orden en que se hayan seleccionado. Si lo desea, se pueden mostrar las variables alfabéticamente, por medias ascendentes o por medias descendentes.

## ***Funciones adicionales del comando DESCRIPTIVES***

Con el lenguaje de sintaxis de comandos también podrá:

- Guardar puntuaciones tipificadas (puntuaciones  $z$ ) para algunas variables, pero no para todas (con el subcomando VARIABLES).
- Especificar nombres para las variables nuevas que contienen puntuaciones tipificadas (mediante el subcomando VARIABLES).

- Excluir del análisis casos con valores perdidos para cualquier variable (mediante el subcomando `MISSING`).
- Ordenar las variables de la presentación por el valor de cualquier estadístico, no sólo por la media (mediante el subcomando `SORT`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

## ***Explorar***

El procedimiento Explorar genera estadísticos de resumen y representaciones gráficas, bien para todos los casos o bien de forma separada para grupos de casos. Existen numerosas razones para utilizar este procedimiento: para inspeccionar los datos, identificar valores atípicos, obtener descripciones, comprobar supuestos y caracterizar diferencias entre subpoblaciones (grupos de casos). La inspección de los datos puede mostrar que existen valores inusuales, valores extremos, discontinuidades en los datos u otras peculiaridades. La exploración de los datos puede ayudar a determinar si son adecuadas las técnicas estadísticas que está teniendo en consideración para el análisis de los datos. La exploración puede indicar que necesita transformar los datos si la técnica necesita una distribución normal. O bien, el usuario puede decidir que necesita utilizar pruebas no paramétricas.

**Ejemplo.** Observe la distribución de los tiempos de aprendizaje de laberintos de una serie de ratas sometidas a cuatro programas de refuerzo diferentes. Para cada uno de los cuatro grupos, se puede observar si la distribución de tiempos es aproximadamente normal y si las cuatro varianzas son iguales. También se pueden identificar los casos con los cinco valores de tiempo mayores y los cinco menores. Los diagramas de caja y los gráficos de tallo y hojas resumen gráficamente la distribución del tiempo de aprendizaje de cada uno de los grupos.

**Estadísticos y gráficos.** Media, mediana, media recortada al 5%, error típico, varianza, desviación típica, mínimo, máximo, amplitud, amplitud intercuartil, asimetría y curtosis y sus errores típicos, intervalo de confianza para la media (y el nivel de confianza especificado), percentiles, estimador-M de Huber, estimador en onda de Andrews, estimador-M redescendente de Hampel, estimador bponderado de Tukey, cinco valores mayores y cinco menores, estadístico de Kolmogorov-Smirnov con el nivel de significación de Lilliefors para contrastar la normalidad y estadístico de Shapiro-Wilk. Diagramas de caja, gráficos de tallo y hojas, histogramas, diagramas de normalidad y diagramas de dispersión por nivel con pruebas de Levene y transformaciones.

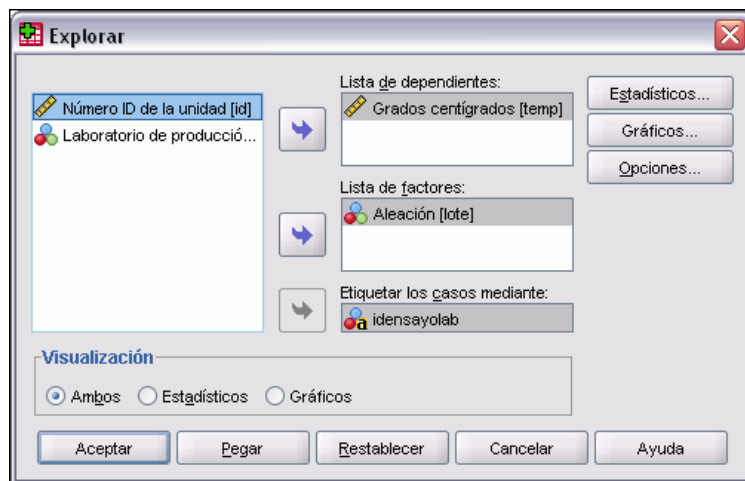
**Datos.** El procedimiento Explorar se puede utilizar para las variables cuantitativas (nivel de medida de razón o de intervalo). Una variable de factor (utilizada para dividir los datos en grupos de casos) debe tener un número razonable de valores distintivos (categorías). Estos valores pueden ser de cadena corta o numéricos. La variable de etiquetas de caso, utilizada para etiquetar valores atípicos en los diagramas de caja, puede ser de cadena corta, de cadena larga (los 15 primeros bytes) o numérica.

**Supuestos.** La distribución de los datos no tiene que ser simétrica ni normal.

**Para explorar los datos**

- ▶ Elija en los menús:
  - Analizar
  - Estadísticos descriptivos
  - Explorar...

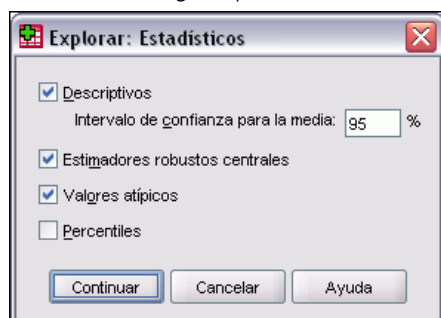
Figura 15-1  
Cuadro de diálogo Explorar



- ▶ Seleccione una o más variables dependientes.
  - Si lo desea, puede:
    - Seleccionar una o más variables de factor, cuyos valores definirán grupos de casos.
    - Seleccionar una variable de identificación para etiquetar los casos.
    - Pulse en Estadísticos para obtener estimadores robustos, valores atípicos, percentiles y tablas de frecuencias.
    - Pulse en Gráficos para obtener histogramas, pruebas y gráficos de probabilidad normal y diagramas de dispersión por nivel con estadísticos de Levene.
    - Pulse en Opciones para manipular los valores perdidos.

**Explorar: Estadísticos**

Figura 15-2  
Cuadro de diálogo Explorar: Estadísticos



**Descriptivos.** Por defecto se muestran estas medidas de dispersión y de tendencia central. Éstas últimas indican la localización de la distribución, e incluyen la media, la mediana y la media recortada al 5%. Las medidas de dispersión muestran la disimilaridad de los valores, incluyen: los errores típicos, la varianza, la desviación típica, el mínimo, el máximo, la amplitud y la amplitud intercuartil. Los estadísticos descriptivos también incluyen medidas de la forma de la distribución: la asimetría y la curtosis se muestran con sus errores típicos. También se muestra el intervalo de confianza a un nivel del 95%; aunque se puede especificar otro nivel.

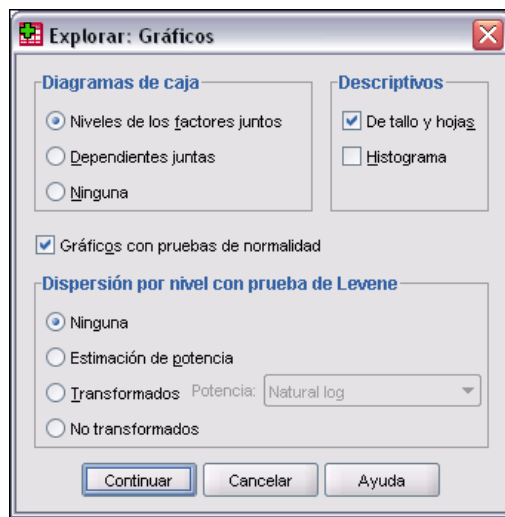
**Estimadores robustos centrales.** Alternativas robustas a la mediana y a la media muestral para estimar la localización. Los estimadores calculados se diferencian por las ponderaciones que aplican a los casos. Se muestran los siguientes: el estimador-M de Huber, el estimador en onda de Andrew, el estimador-M redescendente de Hampel y el estimador bponderado de Tukey.

**Valores atípicos.** Muestra los cinco valores mayores y los cinco menores con las etiquetas de caso.

**Percentiles.** Muestra los valores de los percentiles 5, 10, 25, 50, 75, 90 y 95.

## Explorar: Gráficos

Figura 15-3  
Cuadro de diálogo Explorar: Gráficos



**Diagramas de caja.** Estas alternativas controlan la presentación de los diagramas de caja cuando existe más de una variable dependiente. Niveles de los factores juntos genera una presentación para cada variable dependiente. En cada una se muestran diagramas de caja para cada uno de los grupos definidos por una variable de factor. Dependientes juntas genera una presentación para cada grupo definido por una variable de factor. En cada una se muestran juntos los diagramas de caja de cada variable dependiente. Esta disposición es particularmente útil cuando las variables representan una misma característica medida en momentos distintos.

**Descriptivos.** La sección Descriptivos permite seleccionar gráficos de tallo y hojas e histogramas.

**Gráficos con pruebas de normalidad.** Muestra los diagramas de probabilidad normal y de probabilidad sin tendencia. Se muestra el estadístico de Kolmogorov-Smirnov con un nivel de significación de Lilliefors para contrastar la normalidad. Si se especifican ponderaciones no

enteras, se calculará el estadístico de Shapiro-Wilk cuando el tamaño de la muestra ponderada esté entre 3 y 50. Si no hay ponderaciones o éstas son enteras, se calculará el estadístico cuando el tamaño muestral esté entre 3 y 5.000.

**Dispersión por nivel con prueba de Levene.** Controla la transformación de los datos para los diagramas de dispersión por nivel. Para todos los diagramas de dispersión por nivel se muestra la pendiente de la línea de regresión y las pruebas robustas de Levene sobre la homogeneidad de varianzas. Si selecciona una transformación, las pruebas de Levene se basarán en los datos transformados. Si no selecciona ninguna variable de factor, no se generará ningún diagrama de dispersión por nivel. Estimación de potencia produce un gráfico de los logaritmos naturales de las amplitudes intercuartiles respecto a los logaritmos naturales de las medianas de todas las casillas, así como una estimación de la transformación de potencia necesaria para conseguir varianzas iguales en las casillas. Un diagrama de dispersión por nivel ayuda a determinar la potencia que precisa una transformación para estabilizar (igualar) las varianzas de los grupos. Transformados permite seleccionar una de las alternativas de potencia, quizás siguiendo las recomendaciones de la estimación de potencia, y genera gráficos de los datos transformados. Se trazan la amplitud intercuartil y la mediana de los datos transformados. No transformados genera gráficos de los datos brutos. Es equivalente a una transformación con una potencia de 1.

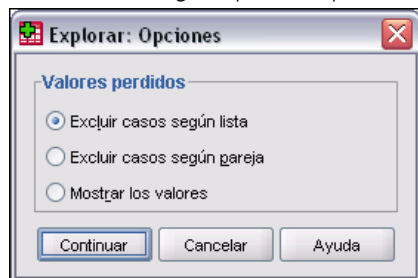
### ***Explorar: Transformaciones de potencia***

A continuación aparecen las transformaciones de potencia para los diagramas de dispersión por nivel. Para transformar los datos, deberá seleccionar una potencia para la transformación. Puede elegir una de las siguientes alternativas:

- **Log natural.** Transformación de logaritmo natural. Este es el método por defecto.
- **1/raíz cuadrada.** Para cada valor de los datos se calcula el inverso de la raíz cuadrada.
- **Recíproco.** Se calcula el inverso de cada valor de los datos.
- **Raíz cuadrada.** Se calcula la raíz cuadrada de cada valor de los datos.
- **Cuadrado.** Se calcula el cuadrado de cada valor de los datos.
- **Cubo.** Se calcula el cubo de cada valor de los datos.

### ***Explorar: Opciones***

Figura 15-4  
Cuadro de diálogo Explorar: Opciones



**Valores perdidos.** Controla el tratamiento de los valores perdidos.

- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable de factor o variable dependiente se excluyen de todos los análisis. Este es el método por defecto.
- **Excluir casos según pareja.** Los casos que no tengan valores perdidos para las variables de un grupo (casilla) se incluyen en el análisis de ese grupo. El caso puede tener valores perdidos para las variables utilizadas en otros grupos.
- **Mostrar los valores.** Los valores perdidos para las variables de factor se tratan como una categoría diferente. Todos los resultados se generan para esta categoría adicional. Las tablas de frecuencias incluyen categorías para los valores perdidos. Los valores perdidos para una variable de factor se incluyen pero se etiquetan como perdidos.

## ***Funciones adicionales del comando EXAMINE***

El procedimiento Explorar utiliza la sintaxis de comandos EXAMINE. Con el lenguaje de sintaxis de comandos también podrá:

- Solicitar los gráficos y resultados totales además de los gráficos y los resultados para los grupos definidos por las variables de factor (con el subcomando TOTAL).
- Especificar una escala común para un grupo de diagramas de caja (con el subcomando SCALE).
- Especificar interacciones de variables de factor (con el subcomando VARIABLES).
- Especificar percentiles distintos de los percentiles por defecto (con el subcomando PERCENTILES).
- Calcular percentiles respecto a cualquiera de los cinco métodos (con el subcomando PERCENTILES).
- Especificar una transformación de potencia para diagramas de dispersión por nivel (con el subcomando PLOT).
- Especificar el número de valores extremos que se van a mostrar (mediante el subcomando STATISTICS).
- Especificar parámetros para los estimadores robustos centrales, los estimadores robustos de ubicación (mediante el subcomando MESTIMATORS).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).



## ***Tablas de contingencia***

El procedimiento Tablas de contingencia crea tablas de clasificación doble y múltiple y, además, proporciona una serie de pruebas y medidas de asociación para las tablas de doble clasificación. La estructura de la tabla y el hecho de que las categorías estén ordenadas o no determinan las pruebas o medidas que se utilizaban.

Los estadísticos de tablas de contingencia y las medidas de asociación sólo se calculan para las tablas de doble clasificación. Si especifica una fila, una columna y un factor de capa (variable de control), el procedimiento Tablas de contingencia crea un panel de medidas y estadísticos asociados para cada valor del factor de capa (o una combinación de valores para dos o más variables de control). Por ejemplo, si *sexo* es un factor de capa para una tabla de *casado* (sí, no) en función de *vida* (vida emocionante, rutinaria o aburrida), los resultados para una tabla de doble clasificación para las mujeres se calculan de forma independiente de los resultados de los hombres y se imprimen en paneles uno detrás del otro.

**Ejemplo.** ¿Es más probable que los clientes de las empresas pequeñas sean más rentables en la venta de servicios (por ejemplo, formación y asesoramiento) que los clientes de las empresas grandes? A partir de una tabla de contingencia podría deducir que la prestación de servicios a la mayoría de las empresas pequeñas (con menos de 500 empleados) produce considerables beneficios, mientras que con la mayoría de las empresas de gran tamaño (con más de 2.500 empleados), los beneficios obtenidos son mucho menores.

**Estadísticos y medidas de asociación.** Chi-cuadrado de Pearson, chi-cuadrado de la razón de verosimilitud, prueba de asociación lineal por lineal, prueba exacta de Fisher, chi-cuadrado corregido de Yates,  $r$  de Pearson, rho de Spearman, coeficiente de contingencia,  $\phi$ ,  $V$  de Cramér, lambdas simétricas y asimétricas, tau de Kruskal y Goodman, coeficiente de incertidumbre, gamma,  $d$ , de Somers, tau- $b$  de Kendall, tau- $c$  de Kendall, coeficiente eta, kappa de Cohen, estimación de riesgo relativo, razón de ventajas, prueba de McNemar y estadísticos de Cochran y Mantel-Haenszel.

**Datos.** Para definir las categorías de cada variable, utilice valores de una variable numérica o de cadena (ocho bytes o menos). Por ejemplo, para *sexo*, podría codificar los datos como 1 y 2 o como *varón* y *mujer*.

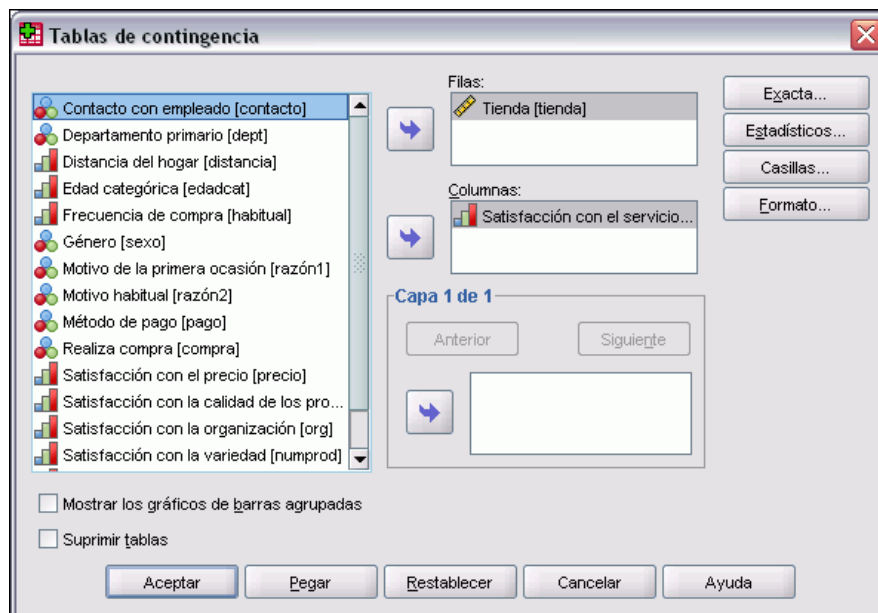
**Supuestos.** En algunos estadísticos y medidas se asume que hay unas categorías ordenadas (datos ordinales) o unos valores cuantitativos (datos de intervalos o de proporciones), como se explica en la sección sobre los estadísticos. Otros estadísticos son válidos cuando las variables de la tabla tienen categorías no ordenadas (datos nominales). Para los estadísticos basados en chi-cuadrado ( $\phi$ ,  $V$  de Cramér y coeficiente de contingencia), los datos deben ser una muestra aleatoria de una distribución multinomial.

*Nota:* Las variables ordinales pueden ser códigos numéricos que representen categorías (por ejemplo, 1 = *bajo*, 2 = *medio*, 3 = *alto*) o valores de cadena. Sin embargo, se supone que el orden alfabético de los valores de cadena indica el orden correcto de las categorías. Por ejemplo, en una variable de cadena cuyos valores sean *bajo*, *medio*, *alto*, se interpreta el orden de las categorías como *alto*, *bajo*, *medio* (orden que no es el correcto). Por norma general, se puede indicar que es más fiable utilizar códigos numéricos para representar datos ordinales.

### **Para obtener tablas de contingencia**

- ▶ Elija en los menús:
  - Analizar
  - Estadísticos descriptivos
  - Tablas de contingencia...

Figura 16-1  
Cuadro de diálogo *Tablas de contingencia*



- ▶ Seleccione una o más variables de fila y una o más variables de columna. Si lo desea, puede:
  - Seleccionar una o más variables de control.
  - Pulsar en Estadísticos para obtener pruebas y medidas de asociación para tablas o subtablas de doble clasificación.
  - Pulsar en Casillas para obtener porcentajes, residuos y valores esperados y observados.
  - Pulsar en Formato para controlar el orden de las categorías.

## Capas de las tablas de contingencia

Si se seleccionan una o más variables de capas, se generará una tabla de contingencia por cada categoría de cada variable de capas (variable de control). Por ejemplo, si emplea una variable de fila, una variable de columna y una variable de capas con dos categorías, obtendrá una tabla de doble clasificación por cada categoría de la variable de capas. Para crear otra capa de variables de control, pulse en Siguiente. Se crean subtablas para cada combinación de categorías para cada variable de la 1ª capa, cada variable de la 2ª capa, y así sucesivamente. Si se solicitan estadísticos y medidas de asociación, se aplicarán sólo a las tablas de doble clasificación.

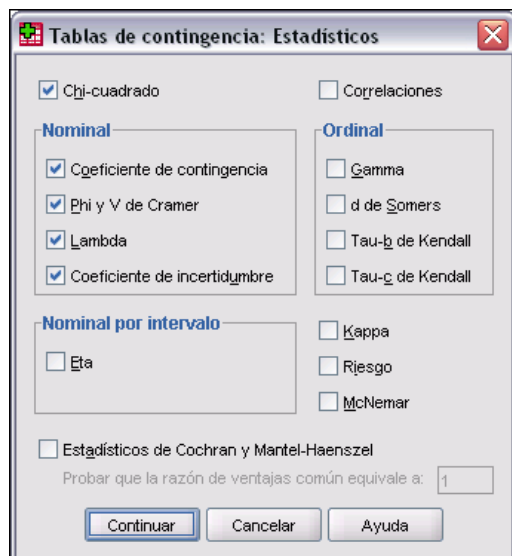
## Gráficos de barras agrupadas de las tablas de contingencia

**Mostrar los gráficos de barras agrupadas.** Los gráficos de barras agrupadas ayudan a resumir los datos por grupos de casos. Hay una agrupación de barras por cada valor de la variable especificada en el cuadro Filas. La variable que define las barras dentro de cada agrupación es la variable especificada en el cuadro Columnas. Por cada valor de esta variable hay un conjunto de barras de distinto color o trama. Si especifica más de una variable en Columnas o en Filas, se generará un gráfico de barras agrupadas por cada combinación de dos variables.

## Tablas de contingencia: Estadísticos

Figura 16-2

Cuadro de diálogo Tablas de contingencia: Estadísticos



**Chi-cuadrado.** Para las tablas con dos filas y dos columnas, seleccione Chi-cuadrado para calcular el chi-cuadrado de Pearson, el chi-cuadrado de la razón de verosimilitud, la prueba exacta de Fisher y el chi-cuadrado corregido de Yates (corrección por continuidad). Para las tablas  $2 \times 2$ , se calcula la prueba exacta de Fisher cuando una tabla (que no resulte de perder columnas o filas en una tabla mayor) presente una casilla con una frecuencia esperada menor que 5. Para las restantes tablas  $2 \times 2$  se calcula el chi-cuadrado corregido de Yates. Para las tablas con cualquier

número de filas y columnas, seleccione Chi-cuadrado para calcular el chi-cuadrado de Pearson y el chi-cuadrado de la razón de verosimilitud. Cuando ambas variables de tabla son cuantitativas, Chi-cuadrado da como resultado la prueba de asociación lineal por lineal.

**Correlaciones.** Para las tablas en las que tanto las columnas como las filas contienen valores ordenados, Correlaciones da como resultado rho, el coeficiente de correlación de Spearman (sólo datos numéricos). La rho de Spearman es una medida de asociación entre órdenes de rangos. Cuando ambas variables de tabla (factores) son cuantitativas, Correlaciones da como resultado  $r$ , el coeficiente de correlación de Pearson, una medida de asociación lineal entre las variables.

**Nominal.** Para los datos nominales (sin orden intrínseco, como católico, protestante o judío), puede seleccionar el Coeficiente de contingencia, coeficiente Phi y V de Cramér, Lambda (lambdas simétricas y asimétricas y tau de Kruskal y Goodman) y el Coeficiente de incertidumbre.

- **Coeficiente de contingencia.** Medida de asociación basada en chi-cuadrado. El valor siempre está comprendido entre 0 y 1. El valor 0 indica que no hay asociación entre la fila y la columna. Los valores cercanos a 1 indican que hay gran relación entre las variables. El valor máximo posible depende del número de filas y columnas de la tabla.
- **Phi y V de Cramer.** Phi es una medida de asociación basada en chi-cuadrado que conlleva dividir el estadístico chi-cuadrado entre el tamaño de la muestra y calcular la raíz cuadrada del resultado. V de Cramer es una medida de asociación basada en chi-cuadrado.
- **Lambda.** Medida de asociación que refleja la reducción proporcional en el error cuando se utilizan los valores de la variable independiente para pronosticar los valores de la variable dependiente. Un valor igual a 1 significa que la variable independiente pronostica perfectamente la variable dependiente. Un valor igual a 0 significa que la variable independiente no ayuda en absoluto a pronosticar la variable dependiente.
- **Coeficiente de incertidumbre.** Medida de asociación que indica la reducción proporcional del error cuando los valores de una variable se emplean para pronosticar valores de la otra variable. Por ejemplo, un valor de 0.83 indica que el conocimiento de una variable reduce en un 83% el error al pronosticar los valores de la otra variable. SPSS calcula tanto la versión simétrica como la asimétrica del coeficiente de incertidumbre.

**Ordinal.** Para las tablas en las que tanto las filas como las columnas contienen valores ordenados, seleccione Gamma (orden cero para tablas de doble clasificación y condicional para tablas cuyo factor de clasificación va de 3 a 10), Tau-b de Kendall y Tau-c de Kendall. Para pronosticar las categorías de columna de las categorías de fila, seleccione d de Somers.

- **Gamma.** Medida de asociación simétrica entre dos variables ordinales cuyo valor siempre está comprendido entre menos -1 y 1. Los valores próximos a 1, en valor absoluto, indican una fuerte relación entre las dos variables. Los valores próximos a cero indican que hay poca o ninguna relación entre las dos variables. Para las tablas de doble clasificación, se muestran las gammas de orden cero. Para las tablas de tres o más factores de clasificación, se muestran las gammas condicionales.
- **d de Somers.** Medida de asociación entre dos variables ordinales que toma un valor comprendido entre -1 y 1. Los valores próximos a 1, en valor absoluto, indican una fuerte relación entre las dos variables. Los valores próximos a cero indican que hay poca o ninguna relación entre las variables. La d de Somers es una extensión asimétrica de gamma que difiere sólo en la inclusión del número de pares no empatados en la variable independiente. También se calcula una versión simétrica de este estadístico.

- **Tau-b de Kendall.** Medida no paramétrica de la correlación para variables ordinales o de rangos que tiene en consideración los empates. El signo del coeficiente indica la dirección de la relación y su valor absoluto indica la magnitud de la misma, de tal modo que los mayores valores absolutos indican relaciones más fuertes. Los valores posibles van de -1 a 1, pero un valor de -1 o +1 sólo se puede obtener a partir de tablas cuadradas.
- **Tau-c de Kendall.** Medida no paramétrica de asociación para variables ordinales que ignora los empates. El signo del coeficiente indica la dirección de la relación y su valor absoluto indica la magnitud de la misma, de tal modo que los mayores valores absolutos indican relaciones más fuertes. Los valores posibles van de -1 a 1, pero un valor de -1 o +1 sólo se puede obtener a partir de tablas cuadradas.

**Nominal por intervalo.** Cuando una variable es categórica y la otra es cuantitativa, seleccione Eta. La variable categórica debe codificarse numéricamente.

- **Eta.** Medida de asociación cuyo valor siempre está comprendido entre 0 y 1. El valor 0 indica que no hay asociación entre las variables de fila y de columna. Los valores cercanos a 1 indican que hay gran relación entre las variables. Eta resulta apropiada para una variable dependiente medida en una escala de intervalo (por ejemplo, ingresos) y una variable independiente con un número limitado de categorías (por ejemplo, sexo). Se calculan dos valores de eta: uno trata la variable de las filas como una variable de intervalo; el otro trata la variable de las columnas como una variable de intervalo.

**Kappa.** La kappa de Cohen mide el acuerdo entre las evaluaciones de dos jueces cuando ambos están valorando el mismo objeto. Un valor igual a 1 indica un acuerdo perfecto. Un valor igual a 0 indica que el acuerdo no es mejor que el que se obtendría por azar. Kappa sólo está disponible para las tablas cuadradas (tablas en las que ambas variables tienen el mismo número de categorías).

**Riesgo.** Para tablas 2x2, una medida del grado de asociación entre la presencia de un factor y la ocurrencia de un evento. Si el intervalo de confianza para el estadístico incluye un valor de 1, no se podrá asumir que el factor está asociado con el evento. Cuando la ocurrencia del factor es rara, se puede utilizar la razón de las ventajas (odds ratio) como estimación del riesgo relativo.

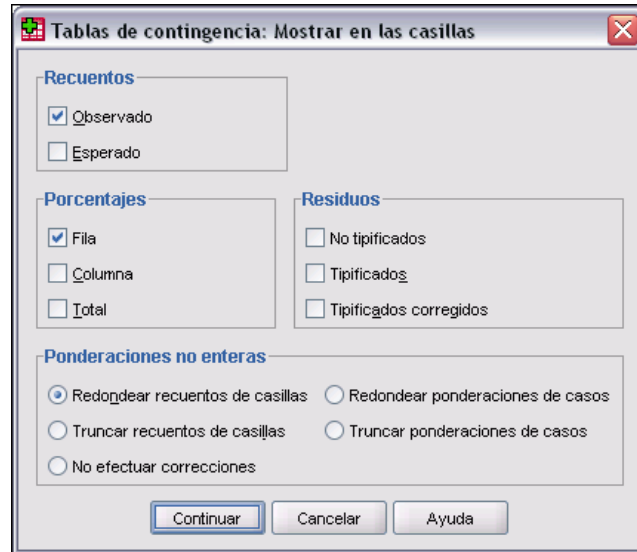
**McNemar.** Prueba no paramétrica para dos variables dicotómicas relacionadas. Contrasta los cambios en las respuestas utilizando la distribución de chi-cuadrado. Es útil para detectar cambios en las respuestas debidas a la intervención experimental en los diseños del tipo "antes-después". Para las tablas cuadradas de mayor orden se informa de la prueba de simetría de McNemar-Bowker.

**Estadísticos de Cochran y de Mantel-Haenszel.** Estos estadísticos se pueden utilizar para contrastar la independencia entre una variable dicotómica de factor y una variable dicotómica de respuesta, condicionada por los patrones en las covariables, los cuales vienen definidos por la variable o variables de las capas (variables de control). Tenga en cuenta que mientras que otros estadísticos se calculan capa por capa, los estadísticos de Cochran y Mantel-Haenszel se calculan una sola vez para todas las capas.

## Tablas de contingencia: Mostrar en las casillas

Figura 16-3

Cuadro de diálogo Tablas de contingencia: Mostrar en las casillas



Para ayudarle a descubrir las tramas en los datos que contribuyen a una prueba de chi-cuadrado significativa, el procedimiento Tablas de contingencia muestra las frecuencias esperadas y tres tipos de residuos (desviaciones) que miden la diferencia entre las frecuencias observadas y las esperadas. Cada casilla de la tabla puede contener cualquier combinación de recuentos, porcentajes y residuos seleccionados.

**Recuentos.** El número de casos realmente observados y el número de casos esperados si las variables de fila y columna son independientes entre sí.

**Porcentajes.** Los porcentajes se pueden sumar a través de las filas o a lo largo de las columnas. También se encuentran disponibles los porcentajes del número total de casos representados en la tabla (una capa).

**Residuos.** Los residuos brutos no tipificados presentan la diferencia entre los valores observados y los esperados. También se encuentran disponibles los residuos tipificados y tipificados corregidos.

- **No tipificados.** La diferencia entre un valor observado y el valor pronosticado por el modelo. El valor pronosticado es el número de casos que se esperarían en la casilla si no hubiera relación entre las dos variables. Un residuo positivo indica que hay más casos en la casilla de los que habría en ella si las variables de fila y columna fueran independientes.
- **Tipificados.** El residuo dividido por una estimación de su error típico. Los residuos tipificados, que son conocidos también como los residuos de Pearson o residuos estandarizados, tienen una media de 0 y una desviación típica de 1.
- **Tipificados corregidos.** El residuo de una casilla (valor observado menos valor pronosticado) dividido por una estimación de su error típico. El residuo tipificado resultante viene expresado en unidades de desviación típica, por encima o por debajo de la media.

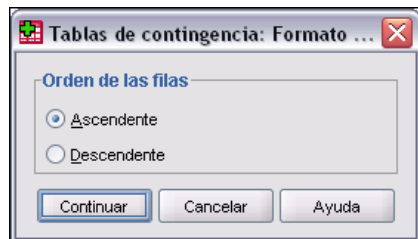
**Ponderaciones no enteras.** Los recuentos de las casillas suelen ser valores enteros, ya que representan el número de casos de cada casilla. Sin embargo, si el archivo de datos está ponderado en un momento determinado por una variable de ponderación con valores fraccionarios (por ejemplo, 1,25), los recuentos de las casillas pueden que también sean valores fraccionarios. Puede truncar o redondear estos valores antes o después de calcular los recuentos de las casillas o bien utilizar recuentos de casillas fraccionarios en la presentación de las tablas y los cálculos de los estadísticos.

- **Redondear frecuencias de casillas.** Las ponderaciones de los casos se utilizan tal cual, pero las ponderaciones acumuladas en las casillas se redondean antes de calcular cualquiera de los estadísticos.
- **Truncar frecuencias de casillas.** Las ponderaciones de los casos se utilizan tal cual, pero las ponderaciones acumuladas en las casillas se truncan antes de calcular cualquiera de los estadísticos.
- **Redondear ponderaciones de casos.** Se redondean las ponderaciones de los casos antes de utilizarlas.
- **Truncar ponderaciones de casos.** Se truncan las ponderaciones de los casos antes de utilizarlas.
- **Sin ajustes.** Las ponderaciones de los casos se utilizan tal cual y se utilizan las frecuencias de casilla fraccionales. Sin embargo, cuando se solicitan Estadísticos exactos (disponibles sólo con la opción Pruebas exactas), los pesos acumulados en las casillas se truncan o redondean antes de calcular los estadísticos de las Pruebas exactas.

## Tablas de contingencia: Formato de tabla

Figura 16-4

Cuadro de diálogo Tablas de contingencia: Formato de tabla



Puede ordenar las filas en orden ascendente o descendente de los valores de la variable de fila.

## Resumir

El procedimiento Resumir calcula estadísticos de subgrupo para las variables dentro de las categorías de una o más variables de agrupación. Se cruzan todos los niveles de las variables de agrupación. Puede elegir el orden en el que se mostrarán los estadísticos. También se muestran estadísticos de resumen para cada variable a través de todas las categorías. Los valores de los datos en cada categoría pueden mostrarse en una lista o suprimirse. Con grandes conjuntos de datos, tiene la opción de listar sólo los primeros  $n$  casos.

**Ejemplo.** ¿Cuál es la media de las ventas por regiones o por tipo de cliente? Podrá descubrir que el importe medio de las ventas es ligeramente superior en la región occidental respecto a las demás regiones, y que la media más alta se da entre los clientes de empresas privadas de la zona occidental.

**Estadísticos.** Suma, número de casos, media, mediana, mediana agrupada, error típico de la media, mínimo, máximo, rango, valor de la variable para la primera categoría de la variable de agrupación, valor de la variable para la última categoría de la variable de agrupación, desviación típica, varianza, curtosis, error típico de curtosis, asimetría, error típico de asimetría, porcentaje de la suma total, porcentaje del  $N$  total, porcentaje de la suma en, porcentaje de  $N$  en, media geométrica y media armónica.

**Datos.** Las variables de agrupación son variables categóricas cuyos valores pueden ser numéricos o de cadena. El número de categorías debe ser razonablemente pequeño. Las otras variables deben poder ordenarse mediante rangos.

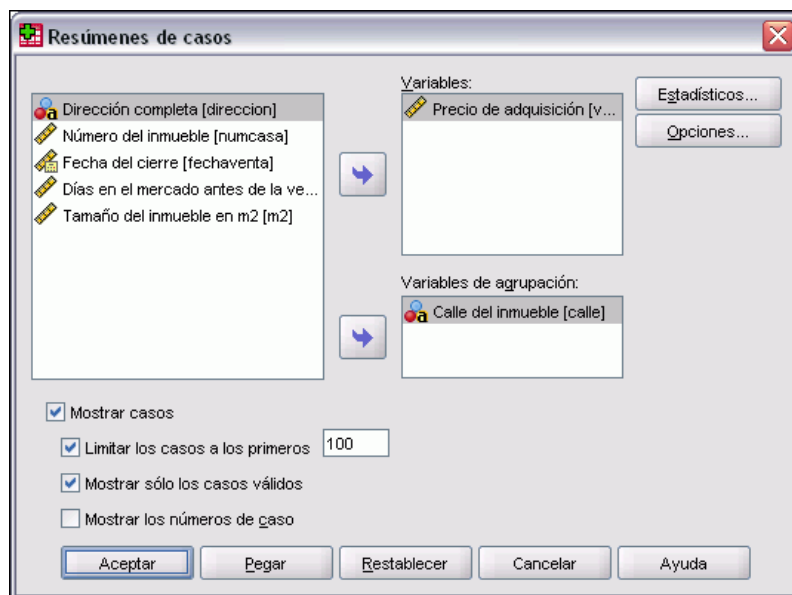
**Supuestos.** Algunos de los estadísticos opcionales de subgrupo, como la media y la desviación típica, se basan en la teoría normal y son adecuados para variables cuantitativas con distribuciones simétricas. Los estadísticos robustos, tales como la mediana y el rango, son adecuados para las variables cuantitativas que pueden o no cumplir el supuesto de normalidad.

### **Para obtener resúmenes de casos**

- ▶ Elija en los menús:
  - Analizar
  - Informes
  - Resúmenes de casos...



Figura 17-1  
Cuadro de diálogo Resúmenes de casos



- Seleccione una o más variables.

Si lo desea, puede:

- Seleccionar una o más variables de agrupación para dividir los datos en subgrupos.
- Pulsar en Opciones para cambiar el título de los resultados, añadir un texto al pie debajo de los resultados o excluir los casos con valores perdidos.
- Pulsar en Estadísticos para acceder a estadísticos adicionales.
- Seleccionar Mostrar los casos para listar los casos en cada subgrupo. Por defecto, el sistema enumera sólo los 100 primeros casos del archivo. Puede aumentar o disminuir el valor de Limitar los casos a los primeros  $n$  o desactivar ese elemento para enumerar todos los casos.

## Resumir: Opciones

Figura 17-2  
Cuadro de diálogo Opciones

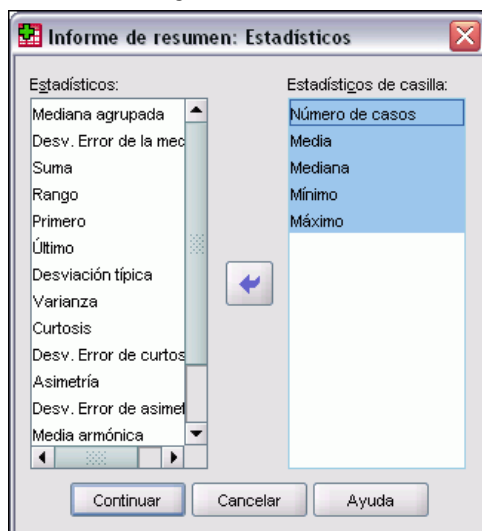


Resumir permite cambiar el título de los resultados o añadir un texto que aparecerá debajo de la tabla de resultados. Puede controlar el ajuste de las líneas en los títulos y textos escribiendo `\n` en el lugar donde desee insertar una línea de separación.

Además, puede elegir entre mostrar o suprimir los subtítulos para los totales e incluir o excluir los casos con valores perdidos para cualquiera de las variables utilizadas en cualquiera de los análisis. A menudo es aconsejable representar los casos perdidos en los resultados con un punto o un asterisco. Introduzca un carácter, frase o código que desee que aparezca cuando haya un valor perdido; de lo contrario, no se aplicará ningún tratamiento especial a los casos perdidos en los resultados.

## Resumir: Estadísticos

Figura 17-3  
Cuadro de diálogo Estadísticos de informe de resumen



Puede elegir uno o más de los siguientes estadísticos de subgrupo para las variables dentro de cada categoría de cada variable de agrupación: suma, número de casos, media, mediana, mediana agrupada, error típico de la media, mínimo, máximo, rango, valor de la variable para la primera categoría de la variable de agrupación, valor de la variable para la última categoría de la variable de agrupación, desviación típica, varianza, curtosis, error típico de curtosis, asimetría, error típico de asimetría, porcentaje de la suma total, porcentaje del  $N$  total, porcentaje de la suma en, porcentaje de  $N$  en, media geométrica y media armónica. El orden en el que aparecen los estadísticos en la lista Estadísticos de casilla es el orden en el que se mostrarán en los resultados. También se muestran estadísticos de resumen para cada variable a través de todas las categorías.

**Primera.** Muestra el primer valor de los datos encontrado en el archivo de datos.

**Media geométrica.** La raíz  $n$ -ésima del producto de los valores de los datos, donde  $n$  representa el número de casos.

**Mediana agrupada.** La mediana calculada para los datos que se codifican en grupos. Por ejemplo, para los datos de edades, si cada valor de una treintena se codifica como 35, cada valor de una cuarentena se codifica como 45 y así sucesivamente, la mediana agrupada es la mediana calculada a partir de los datos codificados.

**Media armónica.** Se utiliza para estimar el tamaño promedio de un grupo cuando los tamaños de las muestras de los grupos no son iguales. La media armónica es el número total de muestras dividido por la suma de los inversos de los tamaños de las muestras.

**Curtosis.** Es una medida del grado en que las observaciones se agrupan en torno a un punto central. En una distribución normal, el valor del estadístico de curtosis es 0. Una curtosis positiva indica que las observaciones se concentran más y presentan colas más largas que las de una distribución normal. Una curtosis negativa indica que las observaciones se agrupan menos y presentan colas más cortas.

**Último.** Muestra el último valor de los datos encontrado en el archivo de datos.

**Máximo.** Valor más grande de una variable numérica.

**Media.** Una medida de tendencia central. El promedio aritmético; la suma dividida por el número de casos.

**Mediana.** Valor por encima y por debajo del cual se encuentran la mitad de los casos; el percentil 50. Si hay un número par de casos, la mediana es la media de los dos valores centrales, cuando los casos se ordenan en orden ascendente o descendente. La mediana es una medida de tendencia central que no es sensible a los valores atípicos (a diferencia de la media, que puede resultar afectada por unos pocos valores extremadamente altos o bajos).

**Mínimo.** Valor más pequeño de una variable numérica.

**N.** Número de casos (observaciones o registros).

**Porcentaje del N total.** Porcentaje del número total de casos en cada categoría.

**Porcentaje de la suma total.** Porcentaje de la suma total en cada categoría.

**Amplitud.** Diferencia entre los valores mayor y menor de una variable numérica; el máximo menos el mínimo. También se denomina el recorrido de la variable.

**Asimetría.** Medida de la asimetría de una distribución. Las distribuciones normales son simétricas, por lo que tienen un valor de asimetría 0. Una distribución que tenga una asimetría significativa positiva tiene una cola derecha larga. Una distribución que tenga una asimetría negativa significativa tiene una cola izquierda larga. Como regla aproximada, un valor de la asimetría mayor que el doble de su error típico se asume que indica una desviación de la simetría.

**Error típico de la curtosis.** La razón de la curtosis sobre su error típico puede utilizarse como contraste de la normalidad (es decir, se puede rechazar la normalidad si la razón es menor que -2 o mayor que +2). Un valor grande y positivo para la curtosis indica que las colas son más largas que las de una distribución normal; por el contrario, un valor extremo y negativo indica que las colas son más cortas (llegando a tener forma de caja como en la distribución uniforme).

**Error típico de la asimetría.** La razón de la asimetría sobre su error típico puede utilizarse como un contraste de la normalidad (es decir, se puede rechazar la normalidad si la razón es menor que -2 o mayor que +2). Un valor grande y positivo para la asimetría indica una cola larga a la derecha; un valor extremo y negativo indica una cola larga por la izquierda.

**Suma.** La suma o total de todos los valores, a lo largo de todos los casos que no tengan valores perdidos.

**Varianza.** Medida de dispersión en torno a la media, igual a la suma de los cuadrados de las desviaciones respecto a la media dividida por el número de casos menos 1. La varianza se mide en unas unidades que son el cuadrado de las de la propia variable.

# Medias

El procedimiento Medias calcula medias de subgrupo y estadísticos univariados relacionados para variables dependientes dentro de las categorías de una o más variables independientes. Si lo desea, puede obtener el análisis de varianza de un factor, la eta y pruebas de linealidad.

**Ejemplo.** Mida la cantidad media de grasa absorbida en función de tres tipos distintos de aceite comestible y realice un análisis de varianza de un factor para comprobar si difieren las medias.

**Estadísticos.** Suma, número de casos, media, mediana, mediana agrupada, error típico de la media, mínimo, máximo, rango, valor de la variable para la primera categoría de la variable de agrupación, valor de la variable para la última categoría de la variable de agrupación, desviación típica, varianza, curtosis, error típico de curtosis, asimetría, error típico de asimetría, porcentaje de la suma total, porcentaje del  $N$  total, porcentaje de la suma en, porcentaje de  $N$  en, media geométrica y media armónica. Las opciones incluyen: análisis de varianza, eta, eta cuadrado y pruebas de linealidad de  $R$  y  $R^2$ .

**Datos.** Las variables dependientes son cuantitativas y las independientes son categóricas. Los valores de las variables categóricas pueden ser numéricos o de cadena.

**Supuestos.** Algunos de los estadísticos opcionales de subgrupo, como la media y la desviación típica, se basan en la teoría normal y son adecuados para variables cuantitativas con distribuciones simétricas. Los estadísticos robustos, tales como la mediana son adecuados para las variables cuantitativas que pueden o no cumplir el supuesto de normalidad. El análisis de varianza es robusto a las desviaciones de la normalidad, aunque los datos de cada casilla deberían ser simétricos. El análisis de varianza también supone que los grupos proceden de poblaciones con la misma varianza. Para comprobar este supuesto, utilice la prueba de homogeneidad de las varianzas de Levene, disponible en el procedimiento ANOVA de un factor.

## **Para obtener medias de subgrupo**

- ▶ Elija en los menús:
  - Analizar
  - Comparar medias
  - Medias...

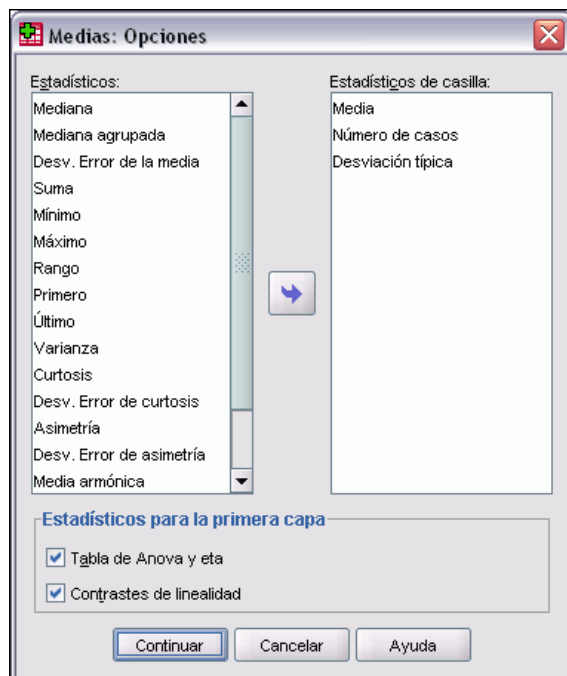
Figura 18-1  
Cuadro de diálogo Medias



- ▶ Seleccione una o más variables dependientes.
- ▶ Utilice uno de los siguientes métodos para seleccionar variables independientes categóricas:
  - Seleccione una o más variables independientes. Se mostrarán resultados individuales para cada variable independiente.
  - Seleccione una o más capas de variables independientes. Cada capa subdivide consecutivamente la muestra. Si tiene una variable independiente en Capa 1 y otra variable independiente en Capa 2, los resultados se mostrarán en una tabla cruzada en contraposición a tablas individuales para cada variable independiente.
- ▶ Si lo desea, pulse en Opciones si desea obtener estadísticos opcionales, una tabla de análisis de varianza, eta, eta cuadrado,  $R$ , y  $R^2$ .

## Medias: Opciones

Figura 18-2  
Cuadro de diálogo Medias: Opciones



Puede elegir uno o más de los siguientes estadísticos de subgrupo para las variables dentro de cada categoría de cada variable de agrupación: suma, número de casos, media, mediana, mediana agrupada, error típico de la media, mínimo, máximo, rango, valor de la variable para la primera categoría de la variable de agrupación, valor de la variable para la última categoría de la variable de agrupación, desviación típica, varianza, curtosis, error típico de curtosis, asimetría, error típico de asimetría, porcentaje de la suma total, porcentaje del  $N$  total, porcentaje de la suma en, porcentaje de  $N$  en, media geométrica, media armónica. Se puede cambiar el orden de aparición de los estadísticos de subgrupo. El orden en el que aparecen en la lista Estadísticos de casilla es el mismo orden que presentarán en los resultados. También se muestran estadísticos de resumen para cada variable a través de todas las categorías.

**Primera.** Muestra el primer valor de los datos encontrado en el archivo de datos.

**Media geométrica.** La raíz  $n$ -ésima del producto de los valores de los datos, donde  $n$  representa el número de casos.

**Mediana agrupada.** La mediana calculada para los datos que se codifican en grupos. Por ejemplo, para los datos de edades, si cada valor de una treintena se codifica como 35, cada valor de una cuarentena se codifica como 45 y así sucesivamente, la mediana agrupada es la mediana calculada a partir de los datos codificados.

**Media armónica.** Se utiliza para estimar el tamaño promedio de un grupo cuando los tamaños de las muestras de los grupos no son iguales. La media armónica es el número total de muestras dividido por la suma de los inversos de los tamaños de las muestras.

**Curtosis.** Es una medida del grado en que las observaciones se agrupan en torno a un punto central. En una distribución normal, el valor del estadístico de curtosis es 0. Una curtosis positiva indica que las observaciones se concentran más y presentan colas más largas que las de una distribución normal. Una curtosis negativa indica que las observaciones se agrupan menos y presentan colas más cortas.

**Último.** Muestra el último valor de los datos encontrado en el archivo de datos.

**Máximo.** Valor más grande de una variable numérica.

**Media.** Una medida de tendencia central. El promedio aritmético; la suma dividida por el número de casos.

**Mediana.** Valor por encima y por debajo del cual se encuentran la mitad de los casos; el percentil 50. Si hay un número par de casos, la mediana es la media de los dos valores centrales, cuando los casos se ordenan en orden ascendente o descendente. La mediana es una medida de tendencia central que no es sensible a los valores atípicos (a diferencia de la media, que puede resultar afectada por unos pocos valores extremadamente altos o bajos).

**Mínimo.** Valor más pequeño de una variable numérica.

**N.** Número de casos (observaciones o registros).

**Porcentaje del N total.** Porcentaje del número total de casos en cada categoría.

**Porcentaje de la suma total.** Porcentaje de la suma total en cada categoría.

**Amplitud.** Diferencia entre los valores mayor y menor de una variable numérica; el máximo menos el mínimo. También se denomina el recorrido de la variable.

**Asimetría.** Medida de la asimetría de una distribución. Las distribuciones normales son simétricas, por lo que tienen un valor de asimetría 0. Una distribución que tenga una asimetría significativa positiva tiene una cola derecha larga. Una distribución que tenga una asimetría negativa significativa tiene una cola izquierda larga. Como regla aproximada, un valor de la asimetría mayor que el doble de su error típico se asume que indica una desviación de la simetría.

**Error típico de la curtosis.** La razón de la curtosis sobre su error típico puede utilizarse como contaste de la normalidad (es decir, se puede rechazar la normalidad si la razón es menor que -2 o mayor que +2). Un valor grande y positivo para la curtosis indica que las colas son más largas que las de una distribución normal; por el contrario, un valor extremo y negativo indica que las colas son más cortas (llegando a tener forma de caja como en la distribución uniforme).

**Error típico de la asimetría.** La razón de la asimetría sobre su error típico puede utilizarse como un contraste de la normalidad (es decir, se puede rechazar la normalidad si la razón es menor que -2 o mayor que +2). Un valor grande y positivo para la asimetría indica una cola larga a la derecha; un valor extremo y negativo indica una cola larga por la izquierda.

**Suma.** La suma o total de todos los valores, a lo largo de todos los casos que no tengan valores perdidos.

**Varianza.** Medida de dispersión en torno a la media, igual a la suma de los cuadrados de las desviaciones respecto a la media dividida por el número de casos menos 1. La varianza se mide en unas unidades que son el cuadrado de las de la propia variable.



***Estadísticos para la primera capa***

**Tabla de Anova y eta.** Muestra una tabla de análisis de varianza de un factor y calcula la eta y la eta cuadrado (son medidas de asociación) para cada variable independiente de la primera capa.

**Contrastes de linealidad.** Calcula la suma de cuadrados, los grados de libertad y la media cuadrática asociados a los componentes lineal y no lineal, así como la razón F, la R y la R cuadrado. Si la variable independiente es una cadena corta entonces la linealidad no se calcula.

## ***Cubos OLAP***

El procedimiento Cubos OLAP (siglas del inglés On-Line Analytic Processing, «Procesamiento analítico interactivo») calcula totales, medias y otros estadísticos univariantes para variables de resumen continuas dentro de las categorías de una o más variables categóricas de agrupación. En la tabla se creará una nueva capa para cada categoría de cada variable de agrupación.

**Ejemplo.** El total y el promedio de ventas para diversas regiones y líneas de producto, dentro de las regiones.

**Estadísticos.** Suma, número de casos, media, mediana, mediana agrupada, error típico de la media, mínimo, máximo, rango, valor de la variable para la primera categoría de la variable de agrupación, valor de la variable para la última categoría de la variable de agrupación, desviación típica, varianza, curtosis, error típico de curtosis, asimetría, error típico de asimetría, porcentaje de casos totales, porcentaje de la suma total, porcentaje de casos totales dentro de las variables agrupadas, porcentaje de la suma total dentro de las variables agrupadas, media geométrica y media armónica.

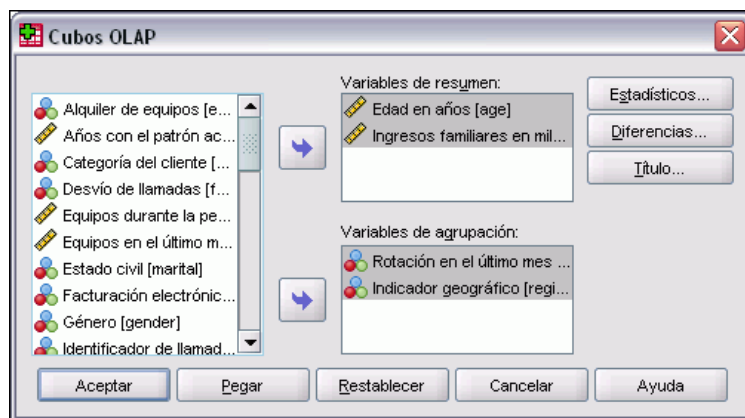
**Datos.** Las variables de resumen son cuantitativas (variables continuas medidas en una escala de intervalo o de razón) y las variables de agrupación son categóricas. Los valores de las variables categóricas pueden ser numéricos o de cadena.

**Supuestos.** Algunos de los estadísticos opcionales de subgrupo, como la media y la desviación típica, se basan en la teoría normal y son adecuados para variables cuantitativas con distribuciones simétricas. Los estadísticos robustos, tales como la mediana y el rango, son adecuados para las variables cuantitativas que pueden o no cumplir el supuesto de normalidad.

### ***Para obtener cubos OLAP***

- ▶ Elija en los menús:
  - Analizar
  - Informes
  - Cubos OLAP..

Figura 19-1  
Cuadro de diálogo Cubos OLAP



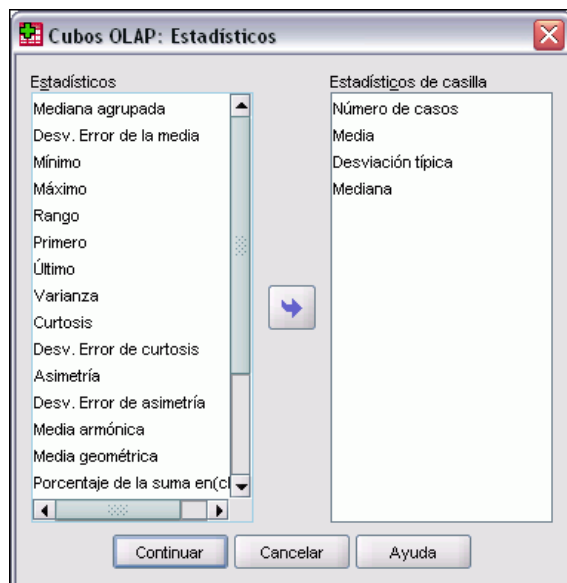
- ▶ Seleccione una o más variables de resumen continuas.
- ▶ Seleccione una o más variables de agrupación categóricas.

Si lo desea:

- Seleccionar diferentes estadísticos de resumen (pulse en Estadísticos). Debe seleccionar una o más variables de agrupación para poder seleccionar estadísticos de resumen.
- Calcule las diferencias existentes entre los pares de variables y los pares de grupos definidos por una variable de agrupación (pulse en Diferencias).
- Crear títulos de tabla personalizados (pulse en Título).

## Cubos OLAP: Estadísticos

Figura 19-2  
Cuadro de diálogo Cubos OLAP: Estadísticos



Puede elegir uno o varios de los siguientes estadísticos de subgrupo para las variables de resumen dentro de cada categoría de cada variable de agrupación: Suma, Número de casos, Media, Mediana, Mediana agrupada, Error típico de la media, Mínimo, Máximo, Rango, Valor de la variable para la primera categoría de la variable de agrupación, Valor de la variable para la última categoría de la variable de agrupación, Desviación típica, Varianza, Curtosis, Error típico de curtosis, Asimetría, Error típico de asimetría, Porcentaje de casos totales, Porcentaje de la suma total, Porcentaje de casos totales dentro de las variables de agrupación, Porcentaje de la suma total dentro de las variables de agrupación, Media geométrica y Media armónica.

Se puede cambiar el orden de aparición de los estadísticos de subgrupo. El orden en el que aparecen en la lista Estadísticos de casilla es el mismo orden que presentarán en los resultados. También se muestran estadísticos de resumen para cada variable a través de todas las categorías.

**Primera.** Muestra el primer valor de los datos encontrado en el archivo de datos.

**Media geométrica.** La raíz enésima del producto de los valores de los datos, donde  $n$  representa el número de casos.

**Mediana agrupada.** La mediana calculada para los datos que se codifican en grupos. Por ejemplo, para los datos de edades, si cada valor de una treintena se codifica como 35, cada valor de una cuarentena se codifica como 45 y así sucesivamente, la mediana agrupada es la mediana calculada a partir de los datos codificados.

**Media armónica.** Se utiliza para estimar el tamaño promedio de un grupo cuando los tamaños de las muestras de los grupos no son iguales. La media armónica es el número total de muestras dividido por la suma de los inversos de los tamaños de las muestras.

**Curtosis.** Es una medida del grado en que las observaciones se agrupan en torno a un punto central. En una distribución normal, el valor del estadístico de curtosis es 0. Una curtosis positiva indica que las observaciones se concentran más y presentan colas más largas que las de una distribución normal. Una curtosis negativa indica que las observaciones se agrupan menos y presentan colas más cortas.

**Último.** Muestra el último valor de los datos encontrado en el archivo de datos.

**Máximo.** Valor más grande de una variable numérica.

**Media.** Una medida de tendencia central. El promedio aritmético; la suma dividida por el número de casos.

**Mediana.** Valor por encima y por debajo del cual se encuentran la mitad de los casos; el percentil 50. Si hay un número par de casos, la mediana es la media de los dos valores centrales, cuando los casos se ordenan en orden ascendente o descendente. La mediana es una medida de tendencia central que no es sensible a los valores atípicos (a diferencia de la media, que puede resultar afectada por unos pocos valores extremadamente altos o bajos).

**Mínimo.** Valor más pequeño de una variable numérica.

**N.** Número de casos (observaciones o registros).

**Porcentaje de N en.** Porcentaje del número de casos para la variable de agrupación especificada dentro de las categorías de otras variables de agrupación. Si sólo tiene una variable de agrupación, este valor es idéntico al porcentaje del número de casos total.

**Porcentaje de la suma en.** Porcentaje de la suma para la variable de agrupación especificada dentro de las categorías de otras variables de agrupación. Si sólo tiene una variable de agrupación, este valor es idéntico al porcentaje de la suma total.

**Porcentaje del N total.** Porcentaje del número total de casos en cada categoría.

**Porcentaje de la suma total.** Porcentaje de la suma total en cada categoría.

**Amplitud.** Diferencia entre los valores mayor y menor de una variable numérica; el máximo menos el mínimo. También se denomina el recorrido de la variable.

**Asimetría.** Medida de la asimetría de una distribución. Las distribuciones normales son simétricas, por lo que tienen un valor de asimetría 0. Una distribución que tenga una asimetría significativa positiva tiene una cola derecha larga. Una distribución que tenga una asimetría negativa significativa tiene una cola izquierda larga. Como regla aproximada, un valor de la asimetría mayor que el doble de su error típico se asume que indica una desviación de la simetría.

**Error típico de la curtosis.** La razón de la curtosis sobre su error típico puede utilizarse como contraste de la normalidad (es decir, se puede rechazar la normalidad si la razón es menor que -2 o mayor que +2). Un valor grande y positivo para la curtosis indica que las colas son más largas que las de una distribución normal; por el contrario, un valor extremo y negativo indica que las colas son más cortas (llegando a tener forma de caja como en la distribución uniforme).

**Error típico de la asimetría.** La razón de la asimetría sobre su error típico puede utilizarse como un contraste de la normalidad (es decir, se puede rechazar la normalidad si la razón es menor que -2 o mayor que +2). Un valor grande y positivo para la asimetría indica una cola larga a la derecha; un valor extremo y negativo indica una cola larga por la izquierda.

**Suma.** La suma o total de todos los valores, a lo largo de todos los casos que no tengan valores perdidos.

**Varianza.** Medida de dispersión en torno a la media, igual a la suma de los cuadrados de las desviaciones respecto a la media dividida por el número de casos menos 1. La varianza se mide en unas unidades que son el cuadrado de las de la propia variable.

## Cubos OLAP: Diferencias

Figura 19-3  
Cuadro de diálogo Cubos OLAP: Diferencias

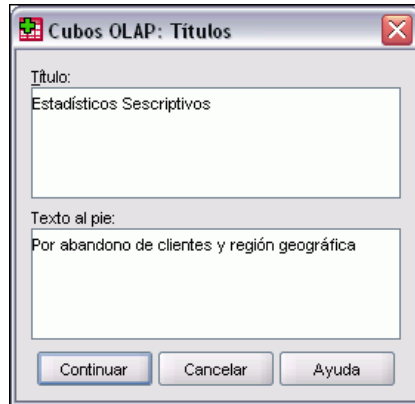
Este cuadro de diálogo le permite calcular el porcentaje y las diferencias aritméticas entre las variables de resumen o entre los grupos definidos por una variable de agrupación. Las diferencias se calculan para todas las medidas seleccionadas en el cuadro de diálogo Cubos OLAP: Estadísticos.

**Diferencias entre variables.** Calcula las diferencias entre pares de variables. Los valores de los estadísticos de resumen para la segunda variable de cada par (la variable Menos) se restan de los valores de los estadísticos de resumen correspondientes a la primera variable del par. En cuanto a las diferencias de porcentaje, el valor de la variable de resumen para la variable Menos es el que se usa como denominador. Debe seleccionar al menos dos variables de resumen en el cuadro de diálogo principal para poder especificar las diferencias entre las variables.

**Diferencias entre grupos de casos.** Calcula las diferencias entre pares de grupos definidos por una variable de agrupación. Los valores de los estadísticos de resumen para la segunda categoría de cada par (la variable Menos) se restan de los valores de los estadísticos de resumen correspondientes a la primera categoría del par. Las diferencias de porcentaje utilizan el valor del estadístico de resumen de la categoría Menos como denominador. Debe seleccionar una o más variables de agrupación en el cuadro de diálogo principal para poder especificar las diferencias entre los grupos.

## Cubos OLAP: Título

Figura 19-4  
Cuadro de diálogo Cubos OLAP: Título



Puede cambiar el título de los resultados o añadir un texto al pie que aparecerá debajo de la tabla de resultados. También puede controlar el ajuste de las líneas de los títulos y de los textos al pie escribiendo `\n` en el lugar del texto donde desee insertar una línea de separación.

# Pruebas T

Hay tres tipos de pruebas  $t$ :

**Prueba T para muestras independientes (prueba T para dos muestras).** Compara las medias de una variable para dos grupos de casos. Se ofrecen estadísticos descriptivos para cada grupo y la prueba de Levene sobre la igualdad de las varianzas, así como valores  $t$  de igualdad de varianzas y varianzas desiguales y un intervalo de confianza al 95% para la diferencia entre las medias.

**Prueba T para muestras relacionadas (prueba T dependiente).** Compara las medias de dos variables en un solo grupo. Esta prueba también se utiliza para pares relacionados o diseños de estudio de control de casos. El resultado incluye estadísticos descriptivos de las variables que se van a contrastar, la correlación entre ellas, estadísticos descriptivos de las diferencias emparejadas, la prueba  $t$  y un intervalo de confianza al 95%.

**Prueba  $t$  para una muestra.** Compara la media de una variable con un valor conocido o hipotetizado. Se muestran estadísticos descriptivos para las variables de contraste junto con la prueba  $t$ . Por defecto, en los resultados se incluye un intervalo de confianza al 95% para la diferencia entre la media de la variable de contraste y el valor hipotetizado de la prueba.

## Prueba T para muestras independientes

El procedimiento Prueba T para muestras independientes compara las medias de dos grupos de casos. Lo ideal es que para esta prueba los sujetos se asignen aleatoriamente a dos grupos, de forma que cualquier diferencia en la respuesta sea debida al tratamiento (o falta de tratamiento) y no a otros factores. Este caso no ocurre si se comparan los ingresos medios para hombres y mujeres. El sexo de una persona no se asigna aleatoriamente. En estas situaciones, debe asegurarse de que las diferencias en otros factores no enmascaren o resalten una diferencia significativa entre las medias. Las diferencias de ingresos medios pueden estar sometidas a la influencia de factores como los estudios (y no solamente el sexo).

**Ejemplo.** Se asigna aleatoriamente un grupo de pacientes con hipertensión arterial a un grupo con placebo y otro con tratamiento. Los sujetos con placebo reciben una pastilla inactiva y los sujetos con tratamiento reciben un nuevo medicamento del cual se espera que reduzca la tensión arterial. Después de tratar a los sujetos durante dos meses, se utiliza la prueba  $t$  para dos muestras para comparar la tensión arterial media del grupo con placebo y del grupo con tratamiento. Cada paciente se mide una sola vez y pertenece a un solo grupo.

**Estadísticos.** Para cada variable: tamaño muestral, media, desviación típica y error típico de la media. Para la diferencia entre las medias: media, error típico e intervalo de confianza (puede especificar el nivel de confianza). Pruebas: prueba de Levene sobre la igualdad de varianzas y pruebas  $t$  de varianzas combinadas y separadas sobre la igualdad de las medias.



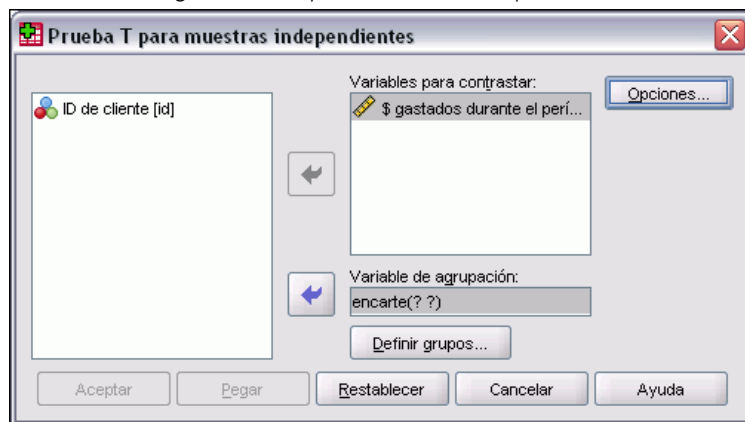
**Datos.** Los valores de la variable cuantitativa de interés se hallan en una única columna del archivo de datos. El procedimiento utiliza una variable de agrupación con dos valores para separar los casos en dos grupos. La variable de agrupación puede ser numérica (valores como 1 y 2, o 6,25 y 12,5) o de cadena corta (como *sí* y *no*). También puede usar una variable cuantitativa, como la *edad*, para dividir los casos en dos grupos especificando un punto de corte (el punto de corte 21 divide la *edad* en un grupo de menos de 21 años y otro de más de 21).

**Supuestos.** Para la prueba *t* de igualdad de varianzas, las observaciones deben ser muestras aleatorias independientes de distribuciones normales con la misma varianza de población. Para la prueba *t* de varianzas desiguales, las observaciones deben ser muestras aleatorias independientes de distribuciones normales. La prueba *t* para dos muestras es bastante robusta a las desviaciones de la normalidad. Al contrastar las distribuciones gráficamente, compruebe que son simétricas y que no contienen valores atípicos.

### **Para obtener una prueba T para muestras independientes**

- ▶ Elija en los menús:
  - Analizar
  - Comparar medias
  - Prueba T para muestras independientes...

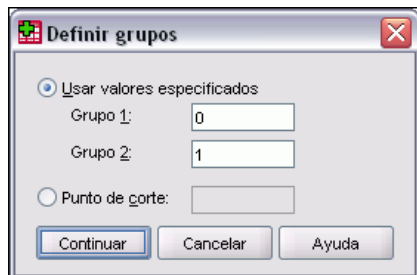
Figura 20-1  
Cuadro de diálogo Prueba T para muestras independientes



- ▶ Seleccione una o más variables de contraste cuantitativas. Se calcula una prueba *t* para cada variable.
- ▶ Seleccione una sola variable de agrupación y pulse en Definir grupos para especificar dos códigos para los grupos que desee comparar.
- ▶ Si lo desea, puede pulsar en Opciones para controlar el tratamiento de los datos perdidos y el nivel del intervalo de confianza.

## Definición de grupos en la prueba *T* para muestras independientes

Figura 20-2  
Cuadro de diálogo Definir grupos para variables numéricas

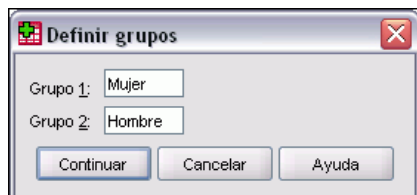


El cuadro de diálogo 'Definir grupos' muestra la opción 'Usar valores especificados' seleccionada. Los campos 'Grupo 1' y 'Grupo 2' contienen los valores '0' y '1' respectivamente. La opción 'Punto de corte' no está seleccionada. En la parte inferior hay tres botones: 'Continuar', 'Cancelar' y 'Ayuda'.

Para las variables de agrupación numéricas, defina los dos grupos de la prueba *t* especificando dos valores o un punto de corte:

- **Usar valores especificados.** Escriba un valor para el Grupo 1 y otro para el Grupo 2. Los casos con otros valores quedarán excluidos del análisis. Los números no tienen que ser enteros (por ejemplo, 6,25 y 12,5 son válidos).
- **Punto de Corte.** Escriba un número que divida los valores de la variable de agrupación en dos conjuntos. Todos los casos con valores menores que el punto de corte forman un grupo y los casos con valores mayores o iguales que el punto de corte forman el otro grupo.

Figura 20-3  
Cuadro de diálogo Definir grupos para variables de cadena

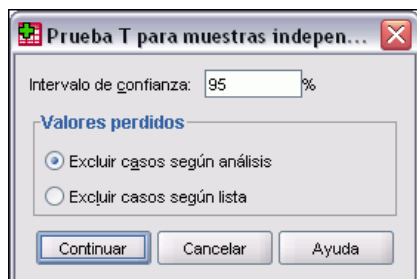


El cuadro de diálogo 'Definir grupos' muestra los campos 'Grupo 1' y 'Grupo 2' con los valores 'Mujer' y 'Hombre' respectivamente. En la parte inferior hay tres botones: 'Continuar', 'Cancelar' y 'Ayuda'.

Para las variables de agrupación de cadena, escriba una cadena para el Grupo 1 y otra para el Grupo 2; por ejemplo *sí* y *no*. Los casos con otras cadenas se excluyen del análisis.

## Prueba *T* para muestras independientes: Opciones

Figura 20-4  
Cuadro de diálogo Prueba *T* para muestras independientes: Opciones



El cuadro de diálogo 'Prueba T para muestras independientes: Opciones' muestra un campo 'Intervalo de confianza' con el valor '95' y el símbolo '%'. En la sección 'Valores perdidos', la opción 'Excluir casos según análisis' está seleccionada. En la parte inferior hay tres botones: 'Continuar', 'Cancelar' y 'Ayuda'.

**Intervalo de confianza.** Por defecto se muestra un intervalo de confianza al 95% para la diferencia entre las medias. Introduzca un valor entre 1 y 99 para solicitar otro nivel de confianza.

**Valores perdidos.** Si ha probado varias variables y se han perdido los datos de una o más de ellas, puede indicar al procedimiento qué casos desea incluir (o excluir).

- **Excluir casos según análisis.** Cada prueba  $t$  utiliza todos los casos que tienen datos válidos para las variables contrastadas. Los tamaños muestrales pueden variar de una prueba a otra.
- **Excluir casos según lista.** Cada prueba  $t$  utiliza sólo aquellos casos que contienen datos válidos para todas las variables utilizadas en las pruebas  $t$  solicitadas. El tamaño muestral es constante en todas las pruebas.

## ***Prueba T para muestras relacionadas***

El procedimiento Prueba T para muestras relacionadas compara las medias de dos variables de un solo grupo. El procedimiento calcula las diferencias entre los valores de las dos variables de cada caso y contrasta si la media difiere de 0.

**Ejemplo.** En un estudio sobre la hipertensión sanguínea, se toma la tensión a todos los pacientes al comienzo del estudio, se les aplica un tratamiento y se les toma la tensión otra vez. De esta manera, a cada paciente le corresponden dos medidas, normalmente denominadas medidas *pre* y *post*. Un diseño alternativo para el que se utiliza esta prueba consiste en un estudio de pares relacionados o un estudio de control de casos en el que cada registro en el archivo de datos contiene la respuesta del paciente y de su sujeto de control correspondiente. En un estudio sobre la tensión sanguínea, pueden emparejarse pacientes y controles por edad (un paciente de 75 años con un miembro del grupo de control de 75 años).

**Estadísticos.** Para cada variable: media, tamaño muestral, desviación típica y error típico de la media. Para cada pareja de variables: correlación, diferencia promedio entre las medias, prueba  $t$  de intervalo de confianza para la diferencia entre las medias (puede especificarse el nivel de confianza). Desviación típica y error típico de la diferencia entre las medias.

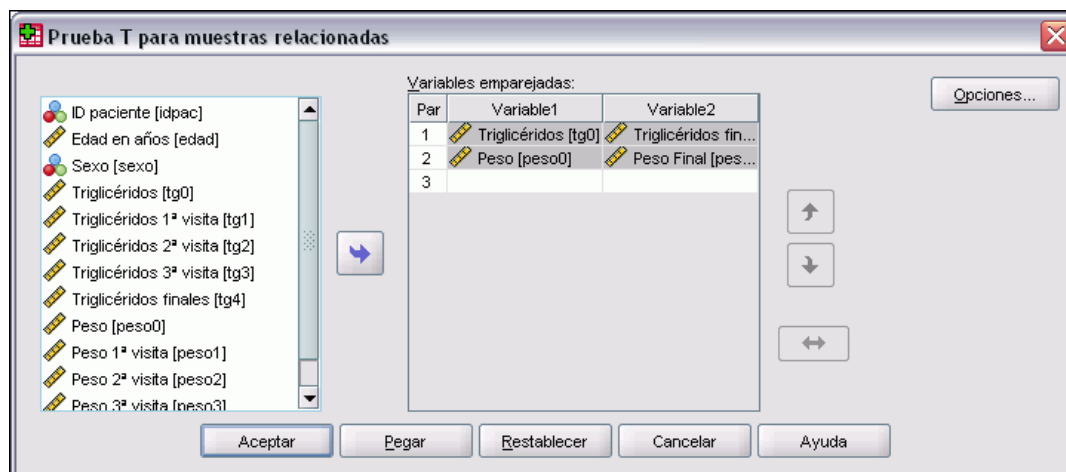
**Datos.** Especifique dos variables cuantitativas (nivel de medida de intervalo o de razón) para cada prueba de pares. En un estudio de pares relacionados o de control de casos, la respuesta de cada sujeto de la prueba y su sujeto de control correspondiente deberán hallarse en el mismo caso en el archivo de datos.

**Supuestos.** Las observaciones de cada par deben hacerse en las mismas condiciones. Las diferencias entre las medias deben estar normalmente distribuidas. Las varianzas de cada variable pueden ser iguales o desiguales.

### ***Para obtener una prueba T para muestras relacionadas***

- ▶ Elija en los menús:
  - Analizar
  - Comparar medias
  - Prueba T para muestras relacionadas...

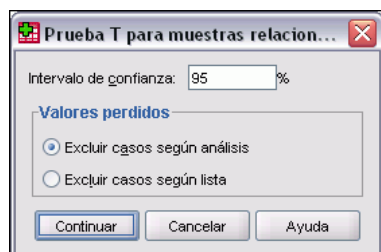
Figura 20-5  
Cuadro de diálogo Prueba T para muestras relacionadas



- ▶ Seleccione uno o más pares de variables
- ▶ Si lo desea, puede pulsar en Opciones para controlar el tratamiento de los datos perdidos y el nivel del intervalo de confianza.

### **Prueba T para muestras relacionadas: Opciones**

Figura 20-6  
Cuadro de diálogo Prueba T para muestras relacionadas: Opciones



**Intervalo de confianza.** Por defecto se muestra un intervalo de confianza al 95% para la diferencia entre las medias. Introduzca un valor entre 1 y 99 para solicitar otro nivel de confianza.

**Valores perdidos.** Si ha probado varias variables y se han perdido los datos de una o más de ellas, puede indicar al procedimiento qué casos desea incluir (o excluir):

- **Excluir casos según análisis.** Cada prueba  $t$  utilizará todos los casos que contienen datos válidos para la pareja de variables contrastadas. Los tamaños muestrales pueden variar de una prueba a otra.
- **Excluir casos según lista.** Cada prueba  $t$  utilizará únicamente los casos que contengan datos válidos para todas las parejas de variables contrastadas. El tamaño muestral es constante en todas las pruebas.

## Prueba T para una muestra

El procedimiento Prueba T para una muestra contrasta si la media de una sola variable difiere de una constante especificada.

**Ejemplos.** Un investigador desea comprobar si la puntuación media del coeficiente intelectual de un grupo de alumnos difiere de 100. O bien, un fabricante de copos de cereales puede tomar una muestra de envases de la línea de producción y comprobar si el peso medio de las muestras difiere de 1 kg con un nivel de confianza al 95%.

**Estadísticos.** Para cada variable a contrastar: media, desviación típica y error típico de la media. La diferencia promedio entre cada valor de los datos y el valor de contraste hipotetizado, una prueba  $t$  que contrasta que esta diferencia es 0 y un intervalo de confianza para la diferencia promedio (para el que puede especificarse el nivel de confianza).

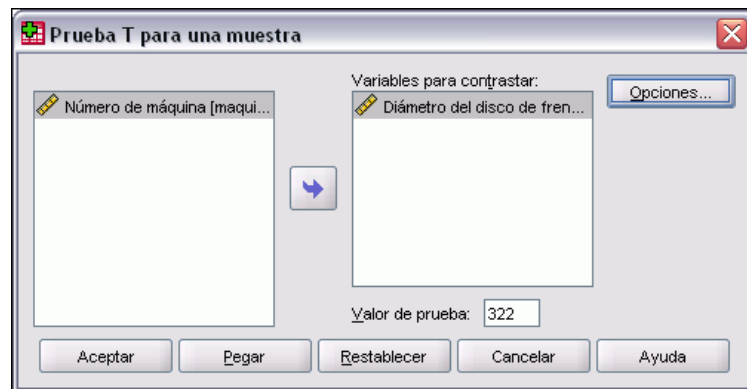
**Datos.** Para contrastar los valores de una variable cuantitativa con un valor de contraste hipotetizado, elija una variable cuantitativa e introduzca un valor de contraste hipotetizado.

**Supuestos.** Esta prueba asume que los datos están normalmente distribuidos; sin embargo, esta prueba es bastante robusto frente a las desviaciones de la normalidad.

### Para obtener una prueba T para una muestra

- ▶ Elija en los menús:
  - Analizar
  - Comparar medias
  - Prueba T para una muestra...

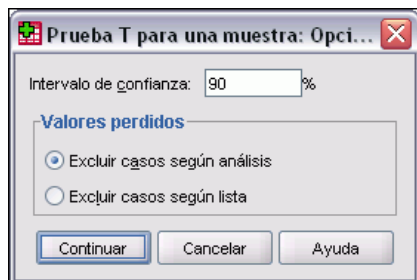
Figura 20-7  
Cuadro de diálogo Prueba T para una muestra



- ▶ Seleccione una o más variables para contrastarlas con el mismo valor hipotetizado.
- ▶ Introduzca un valor de contraste numérico para compararlo con cada media muestral.
- ▶ Si lo desea, puede pulsar en Opciones para controlar el tratamiento de los datos perdidos y el nivel del intervalo de confianza.

## Prueba T para una muestra: Opciones

Figura 20-8  
Cuadro de diálogo Prueba T para una muestra: Opciones



**Intervalo de confianza.** Por defecto se muestra un intervalo de confianza al 95% para la diferencia entre la media y el valor de contraste hipotetizado. Introduzca un valor entre 1 y 99 para solicitar otro nivel de confianza.

**Valores perdidos.** Si ha probado varias variables y se han perdido los datos de una o más de ellas, puede indicar al procedimiento qué casos desea incluir (o excluir).

- **Excluir casos según análisis.** Cada prueba  $t$  utiliza todos los casos que tienen datos válidos para la variable contrastada. Los tamaños muestrales pueden variar de una prueba a otra.
- **Excluir casos según lista.** Cada prueba  $t$  utiliza sólo aquellos casos que contienen datos válidos para todas las variables utilizadas en las pruebas  $t$  solicitadas. El tamaño muestral es constante en todas las pruebas.

## Funciones adicionales del comando T-TEST

Con el lenguaje de sintaxis de comandos también podrá:

- Producir pruebas  $t$  tanto de una sola muestra como de muestras independientes ejecutando un solo comando.
- Contrastar una variable con todas las variables de una lista, en una prueba relacionada (mediante el subcomando `PAIRS`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

## ***ANOVA de un factor***

El procedimiento ANOVA de un factor genera un análisis de varianza de un factor para una variable dependiente cuantitativa respecto a una única variable de factor (la variable independiente). El análisis de varianza se utiliza para contrastar la hipótesis de que varias medias son iguales. Esta técnica es una extensión de la prueba  $t$  para dos muestras.

Además de determinar que existen diferencias entre las medias, es posible que desee saber qué medias difieren. Existen dos tipos de contrastes para comparar medias: a priori y post hoc. Los contrastes a priori se plantean *antes* de ejecutar el experimento y los contrastes post hoc se realizan *después* de haber llevado a cabo el experimento. También puede contrastar las tendencias existentes a través de las categorías.

**Ejemplo.** Las rosquillas absorben diferentes cantidades de grasa cuando se fríen. Se plantea un experimento utilizando tres tipos de grasas: aceite de cacahuete, aceite de maíz y manteca de cerdo. El aceite de cacahuete y el aceite de maíz son grasas no saturadas y la manteca es una grasa saturada. Además de determinar si la cantidad de grasa absorbida depende del tipo de grasa utilizada, también se podría preparar un contraste a priori para determinar si la cantidad de absorción de la grasa difiere para las grasas saturadas y las no saturadas.

**Estadísticos.** Para cada grupo: número de casos, media, desviación típica, error típico de la media, mínimo, máximo, intervalo de confianza al 95% para la media. Prueba de Levene sobre la homogeneidad de varianzas, tabla de análisis de varianza y contrastes robustos de igualdad de medias para cada variable dependiente, contrastes a priori especificados por el usuario y las pruebas de rango y de comparaciones múltiples post hoc: Bonferroni, Sidak, diferencia honestamente significativa de Tukey, GT2 de Hochberg, Gabriel, Dunnett, prueba  $F$  de Ryan-Einot-Gabriel-Welsch, (R-E-G-W  $F$ ), prueba de rango de Ryan-Einot-Gabriel-Welsch (R-E-G-W  $Q$ ), T2 de Tamhane, T3 de Dunnett, Games-Howell,  $C$ , de Dunnett, prueba de rango múltiple de Duncan, Student-Newman-Keuls (S-N-K),  $b$  de Tukey, Waller-Duncan, Scheffé y diferencia mínima significativa.

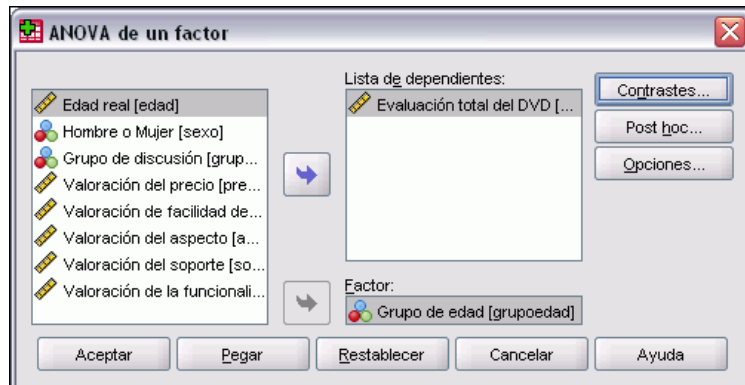
**Datos.** Los valores de la variable de factor deben ser enteros y la variable dependiente debe ser cuantitativa (nivel de medida de intervalo).

**Supuestos.** Cada grupo es una muestra aleatoria independiente procedente de una población normal. El análisis de varianza es robusto a las desviaciones de la normalidad, aunque los datos deberán ser simétricos. Los grupos deben proceder de poblaciones con varianzas iguales. Para contrastar este supuesto, utilice la prueba de Levene de homogeneidad de varianzas.

### Para obtener un análisis de varianza de un factor

- ▶ Elija en los menús:
  - Analizar
  - Comparar medias
  - ANOVA de un factor...

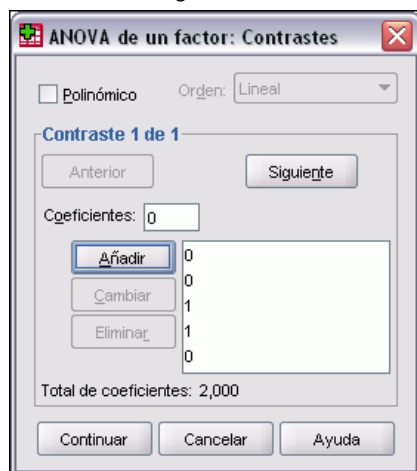
Figura 21-1  
Cuadro de diálogo ANOVA de un factor



- ▶ Seleccione una o más variables dependientes.
- ▶ Seleccione una sola variable de factor independiente.

## ANOVA de un factor: Contrastes

Figura 21-2  
Cuadro de diálogo ANOVA de un factor: Contrastes



Puede dividir las sumas de cuadrados inter-grupos en componentes de tendencia o especificar contrastes a priori.



**Polinómico.** Divide las sumas de cuadrados inter-grupos en componentes de tendencia. Puede contrastar la existencia de tendencia en la variable dependiente a través de los niveles ordenados de la variable de factor. Por ejemplo, podría contrastar si existe una tendencia lineal (creciente o decreciente) en el salario, a través de los niveles ordenados de la titulación mayor obtenida.

■ **Orden.** Se puede elegir un orden polinómico 1°, 2°, 3°, 4° o 5°.

**Coefficientes.** Contrastes a priori especificados por el usuario que serán contrastados mediante el estadístico  $t$ . Introduzca un coeficiente para cada grupo (categoría) de la variable factor y pulse en Añadir después de cada entrada. Cada nuevo valor se añade al final de la lista de coeficientes. Para especificar conjuntos de contrastes adicionales, pulse en Siguiente. Utilice Siguiente y Anterior para desplazarse por los conjuntos de contrastes.

El orden de los coeficientes es importante porque se corresponde con el orden ascendente de los valores de las categorías de la variable de factor. El primer coeficiente en la lista se corresponde con el menor de los valores de grupo en la variable de factor y el último coeficiente se corresponde con el valor más alto. Por ejemplo, si existen seis categorías en la variable factor, los coeficientes -1, 0, 0, 0,5 y 0,5 contrastan el primer grupo con los grupos quinto y sexto. Para la mayoría de las aplicaciones, la suma de los coeficientes debería ser 0. Los conjuntos que no sumen 0 también se pueden utilizar, pero aparecerá un mensaje de advertencia.

## ANOVA de un factor: Contrastes post hoc

Figura 21-3

Cuadro de diálogo ANOVA de un factor: Comparaciones múltiples post hoc

Una vez que se ha determinado que existen diferencias entre las medias, las pruebas de rango post hoc y las comparaciones múltiples por parejas permiten determinar qué medias difieren. Las pruebas de rango identifican subconjuntos homogéneos de medias que no se diferencian entre sí. Las comparaciones múltiples por parejas contrastan la diferencia entre cada pareja de medias y generan una matriz donde los asteriscos indican las medias de grupo significativamente diferentes a un nivel alfa de 0,05.

**Asumiendo varianzas iguales**

La prueba de la diferencia honestamente significativa de Tukey, la GT2 de Hochberg, la prueba de Gabriel y la prueba de Scheffé son pruebas de comparaciones múltiples y pruebas de rango. Otras pruebas de rango disponibles son la *b* de Tukey, S-N-K (Student-Newman-Keuls), Duncan, R-E-G-W *F* (prueba *F* de Ryan-Einot-Gabriel-Welsch), R-E-G-W *Q* (prueba de rango de Ryan-Einot-Gabriel-Welsch) y Waller-Duncan. Las pruebas de comparaciones múltiples disponibles son Bonferroni, prueba de diferencia honestamente significativa de Tukey, Sidak, Gabriel, Hochberg, Dunnett, Scheffé, y DMS (diferencia mínima significativa).

- **LSD.** Utiliza pruebas *t* para llevar a cabo todas las comparaciones por pares entre las medias de los grupos. No se efectúa ninguna corrección de la tasa de error para el hecho de realizar múltiples comparaciones.
- **Bonferroni.** Utiliza las pruebas de *t* para realizar comparaciones por pares entre las medias de los grupos, pero controla la tasa de error global estableciendo que la tasa de error de cada prueba sea igual a la tasa de error por experimento dividida entre el número total de contrastes. Así, se corrige el nivel crítico por el hecho de que se están realizando múltiples comparaciones.
- **Sidak.** Prueba de comparaciones múltiples por parejas basada en un estadístico *t*. La prueba de Sidak corrige el nivel de significación para las comparaciones múltiples y da lugar a límites más estrechos que los de Bonferroni.
- **Scheffe.** Realiza comparaciones conjuntas simultáneas por pares para todas las posibles parejas de combinaciones de las medidas. Usa la distribución muestral *F*. Puede utilizarse para examinar todas las combinaciones lineales de grupos de medias posibles, no sólo las comparaciones por parejas.
- **R-E-G-W *F*.** Procedimiento múltiple por pasos (por tamaño de las distancias) de Ryan-Einot-Gabriel-Welsch que se basa en una prueba *F*.
- **R-E-G-W *Q*.** Procedimiento múltiple por pasos (por tamaño de las distancias) de Ryan-Einot-Gabriel-Welsch que se basa en el rango estudentizado.
- **S-N-K.** Realiza todas las comparaciones por parejas entre las medias utilizando la distribución del rango de Student. Con tamaños de grupo iguales, también compara pares de medias dentro de subconjuntos homogéneos utilizando un procedimiento por pasos. Las medias se ordenan de mayor a menor y se contrastan en primer lugar las diferencias extremas.
- **Tukey.** Utiliza el estadístico del rango estudentizado para realizar todas las comparaciones por pares entre los grupos. Establece la tasa de error por experimento como la tasa de error para el conjunto de todas las comparaciones por pares.
- **Tukey-*b*.** Prueba que emplea la distribución del rango estudentizado para realizar comparaciones por pares entre los grupos. El valor crítico es el promedio de los valores correspondientes a la diferencia honestamente significativa de Tukey y al método de Student-Newman-Keuls.
- **Duncan.** Realiza comparaciones por pares utilizando un orden por pasos idéntico al orden usado por la prueba de Student-Newman-Keuls, pero establece un nivel de protección en la tasa de error para la colección de contrastes, en lugar de usar una tasa de error para los contrastes individuales. Utiliza el estadístico del rango estudentizado.
- **GT2 de Hochberg.** Prueba de comparaciones múltiples y de rango que utiliza el módulo máximo estudentizado. Es similar a la prueba de la diferencia honestamente significativa de Tukey.

- **Gabriel.** Prueba de comparación por parejas que utiliza el módulo máximo estudentizado y que es generalmente más potente que la GT2 de Hochberg, cuando los tamaños de las casillas son desiguales. La prueba de Gabriel se puede convertir en liberal cuando los tamaños de las casillas varían mucho.
- **Waller-Duncan.** Prueba de comparaciones múltiples basada en un estadístico t. Utiliza la aproximación Bayesiana.
- **Dunnnett.** Prueba de comparaciones múltiples por parejas que compara un conjunto de tratamientos respecto a una única media de control. La última categoría es la categoría de control por defecto. Si lo desea, puede seleccionar la primera categoría. Para comprobar que la media de cualquier nivel del factor (excepto la categoría de control) no es igual a la de la categoría de control, utilice una prueba bilateral. Para contrastar si la media en cualquier nivel del factor es menor que la de la categoría de control, seleccione <Control. Para contrastar si la media en cualquier nivel del factor es mayor que la de la categoría de control, seleccione Control.

### ***No asumiendo varianzas iguales***

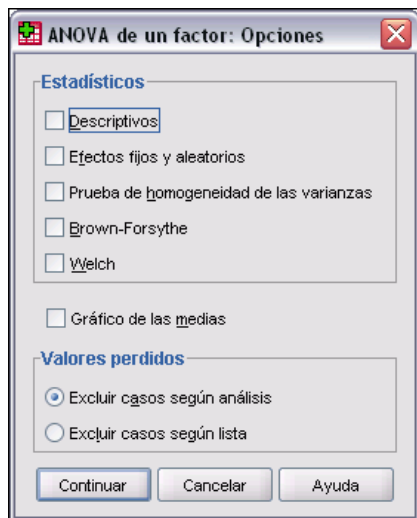
Las pruebas de comparaciones múltiples que no suponen varianzas iguales son T2 de Tamhane, T3 de Dunnnett, Games-Howell y C de Dunnnett.

- **T2 de Tamhane.** Prueba conservadora de comparación por parejas basada en la prueba t. Esta prueba es adecuada cuando las varianzas son desiguales.
- **T3 de Dunnnett.** Prueba de comparación por parejas basada en el módulo máximo estudentizado. Esta prueba es adecuada cuando las varianzas son desiguales.
- **Games-Howell.** Prueba de comparación por parejas que es en ocasiones liberal. Esta prueba es adecuada cuando las varianzas son desiguales.
- **C de Dunnnett.** Prueba de comparación por parejas basada en el rango estudentizado. Esta prueba es adecuada cuando las varianzas son desiguales.

*Nota:* Posiblemente le resulte más fácil interpretar el resultado de los contrastes post hoc si desactiva Ocultar filas y columnas vacías en el cuadro de diálogo Propiedades de tabla (en una tabla pivote activada, seleccione Propiedades de tabla en el menú Formato).

## ANOVA de un factor: Opciones

Figura 21-4  
Cuadro de diálogo ANOVA de un factor: Opciones



**Estadísticos.** Elija uno o más entre los siguientes:

- **Descriptivos.** Calcula los siguientes estadísticos: Número de casos, Media, Desviación típica, Error típico de la media, Mínimo, Máximo y los Intervalos de confianza al 95% de cada variable dependiente para cada grupo.
- **Efectos fijos y aleatorios.** Muestra la desviación típica, el error típico y un intervalo de confianza del 95% para el modelo de efectos fijos, y el error típico, un intervalo de confianza del 95% y una estimación de la varianza entre componentes para el modelo de efectos aleatorios.
- **Prueba de homogeneidad de las varianzas.** Calcula el estadístico de Levene para contrastar la igualdad de las varianzas de grupo. Esta prueba no depende del supuesto de normalidad.
- **Brown-Forsythe.** Calcula el estadístico de Brown-Forsythe para contrastar la igualdad de las medias de grupo. Este estadístico es preferible al estadístico  $F$  si no se supone la igualdad de las varianzas.
- **Welch.** Calcula el estadístico de Welch para contrastar la igualdad de las medias de grupo. Este estadístico es preferible al estadístico  $F$  si no se supone la igualdad de las varianzas.

**Gráfico de las medias.** Muestra un gráfico que representa las medias de los subgrupos (las medias para cada grupo definido por los valores de la variable factor).

**Valores perdidos.** Controla el tratamiento de los valores perdidos.

- **Excluir casos según análisis.** Un caso que tenga un valor perdido para la variable dependiente o la variable de factor en un análisis determinado, no se utiliza en ese análisis. Además, los casos fuera del rango especificado para la variable de factor no se utilizan.
- **Excluir casos según lista.** Se excluyen de todos los análisis los casos con valores perdidos para la variable de factor o para cualquier variable dependiente incluida en la lista de variables dependientes en el cuadro de diálogo principal. Si no se han especificado varias variables dependientes, esta opción no surte efecto.

## ***Funciones adicionales del comando ONEWAY***

Con el lenguaje de sintaxis de comandos también podrá:

- Obtener estadísticos de efectos fijos y aleatorios. Desviación típica, error típico de la media e intervalos de confianza al 95% para el modelo de efectos fijos. Error típico, intervalos de confianza al 95% y estimación de la varianza entre componentes para el modelo de efectos aleatorios (mediante `STATISTICS=EFFECTS`).
- Especificar niveles alfa para la diferencia mínima significativa, Bonferroni, pruebas de comparación múltiple de Duncan y Scheffé (con el subcomando `RANGES`).
- Escribir una matriz de medias, desviaciones típicas y frecuencias, o leer una matriz de medias, frecuencias, varianzas combinadas y grados de libertad para las varianzas combinadas. Estas matrices pueden utilizarse en lugar de los datos brutos para obtener un análisis de varianza de un factor (con el subcomando `MATRIX`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

## ***MLG Análisis univariante***

El procedimiento MLG Univariante proporciona un análisis de regresión y un análisis de varianza para una variable dependiente mediante uno o más factores o variables. Las variables de factor dividen la población en grupos. Con el procedimiento Modelo lineal general se pueden contrastar hipótesis nulas sobre los efectos de otras variables en las medias de varias agrupaciones de una única variable dependiente. Se pueden investigar las interacciones entre los factores así como los efectos de los factores individuales, algunos de los cuales pueden ser aleatorios. Además, se pueden incluir los efectos de las covariables y las interacciones de covariables con los factores. Para el análisis de regresión, las variables (predictoras) independientes se especifican como covariables.

Se pueden contrastar tanto los modelos equilibrados como los no equilibrados. Se considera que un diseño está equilibrado si cada casilla del modelo contiene el mismo número de casos. Además de contrastar hipótesis, MLG Univariante genera estimaciones de los parámetros.

También se encuentran disponibles los contrastes a priori de uso más habitual para contrastar las hipótesis. Además, si una prueba  $F$  global ha mostrado cierta significación, pueden emplearse las pruebas post hoc para evaluar las diferencias entre las medias específicas. Las medias marginales estimadas ofrecen estimaciones de valores de las medias pronosticados para las casillas del modelo; los gráficos de perfil (gráficos de interacciones) de estas medias permiten observar fácilmente algunas de estas relaciones.

En su archivo de datos puede guardar residuos, valores pronosticados, distancia de Cook y valores de influencia como variables nuevas para comprobar los supuestos.

Ponderación MCP permite especificar una variable usada para aplicar a las observaciones una ponderación diferente en un análisis de mínimos cuadrados ponderados (MCP), por ejemplo para compensar la distinta precisión de las medidas.

**Ejemplo.** Se recogen datos de los corredores individuales en el maratón de Chicago durante varios años. El tiempo final de cada corredor es la variable dependiente. Influyen otros factores como el clima (frío, calor o temperatura agradable), los meses de entrenamiento, el número de maratones anteriores y el sexo. La edad se considera una covariable. Observará que el sexo es un efecto significativo y que la interacción del sexo con el clima es significativa.

**Métodos.** Las sumas de cuadrados de Tipo I, Tipo II, Tipo III y Tipo IV pueden emplearse para evaluar las diferentes hipótesis. Tipo III es el valor por defecto.

**Estadísticos.** Las pruebas de rango post hoc y las comparaciones múltiples: Diferencia menos significativa (DMS), Bonferroni, Sidak, Scheffé, Múltiples  $F$  de Ryan-Einot-Gabriel-Welsch (R-E-G-W-F), Rango múltiple de Ryan-Einot-Gabriel-Welsch, Student-Newman-Keuls (S-N-K), Diferencia honestamente significativa de Tukey,  $b$  de Tukey, Duncan, GT2 de Hochberg, Gabriel, Pruebas  $t$  de Waller Duncan, Dunnett (unilateral y bilateral), T2 de Tamhane, T3 de Dunnett, Games-Howell y  $C$  de Dunnett. Estadísticos descriptivos: medias observadas, desviaciones

típicas y frecuencias de todas las variables dependientes en todas las casillas. Prueba de Levene para la homogeneidad de varianzas.

**Diagramas.** Diagramas de dispersión por nivel, gráficos de residuos, gráficos de perfil (interacción).

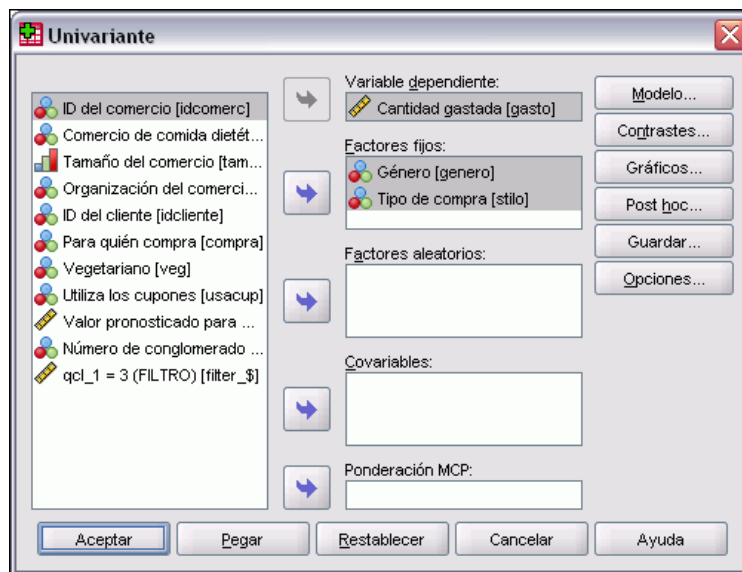
**Datos.** La variable dependiente es cuantitativa. Los factores son categóricos; pueden tener valores numéricos o valores de cadena de hasta ocho caracteres. Pueden tener valores numéricos o valores de cadena de hasta ocho caracteres. Las covariables son variables cuantitativas que están relacionadas con la variable dependiente.

**Supuestos.** Los datos son una muestra aleatoria de una población normal; en la población, todas las varianzas de las casillas son iguales. El análisis de varianza es robusto a las desviaciones de la normalidad, aunque los datos deberán ser simétricos. Para comprobar los supuestos, puede utilizar la prueba de homogeneidad de varianzas y los gráficos de dispersión por nivel. También puede examinar los residuos y los gráficos de residuos.

### Para obtener un análisis MLG Univariante

- ▶ Elija en los menús:
  - Analizar
  - Modelo lineal general
  - Univariante...

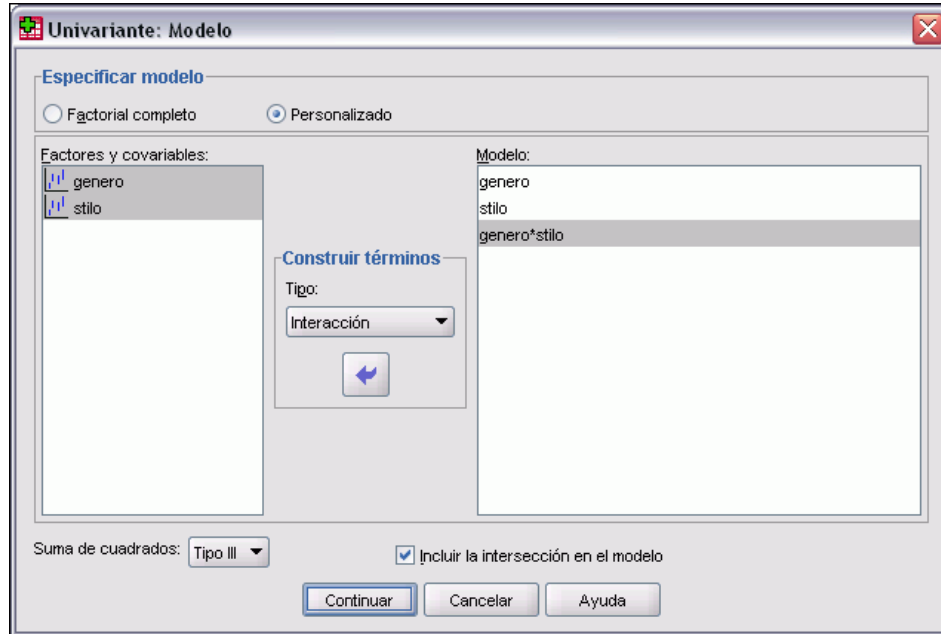
Figura 22-1  
GLM Univariate dialog box



- ▶ Seleccione una variable dependiente.
- ▶ Seleccione variables para Factores fijos, Factores aleatorios y Covariables, en función de los datos.
- ▶ Si lo desea, puede utilizar la Ponderación MCP para especificar una variable de ponderación para el análisis de mínimos cuadrados ponderados. Si el valor de la variable de ponderación es cero, negativo o perdido, el caso queda excluido del análisis. Una variable que ya se haya utilizado en el modelo no puede usarse como variable de ponderación.

## MLG: Modelo

Figura 22-2  
Cuadro de diálogo Univariante: Modelo



**Especificar modelo.** Un modelo factorial completo contiene todos los efectos principales del factor, todos los efectos principales de las covariables y todas las interacciones factor por factor. No contiene interacciones de covariable. Seleccione Personalizado para especificar sólo un subconjunto de interacciones o para especificar interacciones factor por covariable. Indique todos los términos que desee incluir en el modelo.

**Factores y Covariables.** Muestra una lista de los factores y las covariables.

**Modelo.** El modelo depende de la naturaleza de los datos. Después de seleccionar Personalizado, puede elegir los efectos principales y las interacciones que sean de interés para el análisis.

**Suma de cuadrados** Determina el método para calcular las sumas de cuadrados. Para los modelos equilibrados y no equilibrados sin casillas perdidas, el método más utilizado para la suma de cuadrados es el Tipo III.

**Incluir la intersección en el modelo.** La intersección se incluye normalmente en el modelo. Si supone que los datos pasan por el origen, puede excluir la intersección.

### Construir términos

Para las covariables y los factores seleccionados:

**Interacción.** Crea el término de interacción de mayor nivel con todas las variables seleccionadas. Este es el método por defecto.

**Efectos principales.** Crea un término de efectos principales para cada variable seleccionada.

**Todos de 2.** Crea todas las interacciones dobles posibles de las variables seleccionadas.



**Todas de 3.** Crea todas las interacciones triples posibles de las variables seleccionadas.

**Todas de 4.** Crea todas las interacciones cuádruples posibles de las variables seleccionadas.

**Todas de 5.** Crea todas las interacciones quintuples posibles de las variables seleccionadas.

## **Suma de cuadrados**

Para el modelo, puede elegir un tipo de suma de cuadrados. El Tipo III es el más utilizado y es el tipo por defecto.

**Tipo I.** Este método también se conoce como el método de descomposición jerárquica de la suma de cuadrados. Cada término se corrige sólo respecto al término que le precede en el modelo. El método Tipo I para la obtención de sumas de cuadrados se utiliza normalmente para:

- Un modelo ANOVA equilibrado en el que se especifica cualquier efecto principal antes de cualquier efecto de interacción de primer orden, cualquier efecto de interacción de primer orden se especifica antes de cualquier efecto de interacción de segundo orden, y así sucesivamente.
- Un modelo de regresión polinómica en el que se especifica cualquier término de orden inferior antes que cualquier término de orden superior.
- Un modelo puramente anidado en el que el primer efecto especificado está anidado dentro del segundo efecto especificado, el segundo efecto especificado está anidado dentro del tercero, y así sucesivamente. Esta forma de anidamiento solamente puede especificarse utilizando la sintaxis.

**Tipo II.** Este método calcula cada suma de cuadrados del modelo considerando sólo los efectos pertinentes. Un efecto pertinente es el que corresponde a todos los efectos que no contienen el que se está examinando. El método Tipo II para la obtención de sumas de cuadrados se utiliza normalmente para:

- Un modelo ANOVA equilibrado.
- Cualquier modelo que sólo tenga efectos de factor principal.
- Cualquier modelo de regresión.
- Un diseño puramente anidado (esta forma de anidamiento solamente puede especificarse utilizando la sintaxis).

**Tipo III.** Es el método por defecto. Este método calcula las sumas de cuadrados de un efecto del diseño como las sumas de cuadrados corregidas respecto a cualquier otro efecto que no lo contenga y ortogonales a cualquier efecto (si existe) que lo contenga. Las sumas de cuadrados de Tipo III tienen una gran ventaja por ser invariables respecto a las frecuencias de casilla, siempre que la forma general de estimabilidad permanezca constante. Así, este tipo de sumas de cuadrados se suele considerar de gran utilidad para un modelo no equilibrado sin casillas perdidas. En un diseño factorial sin casillas perdidas, este método equivale a la técnica de cuadrados ponderados de las medias de Yates. El método Tipo III para la obtención de sumas de cuadrados se utiliza normalmente para:

- Cualquiera de los modelos que aparecen en los tipos I y II.
- Cualquier modelo equilibrado o desequilibrado sin casillas vacías.

**Tipo IV.** Este método está diseñado para una situación en la que hay casillas perdidas. Para cualquier efecto  $F$  en el diseño, si  $F$  no está contenida en cualquier otro efecto, entonces Tipo IV = Tipo III = Tipo II. Cuando  $F$  está contenida en otros efectos, el Tipo IV distribuye equitativamente los contrastes que se realizan entre los parámetros en  $F$  a todos los efectos de nivel superior. El método Tipo IV para la obtención de sumas de cuadrados se utiliza normalmente para:

- Cualquiera de los modelos que aparecen en los tipos I y II.
- Cualquier modelo equilibrado o no equilibrado con casillas vacías.

## MLG: Contrastes

Figura 22-3  
Cuadro de diálogo Univariante: Contrastes



Los contrastes se utilizan para contrastar las diferencias entre los niveles de un factor. Puede especificar un contraste para cada factor en el modelo (en un modelo de medidas repetidas, para cada factor inter-sujetos). Los contrastes representan las combinaciones lineales de los parámetros.

El contraste de hipótesis se basa en la hipótesis nula  $\mathbf{LB} = 0$ , donde  $\mathbf{L}$  es la matriz de coeficientes de contraste y  $\mathbf{B}$  es el vector de parámetros. Cuando se especifica un contraste, se crea una matriz  $\mathbf{L}$ . Las columnas de la matriz  $\mathbf{L}$  correspondientes al factor coinciden con el contraste. El resto de las columnas se corrigen para que la matriz  $\mathbf{L}$  sea estimable.

Los resultados incluyen un estadístico  $F$  para cada conjunto de contrastes. Para el contraste de diferencias también se muestran los intervalos de confianza simultáneos de tipo Bonferroni basados en la distribución  $t$  de Student.

### Contrastes disponibles

Los contrastes disponibles son de desviación, simples, de diferencias, de Helmert, repetidos y polinómicos. En los contrastes de desviación y los contrastes simples, es posible determinar que la categoría de referencia sea la primera o la última categoría.

### Tipos de contrastes

**Desviación.** Compara la media de cada nivel (excepto una categoría de referencia) con la media de todos los niveles (media global). Los niveles del factor pueden colocarse en cualquier orden.

**Simple.** Compara la media de cada nivel con la media de un nivel especificado. Este tipo de contraste resulta útil cuando existe un grupo de control. Puede seleccionar la primera o la última categoría como referencia.

**Diferencia.** Compara la media de cada nivel (excepto el primero) con la media de los niveles anteriores (a veces también se denominan contrastes de Helmert inversos). (a veces también se denominan contrastes de Helmert inversos).

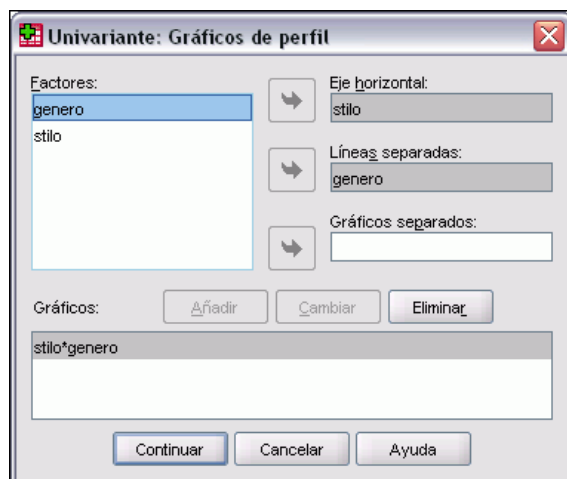
**Helmert.** Compara la media de cada nivel del factor (excepto el último) con la media de los niveles siguientes.

**Repetidas.** Compara la media de cada nivel (excepto el último) con la media del nivel siguiente.

**Polinómico.** Compara el efecto lineal, cuadrático, cúbico, etc. El primer grado de libertad contiene el efecto lineal a través de todas las categorías; el segundo grado de libertad, el efecto cuadrático, y así sucesivamente. Estos contrastes se utilizan a menudo para estimar las tendencias polinómicas.

## MLG: Gráficos de perfil

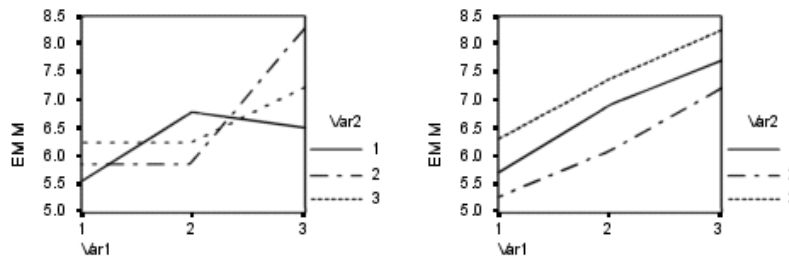
Figura 22-4  
Cuadro de diálogo Univariante: Gráficos de perfil



Los gráficos de perfil (gráficos de interacción) sirven para comparar las medias marginales en el modelo. Un gráfico de perfil es un gráfico de líneas en el que cada punto indica la media marginal estimada de una variable dependiente (corregida respecto a las covariables) en un nivel de un factor. Los niveles de un segundo factor se pueden utilizar para generar líneas diferentes. Cada nivel en un tercer factor se puede utilizar para crear un gráfico diferente. Todos los factores fijos y aleatorios, si existen, están disponibles para los gráficos. Para los análisis multivariantes, los gráficos de perfil se crean para cada variable dependiente. En un análisis de medidas repetidas, es posible utilizar tanto los factores inter-sujetos como los intra-sujetos en los gráficos de perfil. Las opciones MLG - Multivariante y MLG - Medidas repetidas sólo estarán disponibles si tiene instalada la opción Modelos avanzados.

Un gráfico de perfil de un factor muestra si las medias marginales estimadas aumentan o disminuyen a través de los niveles. Para dos o más factores, las líneas paralelas indican que no existe interacción entre los factores, lo que significa que puede investigar los niveles de un único factor. Las líneas no paralelas indican una interacción.

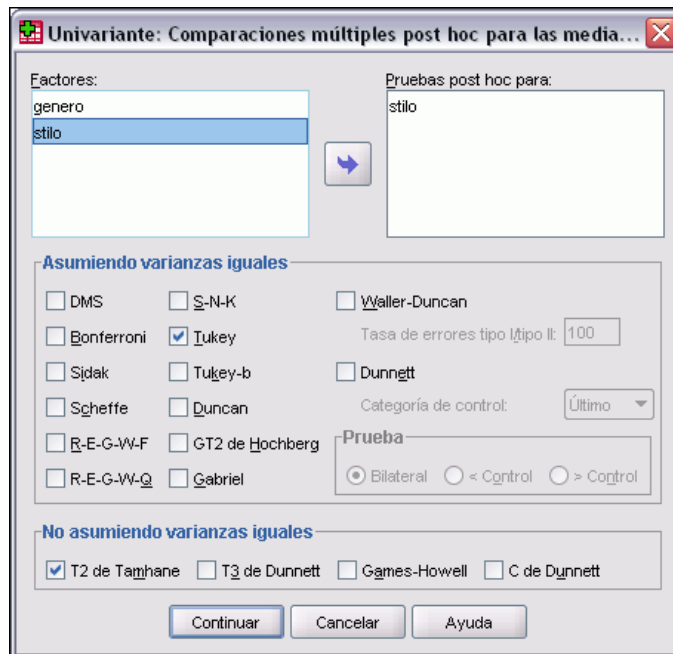
Figura 22-5  
Gráfico no paralelo (izquierda) y gráfico paralelo (derecha)



Después de especificar un gráfico mediante la selección de los factores del eje horizontal y, de manera opcional, los factores para distintas líneas y gráficos, el gráfico deberá añadirse a la lista de gráficos.

## MLG: Comparaciones post hoc

Figura 22-6  
Cuadro de diálogo Post hoc



**Pruebas de comparaciones múltiples post hoc** Una vez que se ha determinado que existen diferencias entre las medias, las pruebas de rango post hoc y las comparaciones múltiples por parejas permiten determinar qué medias difieren. Las comparaciones se realizan sobre valores sin corregir. Estas pruebas se utilizan únicamente para factores inter-sujetos fijos. En MLG Medidas

repetidas, estas pruebas no están disponibles si no existen factores inter-sujetos y las pruebas de comparación múltiple post hoc se realizan para la media a través de los niveles de los factores intra-sujetos. Para MLG - Multivariante, las pruebas post hoc se realizan por separado para cada variable dependiente. Las opciones MLG - Multivariante y MLG - Medidas repetidas sólo estarán disponibles si tiene instalada la opción Modelos avanzados.

Las pruebas de diferencia honestamente significativa de Tukey y de Bonferroni son pruebas de comparación múltiple muy utilizadas. La **prueba de Bonferroni**, basada en el estadístico  $t$  de Student, corrige el nivel de significación observado por el hecho de que se realizan comparaciones múltiples. La **prueba  $t$  de Sidak** también corrige el nivel de significación y da lugar a límites más estrechos que los de Bonferroni. La **prueba de diferencia honestamente significativa de Tukey** utiliza el estadístico del rango estudentizado para realizar todas las comparaciones por pares entre los grupos y establece la tasa de error por experimento como la tasa de error para el conjunto de todas las comparaciones por pares. Cuando se contrasta un gran número de pares de medias, la prueba de la diferencia honestamente significativa de Tukey es más potente que la prueba de Bonferroni. Para un número reducido de pares, Bonferroni es más potente.

**GT2 de Hochberg** es similar a la prueba de la diferencia honestamente significativa de Tukey, pero se utiliza el módulo máximo estudentizado. La prueba de Tukey suele ser más potente. La **prueba de comparación por parejas de Gabriel** también utiliza el módulo máximo estudentizado y es generalmente más potente que la GT2 de Hochberg cuando los tamaños de las casillas son desiguales. La prueba de Gabriel se puede convertir en liberal cuando los tamaños de las casillas varían mucho.

La **prueba  $t$  de comparación múltiple por parejas de Dunnett** compara un conjunto de tratamientos con una media de control simple. La última categoría es la categoría de control por defecto. Si lo desea, puede seleccionar la primera categoría. Asimismo, puede elegir una prueba unilateral o bilateral. Para comprobar que la media de cualquier nivel del factor (excepto la categoría de control) no es igual a la de la categoría de control, utilice una prueba bilateral. Para contrastar si la media en cualquier nivel del factor es menor que la de la categoría de control, seleccione  $<$  Control. Asimismo, para contrastar si la media en cualquier nivel del factor es mayor que la de la categoría de control, seleccione  $>$  Control.

Ryan, Einot, Gabriel y Welsch (R-E-G-W) desarrollaron dos pruebas de rangos múltiples por pasos. Los procedimientos múltiples por pasos (por tamaño de las distancias) contrastan en primer lugar si todas las medias son iguales. Si no son iguales, se contrasta la igualdad en los subconjuntos de medias. **R-E-G-W F** se basa en una prueba  $F$  y **R-E-G-W Q** se basa en un rango estudentizado. Estas pruebas son más potentes que la prueba de rangos múltiples de Duncan y Student-Newman-Keuls (que también son procedimientos múltiples por pasos), pero no se recomiendan para tamaños de casillas desiguales.

Cuando las varianzas son desiguales, utilice **T2 de Tamhane** (prueba conservadora de comparación por parejas basada en una prueba  $t$ ), **T3 de Dunnett** (prueba de comparación por parejas basada en el módulo máximo estudentizado), **prueba de comparación por parejas Games-Howell** (a veces liberal), o **C de Dunnett** (prueba de comparación por parejas basada en el rango estudentizado).

La **prueba de rango múltiple de Duncan**, Student-Newman-Keuls (S-N-K) y **b de Tukey** son pruebas de rango que asignan rangos a medias de grupo y calculan un valor de rango. Estas pruebas no se utilizan con la misma frecuencia que las pruebas anteriormente mencionadas.

La **prueba  $t$  de Waller-Duncan** utiliza la aproximación bayesiana. Esta prueba de rango emplea la media armónica del tamaño muestral cuando los tamaños muestrales no son iguales.

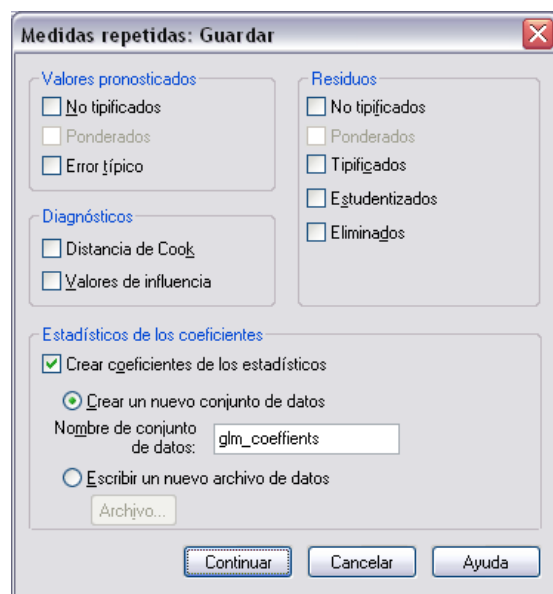
El nivel de significación de la prueba de **Scheffé** está diseñado para permitir todas las combinaciones lineales posibles de las medias de grupo que se van a contrastar, no sólo las comparaciones por parejas disponibles en esta función. El resultado es que la prueba de Scheffé es normalmente más conservadora que otras pruebas, lo que significa que se precisa una mayor diferencia entre las medias para la significación.

La prueba de comparación múltiple por parejas de la diferencia menos significativa (**DMS**) es equivalente a varias pruebas *t* individuales entre todos los pares de grupos. La desventaja de esta prueba es que no se realiza ningún intento de corregir el nivel crítico para realizar las comparaciones múltiples.

**Pruebas mostradas.** Se proporcionan comparaciones por parejas para DMS, Sidak, Bonferroni, Games-Howell, T2 y T3 de Tamhane, *C* de Dunnett y T3 de Dunnett. También se facilitan subconjuntos homogéneos para S-N-K, *b* de Tukey, Duncan, R-E-G-W *F*, R-E-G-W *Q* y Waller. La prueba de la diferencia honestamente significativa de Tukey, la GT2 de Hochberg, la prueba de Gabriel y la prueba de Scheffé son pruebas de comparaciones múltiples y pruebas de rango.

## MLG: Guardar

Figura 22-7  
Cuadro de diálogo Guardar



Es posible guardar los valores pronosticados por el modelo, los residuos y las medidas relacionadas como variables nuevas en el Editor de datos. Muchas de estas variables se pueden utilizar para examinar supuestos sobre los datos. Si desea almacenar los valores para utilizarlos en otra sesión de SPSS, guárdelos en el archivo de datos actual.

**Valores pronosticados.** Son los valores que predice el modelo para cada caso.

- **No tipificados.** Valor pronosticado por el modelo para la variable dependiente.

- **Ponderados.** Los valores pronosticados no tipificados ponderados. Sólo están disponibles si se seleccionó previamente una variable de ponderación WLS.
- **Error típico.** Estimación de la desviación típica del valor promedio de la variable dependiente para los casos que tengan los mismos valores en las variables independientes.

**Diagnósticos.** Son medidas para identificar casos con combinaciones poco usuales de valores para los casos y las variables independientes que puedan tener un gran impacto en el modelo.

- **Distancia de Cook.** Medida de cuánto cambiarían los residuos de todos los casos si se excluyera un caso determinado del cálculo de los coeficientes de regresión. Una Distancia de Cook grande indica que la exclusión de ese caso del cálculo de los estadísticos de regresión hará variar substancialmente los coeficientes.
- **Valores de influencia.** Los valores de influencia no centrados. La influencia relativa de una observación en el ajuste del modelo.

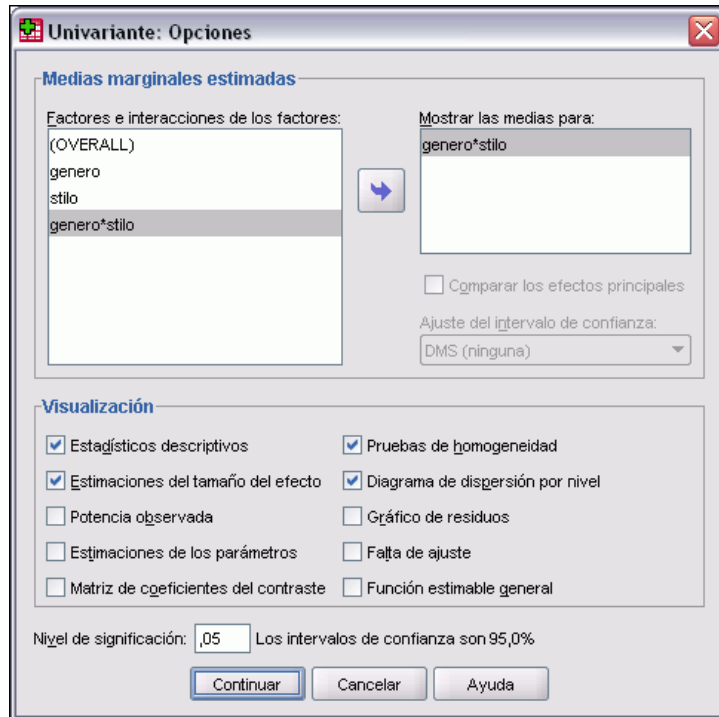
**Residuos.** Un residuo no tipificado es el valor real de la variable dependiente menos el valor pronosticado por el modelo. También se encuentran disponibles residuos eliminados, estudentizados y tipificados. Si ha seleccionado una variable MCP, contará además con residuos no tipificados ponderados.

- **No tipificados.** Diferencia entre un valor observado y el valor pronosticado por el modelo.
- **Ponderados.** Los residuos no tipificados ponderados. Sólo están disponibles si se seleccionó previamente una variable de ponderación WLS.
- **Tipificados.** El residuo dividido por una estimación de su error típico. Los residuos tipificados, que son conocidos también como los residuos de Pearson o residuos estandarizados, tienen una media de 0 y una desviación típica de 1.
- **Método de Student.** Residuo dividido por una estimación de su desviación típica que varía de caso en caso, dependiendo de la distancia de los valores de cada caso en las variables independientes respecto a las medias en las variables independientes.
- **Eliminados.** Residuo para un caso cuando éste se excluye del cálculo de los coeficientes de la regresión. Es igual a la diferencia entre el valor de la variable dependiente y el valor pronosticado corregido.

**Estadísticos de los coeficientes.** Escribe una matriz varianza-covarianza de las estimaciones de los parámetros del modelo en un nuevo conjunto de datos de la sesión actual o un archivo de datos externo con formato SPSS. Asimismo, para cada variable dependiente habrá una fila de estimaciones de los parámetros, una fila de valores de significación para los estadísticos  $t$  correspondientes a las estimaciones de los parámetros y una fila de grados de libertad de los residuos. En un modelo multivariante, existen filas similares para cada variable dependiente. Si lo desea, puede usar este archivo matricial en otros procedimientos que lean archivos matriciales.

## Opciones MLG

Figura 22-8  
Cuadro de diálogo Opciones



Este cuadro de diálogo contiene estadísticos opcionales. Los estadísticos se calculan utilizando un modelo de efectos fijos.

**Medias marginales estimadas.** Seleccione los factores e interacciones para los que desee obtener estimaciones de las medias marginales de la población en las casillas. Estas medias se corrigen respecto a las covariables, si las hay.

- **Comparar los efectos principales.** Proporciona comparaciones por parejas no corregidas entre las medias marginales estimadas para cualquier efecto principal del modelo, tanto para los factores inter-sujetos como para los intra-sujetos. Este elemento sólo se encuentra disponible si los efectos principales están seleccionados en la lista Mostrar las medias para.
- **Ajuste del intervalo de confianza.** Seleccione un ajuste de diferencia menor significativa (DMS), Bonferroni o Sidak para los intervalos de confianza y la significación. Este elemento sólo estará disponible si se selecciona Comparar los efectos principales.

**Mostrar.** Seleccione Estadísticos descriptivos para generar medias observadas, desviaciones típicas y frecuencias para cada variable dependiente en todas las casillas. La opción Estimaciones del tamaño del efecto ofrece un valor parcial de eta-cuadrado para cada efecto y cada estimación de parámetros. El estadístico eta cuadrado describe la proporción de variabilidad total atribuible a un factor. Seleccione Potencia observada para obtener la potencia de la prueba cuando la hipótesis alternativa se ha establecido basándose en el valor observado. Seleccione Estimaciones de los parámetros para generar las estimaciones de los parámetros, los errores típicos, las pruebas  $t$ , los



intervalos de confianza y la potencia observada para cada prueba. Seleccione Matriz de coeficientes de contraste para obtener la matriz **L**.

Las pruebas de homogeneidad producen la prueba de homogeneidad de varianzas de Levene para cada variable dependiente en todas las combinaciones de nivel de los factores inter-sujetos sólo para factores inter-sujetos. Las opciones de diagramas de dispersión por nivel y gráfico de los residuos son útiles para comprobar los supuestos sobre los datos. Estos elementos no estarán activado si no hay factores. Seleccione Gráficos de los residuos para generar un gráfico de los residuos observados respecto a los pronosticados respecto a los tipificados para cada variable dependiente. Estos gráficos son útiles para investigar el supuesto de varianzas iguales. Seleccione Falta de ajuste para comprobar si el modelo puede describir de forma adecuada la relación entre la variable dependiente y las variables independientes. La función estimable general permite construir pruebas de hipótesis personales basadas en la función estimable general. Las filas en las matrices de coeficientes de contraste son combinaciones lineales de la función estimable general.

**Nivel de significación.** Puede que le interese corregir el nivel de significación usado en las pruebas post hoc y el nivel de confianza empleado para construir intervalos de confianza. El valor especificado también se utiliza para calcular la potencia observada para la prueba. Si especifica un nivel de significación, el cuadro de diálogo mostrará el nivel asociado de los intervalos de confianza.

## ***Funciones adicionales de los comandos UNIANOVA***

Con el lenguaje de sintaxis de comandos también podrá:

- Especificar efectos anidados en el diseño (utilizando el subcomando DESIGN).
- Especificar contrastes de los efectos respecto a una combinación lineal de efectos o un valor (utilizando el subcomando TEST).
- Especificar contrastes múltiples (utilizando el subcomando CONTRAST).
- Incluir los valores perdidos definidos por el usuario (utilizando el subcomando MISSING).
- Especificar criterios EPS (mediante el subcomando CRITERIA).
- Construir una matriz **L**, una matriz **M** o una matriz **K** (utilizando los subcomandos LMATRIX, MMATRIX y KMATRIX).
- Especificar una categoría de referencia intermedia (utilizando el subcomando CONTRAST para los contrastes de desviación o simples).
- Especificar la métrica para los contrastes polinómicos (utilizando el subcomando CONTRAST).
- Especificar términos de error para las comparaciones post hoc (utilizando el subcomando POSTHOC).
- Calcular medias marginales estimadas para cualquier factor o interacción entre los factores en la lista de factores (utilizando el subcomando EMMEANS).
- Especificar nombres para las variables temporales (utilizando el subcomando SAVE).
- Construir un archivo de datos matricial de correlaciones (utilizando el subcomando OUTFILE).

- Construir un archivo de datos matricial que contenga estadísticos de la tabla de ANOVA inter-sujetos (utilizando el subcomando `OUTFILE`).
- Guardar la matriz del diseño en un nuevo archivo de datos (utilizando el subcomando `OUTFILE`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

# ***Correlaciones bivariadas***

El procedimiento Correlaciones bivariadas calcula el coeficiente de correlación de Pearson, la rho de Spearman y la tau-*b* de Kendall con sus niveles de significación. Las correlaciones miden cómo están relacionadas las variables o los órdenes de los rangos. Antes de calcular un coeficiente de correlación, inspeccione los datos para detectar valores atípicos (que pueden generar resultados equívocos) y evidencias de una relación lineal. El coeficiente de correlación de Pearson es una medida de asociación lineal. Dos variables pueden estar perfectamente relacionadas, pero si la relación no es lineal, el coeficiente de correlación de Pearson no será un estadístico adecuado para medir su asociación.

**Ejemplo.** ¿Está el número de partidos ganados por un equipo de baloncesto correlacionado con el número medio de puntos anotados por partido? Un diagrama de dispersión indica que existe una relación lineal. Al analizar los datos de la temporada 94/95 de la NBA, se descubre que el coeficiente de correlación de Pearson (0,581) es significativo al nivel 0,01. Se puede sospechar que cuantos más partidos se ganen por temporada, menos puntos habrán anotado los adversarios. Estas variables están correlacionadas negativamente (-0,401) y la correlación es significativa al nivel 0,05.

**Estadísticos.** Para cada variable: número de casos sin valores perdidos, desviación típica y media. Para cada pareja de variables: coeficiente de correlación de Pearson, rho de Spearman, tau-*b* de Kendall, productos cruzados de las desviaciones y covarianzas.

**Datos.** Utilice variables cuantitativas simétricas para el coeficiente de correlación de Pearson y variables cuantitativas o variables con categorías ordenadas para la rho de Spearman y la tau-*b* de Kendall.

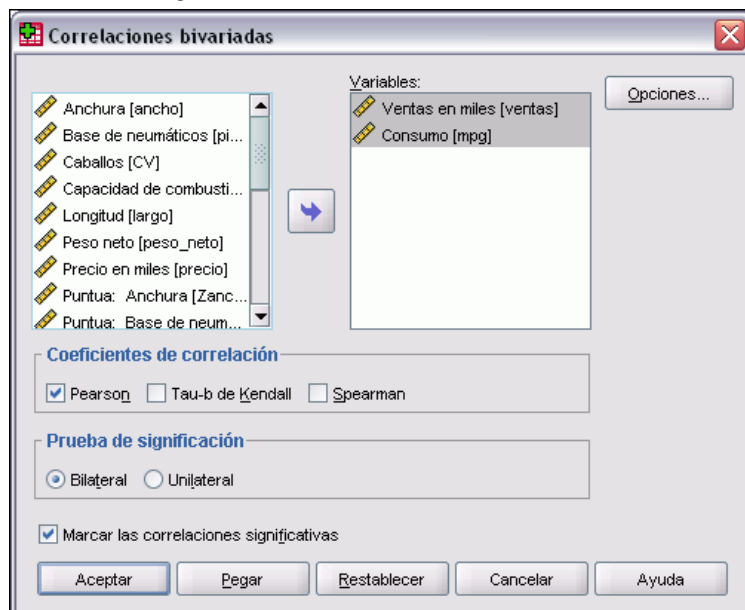
**Supuestos.** El coeficiente de correlación de Pearson asume que cada pareja de variables es normal bivariada.

## ***Para obtener correlaciones bivariadas***

Elija en los menús:

Analizar  
Correlaciones  
Bivariadas...

Figura 23-1  
Cuadro de diálogo *Correlaciones bivariadas*



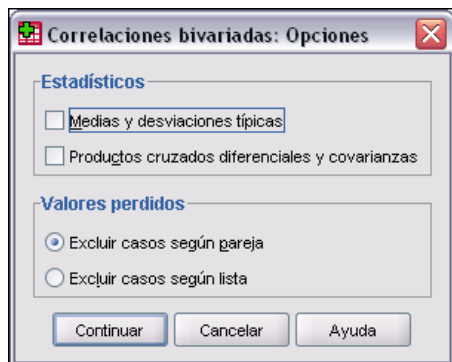
- Seleccione dos o más variables numéricas.

También se encuentran disponibles las siguientes opciones:

- **Coeficientes de correlación.** Para las variables cuantitativas, normalmente distribuidas, seleccione el coeficiente de correlación de Pearson. Si los datos no están normalmente distribuidos o tienen categorías ordenadas, seleccione los correspondientes a la Tau-b de Kendall o Spearman, que miden la asociación entre órdenes de rangos. Los coeficientes de correlación pueden estar entre -1 (una relación negativa perfecta) y +1 (una relación positiva perfecta). Un valor 0 indica que no existe una relación lineal. Al interpretar los resultados, se debe evitar extraer conclusiones de causa-efecto a partir de una correlación significativa.
- **Prueba de significación.** Se pueden seleccionar las probabilidades bilaterales o las unilaterales. Si conoce de antemano la dirección de la asociación, seleccione Unilateral. Si no es así, seleccione Bilateral.
- **Marcar las correlaciones significativas.** Los coeficientes de correlación significativos al nivel 0,05 se identifican por medio de un solo asterisco y los significativos al nivel 0,01 se identifican con dos asteriscos.

## Correlaciones bivariadas: Opciones

Figura 23-2  
Cuadro de diálogo Correlaciones bivariadas: Opciones



**Estadísticos.** Para las correlaciones de Pearson, se puede elegir una o ambas de estas opciones:

- **Medias y desviaciones típicas.** Se muestran para cada variable. También se muestra el número de casos que no tienen valores perdidos. Los valores perdidos se consideran según cada variable individual, sin tener en cuenta la opción elegida para la manipulación de los valores perdidos.
- **Productos cruzados diferenciales y covarianzas.** Se muestran para cada pareja de variables. Cada producto cruzado de las desviaciones es igual a la suma de los productos de las variables corregidas respecto a la media. Éste es el numerador del coeficiente de correlación de Pearson. La covarianza es una medida no tipificada de la relación entre dos variables, igual al producto cruzado diferencial dividido por  $N-1$ .

**Valores perdidos.** Puede elegir uno de los siguientes:

- **Excluir casos según pareja.** Se excluyen del análisis los casos con valores perdidos para una o ambas variables de la pareja que forma un coeficiente de correlación. Debido a que cada coeficiente está basado en todos los casos que tienen códigos válidos para esa pareja concreta de variables, en cada cálculo se utiliza la mayor cantidad de información disponible. Esto puede dar como resultado un grupo de coeficientes basados en un número de casos variable.
- **Excluir casos según lista.** Se excluyen de todas las correlaciones los casos con valores perdidos para cualquier variable.

## Funciones adicionales de los comandos **CORRELATIONS** y **NONPAR CORR**

Con el lenguaje de sintaxis de comandos también podrá:

- Escribir una matriz de correlaciones para correlaciones de Pearson que pueda ser utilizada en lugar de los datos brutos, con el fin de obtener otros análisis como el análisis factorial (con el subcomando `MATRIX`).
- Obtener correlaciones de todas las variables de una lista con todas las variables de una segunda lista (utilizando la palabra clave `WITH` en el subcomando `VARIABLES`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

# *Correlaciones parciales*

El procedimiento Correlaciones parciales calcula los coeficientes de correlación parcial, los cuales describen la relación lineal existente entre dos variables mientras se controlan los efectos de una o más variables adicionales. Las correlaciones son medidas de asociación lineal. Dos variables pueden estar perfectamente relacionadas, pero si la relación no es lineal, el coeficiente de correlación no es un estadístico adecuado para medir su asociación.

**Ejemplo.** ¿Existe alguna relación entre la financiación sanitaria y las tasas de enfermedad? Aunque cabe esperar que dicha relación sea negativa, un estudio describe una correlación *positiva* significativa: si la financiación sanitaria aumenta, las tasas de enfermedad parecen disminuir. Sin embargo, si se controla la tasa de visitas de visitantes médicos, se elimina prácticamente la correlación positiva observada. La financiación sanitaria y las tasas de enfermedad sólo parecen estar relacionadas positivamente debido a que más personas tienen acceso a la sanidad si la financiación aumenta, lo que tiene como resultado que los médicos y hospitales informen de más enfermedades.

**Estadísticos.** Para cada variable: número de casos sin valores perdidos, desviación típica y media. Matrices de correlación de orden cero y parcial, con grados de libertad y niveles de significación.

**Datos.** Utilice variables cuantitativas y simétricas.

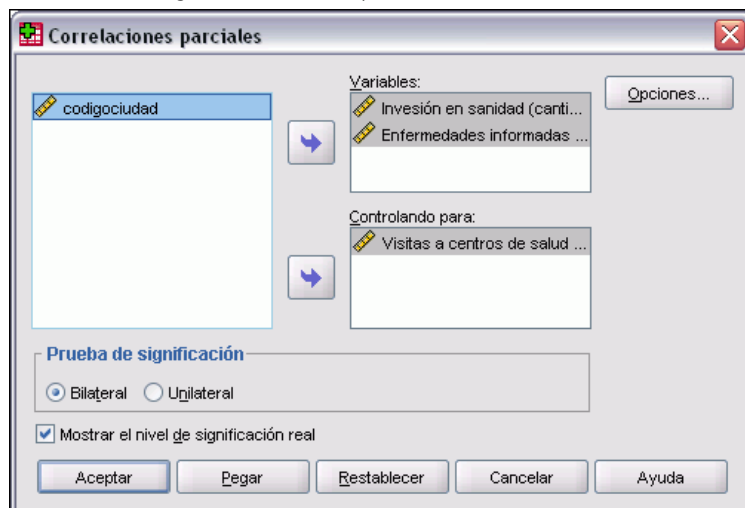
**Supuestos.** El procedimiento Correlaciones parciales supone que cada par de variables es normal bivalente.

## ***Para obtener correlaciones parciales***

► Elija en los menús:

Analizar  
Correlaciones  
Parciales...

Figura 24-1  
Cuadro de diálogo *Correlaciones parciales*



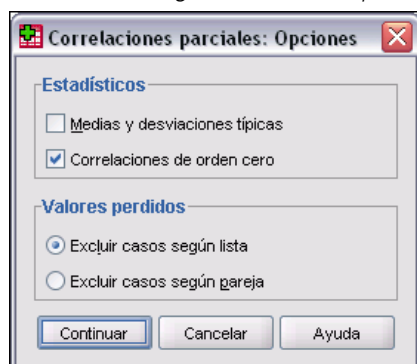
- ▶ Seleccione dos o más variables numéricas para las que se van a calcular las correlaciones parciales.
- ▶ Elija una o más variables numéricas de control.

También se encuentran disponibles las siguientes opciones:

- **Prueba de significación.** Se pueden seleccionar las probabilidades bilaterales o las unilaterales. Si conoce de antemano la dirección de la asociación, seleccione Unilateral. Si no es así, seleccione Bilateral.
- **Mostrar el nivel de significación real.** Por defecto, se muestran la probabilidad y los grados de libertad para cada coeficiente de correlación. Si anula la selección de este elemento, los coeficientes significativos al nivel 0,05 se identifican con un asterisco, los coeficientes significativos al nivel 0,01 se identifican con un asterisco doble y se eliminan los grados de libertad. Este ajuste afecta a las matrices de correlación parcial y de orden cero.

## ***Correlaciones parciales: Opciones***

Figura 24-2  
Cuadro de diálogo *Correlaciones parciales: Opciones*





**Estadísticos.** Puede elegir una o ambas de las siguientes opciones:

- **Medias y desviaciones típicas.** Se muestran para cada variable. También se muestra el número de casos que no tienen valores perdidos.
- **Correlaciones de orden cero.** Se muestra una matriz de las correlaciones simples entre todas las variables, incluyendo las variables de control.

**Valores perdidos.** Puede elegir una de las siguientes alternativas:

- **Excluir casos según lista.** Se excluyen de todos los cálculos los casos que presenten valores perdidos para cualquier variable, incluso si es para las variables de control.
- **Excluir casos según pareja.** Para el cálculo de las correlaciones de orden cero, en las que se basan las correlaciones parciales, no se utilizará un caso si tiene valores perdidos en una o ambas variables de un par. La eliminación según pareja aprovecha el máximo de los datos que sean posibles. Sin embargo, el número de casos puede variar de unos coeficientes a otros. Cuando se activa esta opción, los grados de libertad para un coeficiente parcial determinado se basan en el número menor de casos utilizado en el cálculo de cualquiera de las correlaciones de orden cero necesarias para el cálculo de dicho coeficiente parcial.

## ***Funciones adicionales del comando PARTIAL CORR***

Con el lenguaje de sintaxis de comandos también podrá:

- Leer una matriz de correlaciones de orden cero o escribir una matriz de correlaciones parciales (mediante el subcomando `MATRIX`).
- Obtener correlaciones parciales entre dos listas de variables (mediante la palabra clave `WITH` en el subcomando `VARIABLES`).
- Obtener análisis múltiples (mediante varios subcomandos `VARIABLES`).
- Especificar otros valores para solicitar (por ejemplo, las correlaciones parciales tanto de primer como de segundo orden) cuando tiene dos variables de control (mediante el subcomando `VARIABLES`).
- Suprimir coeficientes redundantes (mediante el subcomando `FORMAT`).
- Mostrar una matriz de correlaciones simples cuando algunos coeficientes no se pueden calcular (mediante el subcomando `STATISTICS`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

# ***Distancias***

Este procedimiento calcula una variedad de estadísticos que miden las similitudes o diferencias (distancias), entre pares de variables o entre pares de casos. Estas medidas de similitud o distancia se pueden utilizar después con otros procedimientos, como análisis factorial, análisis de conglomerados o escalamiento multidimensional, para ayudar en el análisis de conjuntos de datos complejos.

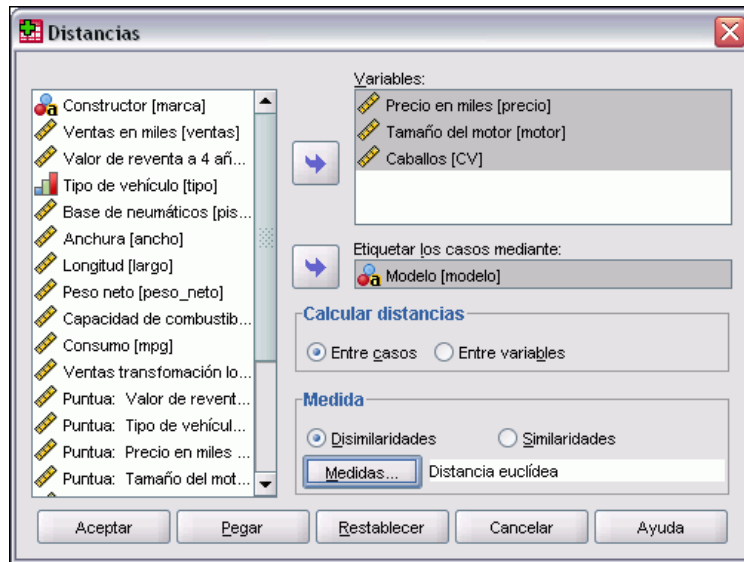
**Ejemplo.** ¿Es posible medir similaridades entre pares de automóviles en función de ciertas características, como tipo de motor, consumo y potencia? Al calcular las similitudes entre los coches, se puede obtener una noción de qué coches son similares entre sí y cuáles son diferentes. Para un análisis más formal, puede considerar la aplicación de un análisis de conglomeración jerárquico o escalamiento multidimensional a las similitudes para explorar la estructura subyacente.

**Estadísticos.** Las medidas de diferencia (distancia) para datos de un intervalo son Distancia euclídea, Distancia euclídea al cuadrado, Chebychev, bloque, Minkowski o personalizada; para datos de recuento, medida de chi-cuadrado o phi-cuadrado; para datos binarios, Distancia euclídea, Distancia euclídea al cuadrado, diferencia de tamaño, diferencia de configuración, varianza, forma o Lance y Williams. Las medidas de similitud para datos de intervalos son correlación de Pearson o coseno; para datos binarios, Russel y Rao, concordancia simple, Jaccard, Dice, Rogers y Tanimoto, Sokal y Sneath 1, Sokal y Sneath 2, Sokal y Sneath 3, Kulczynski 1, Kulczynski 2, Sokal y Sneath 4, Hamann, Lambda,  $D$  de Anderberg,  $Y$  de Yule,  $Q$  de Yule, Ochiai, Sokal y Sneath 5, correlación Phi de 4 puntos o dispersión.

## ***Para obtener matrices de distancias***

- ▶ Elija en los menús:
  - Analizar
  - Correlaciones
  - Distancias...

Figura 25-1  
Cuadro de diálogo Distancias



- ▶ Seleccione al menos una variable numérica para calcular distancias entre casos o seleccione al menos dos variables numéricas para calcular distancias entre variables.
- ▶ Seleccione una alternativa en el grupo Calcular distancias para calcular proximidades entre casos o entre variables.

## Distancias: Medidas de disimilaridad

Figura 25-2  
Cuadro de diálogo Distancias: Medidas de disimilaridad

En el grupo Medida, seleccione la alternativa que corresponda al tipo de datos (intervalo, recuento o binario); a continuación, de la lista desplegable, seleccione una las medidas que corresponda a dicho tipo de datos. Las medidas disponibles, por tipo de dato, son:

- **Datos de intervalo.** Distancia euclídea, Distancia euclídea al cuadrado, Chebychev, Bloque, Minkowski o Personalizada.
- **Datos de frecuencias.** Medida de chi-cuadrado o Medida de phi-cuadrado.
- **Datos binarios.** Distancia euclídea, Distancia euclídea al cuadrado, Diferencia de tamaño, Diferencia de configuración, Varianza, Forma o Lance y Williams. (Introduzca valores para Presente y Ausente para especificar cuáles son los dos valores representativos; las Distancias ignorarán todos los demás valores.)

El grupo Transformar valores permite estandarizar los valores de los datos para casos o variables *antes* de calcular proximidades. Estas transformaciones no se pueden aplicar a los datos binarios. Los métodos disponibles de estandarización son: Puntuaciones  $z$ , Rango -1 a -1, Rango 0 a 1, Magnitud máxima de 1, Media de 1 y Desviación típica 1.

El grupo Transformar medidas permite transformar los valores generados por la medida de distancia. Se aplican después de calcular la medida de distancia. Las opciones disponibles son: Valores absolutos, Cambiar el signo y Cambiar la escala al rango 0-1.

## Distancias: Medidas de similitud

Figura 25-3  
Cuadro de diálogo Distancias: Medidas de similitud

En el grupo Medida, seleccione la alternativa que corresponda al tipo de datos (intervalo o binario); a continuación, de la lista desplegable, seleccione una de las medidas que corresponda a dicho tipo de datos. Las medidas disponibles, por tipo de dato, son:

- **Datos de intervalo.** Correlación de Pearson o Coseno.
- **Datos binarios.** Russel y Rao, Concordancia simple, Jaccard, Dice, Rogers y Tanimoto, Sokal y Sneath 1, Sokal y Sneath 2, Sokal y Sneath 3, Kulczynski 1, Kulczynski 2, Sokal y Sneath 4, Hamann, Lambda,  $D$  de Anderberg,  $Y$  de Yule,  $Q$  de Yule, Ochiai, Sokal y Sneath 5, Correlación Phi de 4 puntos o Dispersión. (Introduzca valores para Presente y Ausente para especificar cuáles son los dos valores representativos; las Distancias ignorarán todos los demás valores.)

El grupo Transformar valores permite estandarizar los valores de los datos para casos o variables antes de calcular proximidades. Estas transformaciones no se pueden aplicar a los datos binarios. Los métodos disponibles de estandarización son: Puntuaciones  $z$ , Rango -1 a 1, Rango 0 a 1, Magnitud máxima de 1, Media de 1 y Desviación típica 1.

El grupo Transformar medidas permite transformar los valores generados por la medida de distancia. Se aplican después de calcular la medida de distancia. Las opciones disponibles son: Valores absolutos, Cambiar el signo y Cambiar la escala al rango 0-1.

## ***Funciones adicionales del comando PROXIMITIES***

El procedimiento Distancias utiliza la sintaxis de comandos PROXIMITIES. Con el lenguaje de sintaxis de comandos también podrá:

- Especificar cualquier número entero como la potencia para la medida de distancia de Minkowski.
- Especificar cualquier número entero como la potencia y la raíz para una medida de distancia personalizada.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

# Regresión lineal

La regresión lineal estima los coeficientes de la ecuación lineal, con una o más variables independientes, que mejor prediga el valor de la variable dependiente. Por ejemplo, puede intentar predecir el total de ventas anuales de un vendedor (la variable dependiente) a partir de variables independientes tales como la edad, la formación y los años de experiencia.

**Ejemplo.** ¿Están relacionados el número de partidos ganados por un equipo de baloncesto en una temporada con la media de puntos que el equipo marca por partido? Un diagrama de dispersión indica que estas variables están relacionadas linealmente. El número de partidos ganados y la media de puntos marcados por el equipo adversario también están relacionados linealmente. Estas variables tienen una relación negativa. A medida que el número de partidos ganados aumenta, la media de puntos marcados por el equipo adversario disminuye. Con la regresión lineal es posible modelar la relación entre estas variables. Puede utilizarse un buen modelo para predecir cuántos partidos ganarán los equipos.

**Estadísticos.** Para cada variable: número de casos válidos, media y desviación típica. Para cada modelo: coeficientes de regresión, matriz de correlaciones, correlaciones parciales y semiparciales,  $R$  múltiple,  $R^2$ ,  $R^2$  corregida, cambio en  $R^2$ , error típico de la estimación, tabla de análisis de varianza, valores pronosticados y residuos. Además, intervalos de confianza al 95% para cada coeficiente de regresión, matriz de varianzas-covarianzas, factor de inflación de la varianza, tolerancia, prueba de Durbin-Watson, medidas de distancia (Mahalanobis, Cook y valores de influencia),  $Df_{\text{Beta}}$ ,  $Df_{\text{Ajuste}}$ , intervalos de pronóstico y diagnósticos por caso. Gráficos: diagramas de dispersión, gráficos parciales, histogramas y gráficos de probabilidad normal.

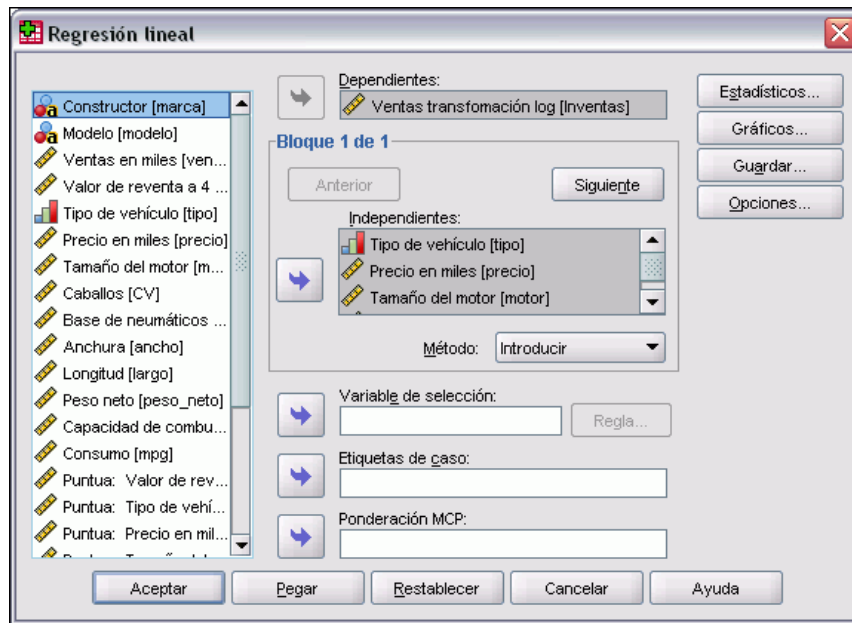
**Datos.** Las variables dependiente e independientes deben ser cuantitativas. Las variables categóricas, como la religión, estudios principales o el lugar de residencia, han de recodificarse como variables binarias (dummy) o como otros tipos de variables de contraste.

**Supuestos.** Para cada valor de la variable independiente, la distribución de la variable dependiente debe ser normal. La varianza de distribución de la variable dependiente debe ser constante para todos los valores de la variable independiente. La relación entre la variable dependiente y cada variable independiente debe ser lineal y todas las observaciones deben ser independientes.

## **Para obtener un análisis de regresión lineal**

- Elija en los menús:
  - Analizar
  - Regresión
  - Lineal...

Figura 26-1  
Cuadro de diálogo Regresión lineal



- ▶ En el cuadro de diálogo Regresión lineal, seleccione una variable numérica dependiente.
- ▶ Seleccione una más variables numéricas independientes.

Si lo desea, puede:

- Agrupar variables independientes en bloques y especificar distintos métodos de entrada para diferentes subconjuntos de variables.
- Elegir una variable de selección para limitar el análisis a un subconjunto de casos que tengan valores particulares para esta variable.
- Seleccionar una variable de identificación de casos para identificar los puntos en los diagramas.
- Seleccione una variable numérica de Ponderación MCP para el análisis de mínimos cuadrados ponderados.

**MCP (WLS).** Permite obtener un modelo de mínimos cuadrados ponderados. Los puntos de datos se ponderan por los inversos de sus varianzas. Esto significa que las observaciones con varianzas grandes tienen menor impacto en el análisis que las observaciones asociadas a varianzas pequeñas. Si el valor de la variable de ponderación es cero, negativo o perdido, el caso queda excluido del análisis.

## ***Métodos de selección de variables en el análisis de regresión lineal***

La selección del método permite especificar cómo se introducen las variables independientes en el análisis. Utilizando distintos métodos se pueden construir diversos modelos de regresión a partir del mismo conjunto de variables.



- **Introducir (Regresión).** Procedimiento para la selección de variables en el que todas las variables de un bloque se introducen en un solo paso.
- **Pasos suc..** En cada paso se introduce la variable independiente que no se encuentre ya en la ecuación y que tenga la probabilidad para F más pequeña, si esa probabilidad es suficientemente pequeña. Las variables ya introducidas en la ecuación de regresión se eliminan de ella si su probabilidad para F llega a ser suficientemente grande. El método termina cuando ya no hay más variables candidatas a ser incluidas o eliminadas.
- **Borrar.** Procedimiento para la selección de variables en el que las variables de un bloque se eliminan en un solo paso.
- **Eliminación hacia atrás.** Procedimiento de selección de variables en el que se introducen todas las variables en la ecuación y después se van excluyendo una tras otra. Aquella variable que tenga la menor correlación parcial con la variable dependiente será la primera en ser considerada para su exclusión. Si satisface el criterio de eliminación, será eliminada. Tras haber excluido la primera variable, se pondrá a prueba aquella variable, de las que queden en la ecuación, que presente una correlación parcial más pequeña. El procedimiento termina cuando ya no quedan en la ecuación variables que satisfagan el criterio de exclusión.
- **Selección hacia adelante.** Procedimiento de selección de variables en el que éstas son introducidas secuencialmente en el modelo. La primera variable que se considerará para ser introducida en la ecuación será aquella que tenga mayor correlación, positiva o negativa, con la variable dependiente. Dicha variable será introducida en la ecuación sólo si satisface el criterio de entrada. Si ha entrado la primera variable, se considerará como próxima candidata la variable independiente que no esté en la ecuación y cuya correlación parcial sea la mayor. El procedimiento termina cuando ya no quedan variables que satisfagan el criterio de entrada.

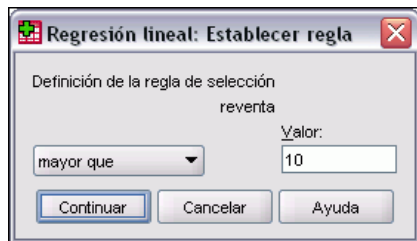
Los valores de significación de los resultados se basan en el ajuste de un único modelo. Por ello, estos valores no suele ser válidos cuando se emplea un método por pasos (pasos sucesivos, hacia adelante o hacia atrás).

Todas las variables deben superar el criterio de tolerancia para que puedan ser introducidas en la ecuación, independientemente del método de entrada especificado. El nivel de tolerancia por defecto es 0,0001. Tampoco se introduce una variable si esto provoca que la tolerancia de otra ya presente en el modelo se sitúe por debajo del criterio de tolerancia.

Todas las variables independientes seleccionadas se añaden a un mismo modelo de regresión. Sin embargo, puede especificar distintos métodos de introducción para diferentes subconjuntos de variables. Por ejemplo, puede introducir en el modelo de regresión un bloque de variables que utilice la selección por pasos sucesivos, y un segundo bloque que emplee la selección hacia adelante. Para añadir un segundo bloque de variables al modelo de regresión, pulse en Siguiente.

## Regresión lineal: Establecer regla

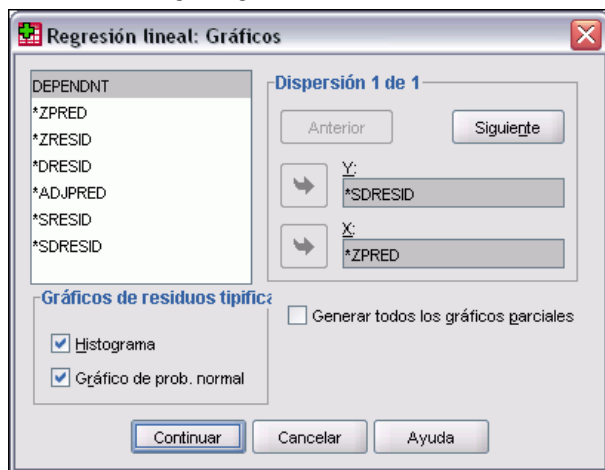
Figura 26-2  
Cuadro de diálogo Regresión lineal: Establecer regla



Los casos definidos por la regla de selección se incluyen en el análisis. Por ejemplo, si selecciona una variable, elija igual que y escriba 5 para el valor; de este modo, solamente se incluirán en el análisis los casos para los cuales la variable seleccionada tenga un valor igual a 5. También se permite un valor de cadena.

## Regresión lineal: Gráficos

Figura 26-3  
Cuadro de diálogo Regresión lineal: Gráficos



Los gráficos pueden ayudar a validar los supuestos de normalidad, linealidad e igualdad de las varianzas. También son útiles para detectar valores atípicos, observaciones poco usuales y casos de influencia. Tras guardarlos como nuevas variables, dispondrá en el Editor de datos de los valores pronosticados, los residuos y otros valores diagnósticos, con los cuales podrá poder crear gráficos respecto a las variables independientes. Se encuentran disponibles los siguientes gráficos:

**Diagramas de dispersión.** Puede representar cualquier combinación por parejas de la lista siguiente: la variable dependiente, los valores pronosticados tipificados, los residuos tipificados, los residuos eliminados, los valores pronosticados corregidos, los residuos estudentizados o los residuos eliminados estudentizados. Represente los residuos tipificados frente a los valores pronosticados tipificados para contrastar la linealidad y la igualdad de las varianzas.

**Lista de variables de origen.** Muestra una lista con la variable dependiente (DEPENDNT) y las siguientes variables pronosticadas y residuales: valores pronosticados tipificados (\*ZPRED), residuos tipificados (\*ZRESID), residuos eliminados (\*DRESID), valores pronosticados corregidos (\*ADJPRED), residuos estudentizados (\*SRESID) y residuos estudentizados eliminados (\*SDRESID).

**Generar todos los gráficos parciales.** Muestra los diagramas de dispersión de los residuos de cada variable independiente y los residuos de la variable dependiente cuando se regresan ambas variables por separado sobre las restantes variables independientes. En la ecuación debe haber al menos dos variables independientes para que se generen los gráficos parciales.

**Gráficos de residuos tipificados.** Puede obtener histogramas de los residuos tipificados y gráficos de probabilidad normal que comparen la distribución de los residuos tipificados con una distribución normal.

Si se solicita cualquier gráfico, se muestran los estadísticos de resumen para los valores pronosticados tipificados y los residuos tipificados (\*ZPRED y \*ZRESID).

## Regresión lineal: Para guardar variables nuevas

Figura 26-4  
Cuadro de diálogo Regresión lineal: Guardar

**Regresión lineal: Guardar**

**Valores pronosticados**

- No tipificados
- Tipificados
- Corregidos
- E.T. del pronóstico promedio

**Residuos**

- No tipificados
- Tipificados
- Método de Student
- Eliminados
- Eliminados estudentizados

**Distancias**

- Mahalanobis
- De Cook
- Valores de influencia

**Estadísticos de influencia**

- DfBetas
- DfBetas tipificadas
- DfAjuste
- DfAjuste tipificada
- Razón entre covarianzas

**Intervalos de pronóstico**

- Media
- Individuos

Intervalo de confianza:  %

**Estadísticos de los coeficientes**

- Crear estadísticos de los coeficientes
- Crear un nuevo conjunto de datos
  - Nombre de conjunto de datos:
- Escribir un nuevo archivo de datos
  - Archivo...

**Exportar información del modelo a un archivo XML**

Incluir la matriz de covarianzas

Puede guardar los valores pronosticados, los residuos y otros estadísticos útiles para los diagnósticos. Cada selección añade una o más variables nuevas a su archivo de datos activo.

**Valores pronosticados.** Son los valores que el modelo de regresión pronostica para cada caso.

- **No tipificados.** Valor pronosticado por el modelo para la variable dependiente.
- **Tipificados.** Transformación de cada valor pronosticado a su forma tipificada. Es decir, se sustrae el valor pronosticado medio al valor pronosticado y el resultado se divide por la desviación típica de los valores pronosticados. Los valores pronosticados tipificados tienen una media de 0 y una desviación típica de 1.
- **Corregidos.** Valor pronosticado para un caso cuando dicho caso no se incluye en los cálculos de los coeficientes de regresión.
- **E.T. del pronóstico promedio.** Error típico de los valores pronosticados. Estimación de la desviación típica del valor promedio de la variable dependiente para los casos que tengan los mismos valores en las variables independientes.

**Distancias.** Son medidas para identificar casos con combinaciones poco usuales de valores para las variables independientes y casos que puedan tener un gran impacto en el modelo.

- **Mahalanobis.** Medida de cuánto difieren los valores de un caso en las variables independientes respecto al promedio para todos los casos. Una distancia de Mahalanobis grande identificará a un caso que tenga valores extremos en una o más de las variables independientes.
- **De Cook.** Medida de cuánto cambiarían los residuos de todos los casos si se excluyera un caso determinado del cálculo de los coeficientes de regresión. Una Distancia de Cook grande indica que la exclusión de ese caso del cálculo de los estadísticos de regresión hará variar substancialmente los coeficientes.
- **Valores de influencia.** Medida de la influencia de un punto en el ajuste de la regresión. La influencia centrada varía entre 0 (no influye en el ajuste) y  $(N-1)/N$ .

**Intervalos de pronóstico.** Los límites superior e inferior para los intervalos de pronóstico individual y promedio.

- **Media.** Límites inferior y superior (dos variables) para el intervalo de predicción de la respuesta pronosticada promedio.
- **Individual.** Límites superior e inferior (dos variables) del intervalo de predicción para la variable dependiente para un caso individual.
- **Intervalo de confianza.** Introduzca un valor entre 1 y 99,99 para especificar el nivel de confianza para los dos intervalos de predicción. Debe seleccionar Media o Individuos antes de introducir este valor. Los valores habituales para los intervalos de confianza son 90, 95 y 99.

**Residuos.** El valor actual de la variable dependiente menos el valor pronosticado por la ecuación de regresión.

- **No tipificados.** Diferencia entre un valor observado y el valor pronosticado por el modelo.
- **Tipificados.** El residuo dividido por una estimación de su error típico. Los residuos tipificados, que son conocidos también como los residuos de Pearson o residuos estandarizados, tienen una media de 0 y una desviación típica de 1.
- **Método de Student.** Residuo dividido por una estimación de su desviación típica que varía de caso en caso, dependiendo de la distancia de los valores de cada caso en las variables independientes respecto a las medias en las variables independientes.

- **Eliminados.** Residuo para un caso cuando éste se excluye del cálculo de los coeficientes de la regresión. Es igual a la diferencia entre el valor de la variable dependiente y el valor pronosticado corregido.
- **Eliminados estudentizados.** Residuo eliminado para un caso dividido por su error típico. La diferencia entre un residuo eliminado estudentizado y su residuo estudentizado asociado indica la diferencia que implica el eliminar un caso sobre su propia predicción.

**Estadísticos de influencia.** El cambio en los coeficientes de regresión (DfBeta) y en los valores pronosticados (DfAjuste) que resulta de la exclusión de un caso particular. También están disponibles los valores tipificados para las DfBeta y para las DfAjuste, junto con la razón entre covarianzas.

- **DfBeta(s).** La diferencia en el valor de beta es el cambio en el valor de un coeficiente de regresión que resulta de la exclusión de un caso particular. Se calcula un valor para cada término del modelo, incluyendo la constante.
- **DfBeta tipificada.** Valor de la diferencia en beta tipificada. El cambio tipificado en un coeficiente de regresión cuando se elimina del análisis un caso particular. Puede interesarle examinar aquellos casos cuyos valores absolutos sean mayores que 2 dividido por la raíz cuadrada de N, donde N es el número de casos. Se calcula un valor para cada término del modelo, incluyendo la constante.
- **DfAjuste.** La diferencia en el valor ajustado es el cambio en el valor pronosticado que resulta de la exclusión de un caso particular.
- **DfAjuste tipificado.** Diferencia tipificada en el valor ajustado. El cambio, tipificado, en el valor pronosticado que resulta de la exclusión de un caso particular. Puede interesarle examinar aquellos valores tipificados cuyo valor absoluto sea mayor que 2 dividido por la raíz cuadrada de  $p/N$ , donde p es el número de variables independientes en la ecuación y N es el número de casos.
- **Razón entre covarianzas.** Razón del determinante de la matriz de covarianza con un caso particular excluido del cálculo de los coeficientes de regresión, respecto al determinante de la matriz de covarianza con todos los casos incluidos. Si la razón se aproxima a 1, el caso no altera significativamente la matriz de covarianza.

**Estadísticos de los coeficientes.** Almacena los coeficientes de regresión en un conjunto de datos o en un archivo de datos. Los conjuntos de datos están disponibles para su uso posterior durante la misma sesión, pero no se guardarán como archivos a menos que se hayan guardado explícitamente antes de que finalice la sesión. El nombre de un conjunto de datos debe cumplir las normas de denominación de variables. Si desea obtener más información, consulte [Nombres de variable](#) en Capítulo 5 en p. 82.

**Exportar información del modelo a un archivo XML.** Las estimaciones de los parámetros y (si lo desea) sus covarianzas se exportan al archivo especificado en formato XML (PMML). SmartScore y servidor de SPSS (un producto independiente) pueden utilizar este archivo del modelo para aplicar la información del modelo en otros archivos de datos con fines de puntuación.

## Regresión lineal: Estadísticos

Figura 26-5  
Cuadro de diálogo Estadísticos



Se encuentran disponibles los siguientes estadísticos:

**Coeficientes de regresión.** La opción Estimaciones muestra el coeficiente de regresión  $B$ , el error típico de  $B$ , el coeficiente beta tipificado, el valor de  $t$  para  $B$  y el nivel de significación bilateral  $z$ . Intervalos de confianza muestra intervalos de confianza al 95% para cada coeficiente de regresión o una matriz de covarianzas. Matriz de covarianzas muestra una matriz de varianzas-covarianzas de los coeficientes de regresión, con las covarianzas fuera de la diagonal y las varianzas en la diagonal. También se muestra una matriz de correlaciones.

**Ajuste del modelo.** Presenta una lista de las variables introducidas y eliminadas del modelo y muestra los siguientes estadísticos de bondad de ajuste:  $R$  múltiple,  $R^2$  cuadrado y  $R^2$  cuadrado corregido, error típico de la estimación y tabla de análisis de la varianza.

**Cambio en R cuadrado.** Cambio en el estadístico  $R^2$  cuadrado que se produce al añadir o eliminar una variable independiente. Si es grande el cambio en  $R^2$  cuadrado asociado a una variable, esto significa que esa variable es un buen predictor de la variable dependiente.

**Descriptivos.** Proporciona el número de casos válidos, la media y la desviación típica para cada variable en el análisis. También muestra una matriz de correlaciones con el nivel de significación unilateral y el número de casos para cada correlación.

**Correlación parcial.** La correlación que permanece entre dos variables después de eliminar la correlación que es debida a su relación mutua con las otras variables. La correlación entre la variable dependiente y una variable independiente cuando se han eliminado de ambas los efectos lineales de las otras variables independientes presentes en el modelo.

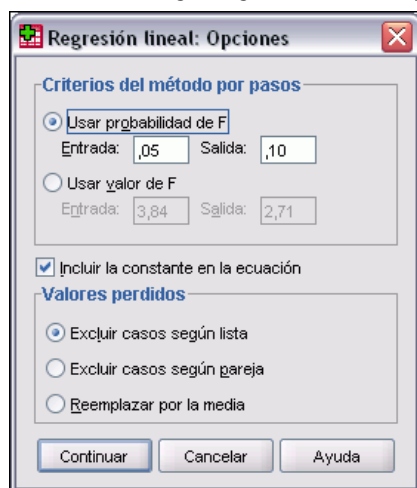
**Correlación semi-parcial.** La correlación entre una variable dependiente y una variable independiente cuando se han eliminado de la variable independiente los efectos lineales de las otras variables independientes del modelo. Está relacionada con el cambio en  $R^2$  cuadrado cuando una variable se añade a una ecuación. En ocasiones se denomina correlación semiparcial.

**Diagnósticos de colinealidad.** La colinealidad (o multicolinealidad) es una situación no deseable en la que una de las variables independientes es una función lineal de otras variables independientes. Muestra los autovalores de la matriz de productos cruzados no centrada y escalada, los índices de condición y las proporciones de la descomposición de la varianza junto con los factores de inflación de la varianza (FIV) y las tolerancias para las variables individuales.

**Residuos.** Presenta la prueba de Durbin-Watson sobre la correlación serial de los residuos y los diagnósticos por casos para los casos que cumplan el criterio de selección (los valores atípicos por encima de  $n$  desviaciones típicas).

## Regresión lineal: Opciones

Figura 26-6  
Cuadro de diálogo Regresión lineal: Opciones



Se encuentran disponibles las siguientes opciones:

**Criterios del método por pasos.** Estas opciones son aplicables si se ha especificado el método de selección de variables hacia adelante, hacia atrás o por pasos. Las variables se pueden introducir o eliminar del modelo dependiendo de la significación (probabilidad) del valor de  $F$  o del propio valor de  $F$ .

- **Usar probabilidad de F.** Una variable se introduce en el modelo si el nivel de significación de su valor de  $F$  es menor que el valor de entrada, y se elimina si el nivel de significación de su valor de  $F$  es mayor que el valor de salida. La entrada debe ser menor que la salida y ambos valores deben ser positivos. Para introducir más variables en el modelo, eleve el valor de entrada. Para eliminar más variables del modelo, disminuya el valor de salida.
- **Usar valor de F.** Una variable se introduce en el modelo si su valor de  $F$  es mayor que el valor de entrada, y se elimina si su valor de  $F$  es menor que el valor de salida. La entrada debe ser mayor que la salida y ambos valores deben ser positivos. Para introducir más variables en el modelo, disminuya el valor de entrada. Para eliminar más variables del modelo, eleve el valor de salida.

**Incluir la constante en la ecuación.** Por defecto, el modelo de regresión incluye un término constante. Si se anula la selección de esta opción se obtiene la regresión que pasan por el origen, lo cual se hace raramente. Algunos resultados de la regresión que pasan por el origen no son comparables con los resultados de la regresión que sí incluyen una constante. Por ejemplo,  $R^{\text{cuadrado}}$  no puede interpretarse de la manera usual.

**Valores perdidos.** Puede elegir uno de los siguientes:

- **Excluir casos según lista.** Sólo se incluirán en el análisis los casos con valores válidos para todas las variables.
- **Excluir casos según pareja.** Los casos con datos completos para la pareja de variables correlacionadas se utilizan para calcular el coeficiente de correlación en el cual se basa el análisis de regresión. Los grados de libertad se basan en el  $N$  mínimo de las parejas.
- **Reemplazar por la media.** Se emplean todos los casos en los cálculos, sustituyendo las observaciones perdidas por la media de la variable.

## ***Funciones adicionales del comando REGRESSION***

Con el lenguaje de sintaxis de comandos también podrá:

- Escribir una matriz de correlaciones o leer una matriz (en lugar de los datos brutos) con el fin de obtener el análisis de regresión (mediante el subcomando `MATRIX`).
- Especificar los niveles de tolerancia (mediante el subcomando `CRITERIA`).
- Obtener múltiples modelos para las mismas variables dependientes u otras diferentes (mediante los subcomandos `METHOD` y `DEPENDENT`).
- Obtener estadísticos múltiples (mediante los subcomandos `DESCRIPTIVE` y `VARIABLES`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).



## ***Regresión ordinal***

La regresión ordinal permite dar forma a la dependencia de una respuesta ordinal politómica sobre un conjunto de predictores, que pueden ser factores o covariables. El diseño de la regresión ordinal se basa en la metodología de McCullagh (1980, 1998) y en la sintaxis se hace referencia al procedimiento como `PLUM`.

El análisis de regresión lineal ordinario implica minimizar las diferencias de la suma de los cuadrados entre una variable de respuesta (la dependiente) y una combinación ponderada de las variables predictoras (las independientes). Los coeficientes estimados reflejan cómo los cambios en los predictores afectan a la respuesta. Se considera que la respuesta es numérica, en el sentido en que los cambios en el nivel de la respuesta son equivalentes en todo el rango de la respuesta. Por ejemplo, la diferencia de altura entre una persona que mide 150 cm y una que mide 140 cm es de 10 cm, que tiene el mismo significado que la diferencia de altura entre una persona que mide 210 cm y una que mide 200 cm. Estas relaciones no se mantienen necesariamente con las variables ordinales, en las que la elección y el número de categorías de respuesta pueden ser bastante arbitrarios.

**Ejemplo.** La regresión ordinal podría utilizarse para estudiar la reacción de los pacientes con respecto a una dosis de un fármaco. Las reacciones posibles podrían clasificarse como *ninguna*, *ligera*, *moderada* o *grave*. La diferencia entre una reacción ligera y una moderada es difícil o imposible de cuantificar y se basa en la apreciación. Además, la diferencia entre una respuesta ligera y una moderada podría ser superior o inferior a la diferencia entre una respuesta moderada y una grave.

**Estadísticos y gráficos.** Frecuencias observadas y esperadas y frecuencias acumuladas, residuos de Pearson para las frecuencias y las frecuencias acumuladas, probabilidades observadas y esperadas, probabilidades acumuladas observadas y esperadas para cada categoría de respuesta por patrón en las covariables, matrices de correlaciones asintóticas y de covarianzas entre las estimaciones de los parámetros, chi-cuadrado de Pearson y chi-cuadrado de la razón de verosimilitud, estadísticos de bondad de ajuste, historial de iteraciones, contraste del supuesto de líneas paralelas, estimaciones de los parámetros, errores típicos, intervalos de confianza y estadísticos  $R^2$  de Cox y Snell, de Nagelkerke y de McFadden.

**Datos.** Se asume que la variable dependiente es ordinal y puede ser numérica o de cadena. El orden se determina al clasificar los valores de la variable dependiente en orden ascendente. El valor inferior define la primera categoría. Se asume que las variables de factor son categóricas. Las covariables deben ser numéricas. Observe que al usar más de una covariable continua, se puede llegar a crear una tabla de probabilidades de casilla muy grande.

**Supuestos.** Sólo se permite una variable de respuesta y debe especificarse. Además, para cada patrón distinto de valores en las variables independientes, se supone que las respuestas son variables multinomiales independientes.

**Procedimientos relacionados.** La regresión logística nominal utiliza modelos similares para las variables dependientes nominales.

### **Obtener una regresión ordinal**

- ▶ Elija en los menús:  
Analizar  
Regresión  
Ordinal...

Figura 27-1  
Cuadro de diálogo Regresión ordinal



- ▶ Seleccione una variable dependiente.
- ▶ Pulse en Aceptar.

## **Regresión ordinal: Opciones**

El cuadro de diálogo Opciones le permite ajustar los parámetros utilizados en el algoritmo de estimación iterativo, seleccionar un nivel de confianza para las estimaciones de los parámetros y seleccionar una función de enlace.

Figura 27-2  
Cuadro de diálogo Regresión ordinal de muestras complejas: Opciones

**Iteraciones.** El algoritmo iterativo puede personalizarse.

- **Iteraciones máximas.** Especifique un número entero no negativo. Si se especifica el 0, el procedimiento devolverá las estimaciones iniciales.
- **Máxima subdivisión por pasos.** Especifique un número entero positivo.
- **Convergencia del logaritmo de la verosimilitud.** El algoritmo se detiene si el cambio absoluto o relativo en el logaritmo de la verosimilitud es inferior a este valor. Si se especifica 0, no se utiliza el criterio.
- **Convergencia de los parámetros.** El algoritmo se detiene si el cambio absoluto o relativo en cada una de las estimaciones de los parámetros es inferior a este valor. Si se especifica 0, no se utiliza el criterio.

**Intervalo de confianza.** Especifique un valor mayor o igual a 0 e inferior a 100.

**Delta.** El valor añadido a las frecuencias de casilla de cero. Especifique un valor no negativo inferior a 1.

**Tolerancia para la singularidad.** Utilizada para comprobar los predictores con alta dependencia. Seleccione un valor en la lista de opciones.

**Función de enlace.** La función de enlace es una transformación de las probabilidades acumuladas que permiten la estimación del modelo. Existen cinco funciones de enlace que se resumen en la siguiente tabla.

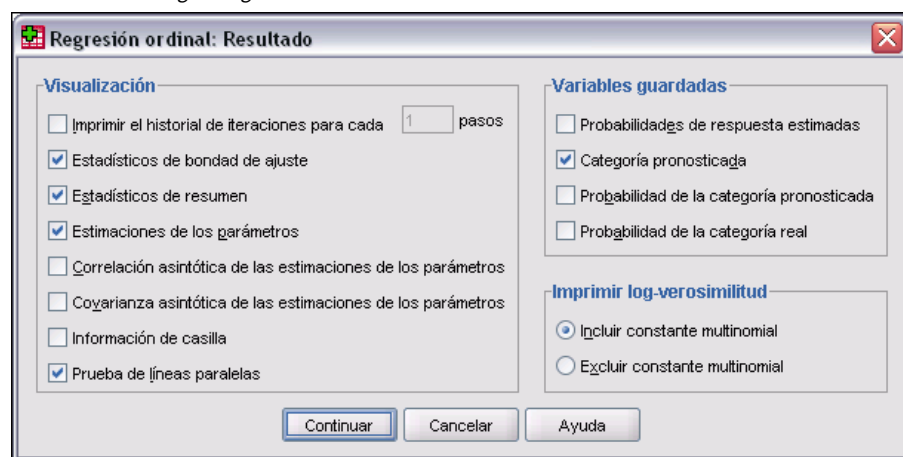
Función	Forma	Aplicación habitual
Logit	$\log\left(\frac{p}{1-p}\right)$	Categorías distribuidas de forma uniforme
Log-log complementario	$\log(\log(1-p))$	Categorías más altas más probables
Log-log negativo	$\log(-\log(p))$	Categorías más bajas más probables

Función	Forma	Aplicación habitual
Probit	$I(\cdot)$	La variable latente sigue una distribución normal
Cauchit (Cauchy inversa)	$\tan((\cdot - 0.5))$	La variable latente tiene muchos valores extremos

## Resultados de la regresión ordinal

El cuadro de diálogo Resultados le permite generar tablas que se pueden visualizar en el Visor y guardar variables en el archivo de trabajo.

Figura 27-3  
Cuadro de diálogo Regresión ordinal: Resultados



**Mostrar.** Genera tablas correspondientes a:

- **Imprimir el historial de iteraciones.** El logaritmo de la verosimilitud y las estimaciones de los parámetros se imprimen con la frecuencia de iteraciones a imprimir especificada. Siempre se imprimen la primera y la última iteración.
- **Estadísticos de bondad de ajuste.** Estadísticos chi-cuadrado de Pearson y chi-cuadrado de la razón de verosimilitud. Estos estadísticos se calculan según la clasificación especificada en la lista de variables.
- **Estadísticos de resumen.**  $R^2$  de Cox y Snell, de Nagelkerke y de McFadden.
- **Estimaciones de los parámetros.** Estimaciones de los parámetros, errores típicos e intervalos de confianza.
- **Correlación asintótica de las estimaciones de los parámetros.** Matriz de las correlaciones entre las estimaciones de los parámetros.
- **Covarianza asintótica de las estimaciones de los parámetros.** Matriz de las covarianzas entre las estimaciones de los parámetros.
- **Información de casilla.** Frecuencias esperadas y observadas y frecuencias acumuladas, residuos de Pearson para las frecuencias y las frecuencias acumuladas, probabilidades esperadas y observadas y probabilidades esperadas y observadas de cada categoría de respuesta según el patrón en las covariables. Tenga en cuenta que en el caso de modelos con muchos patrones

de covariables (por ejemplo, modelos con covariables continuas), esta opción puede generar una tabla grande y poco manejable.

- **Prueba de líneas paralelas.** Prueba correspondiente a la hipótesis de que los parámetros de ubicación son equivalentes en todos los niveles de la variable dependiente. Esta prueba está disponible únicamente para el modelo de sólo ubicación.

**Variables guardadas.** Guarda las siguientes variables en el archivo de trabajo:

- **Probabilidades de respuesta estimadas.** Probabilidades estimadas por el modelo para la clasificación de un patrón de factor/covariable en las categorías de respuesta. El número de probabilidades es igual al número de categorías de respuesta.
- **Categoría pronosticada.** La categoría de respuesta con la mayor probabilidad estimada para un patrón de factor/covariable.
- **Probabilidad de la categoría pronosticada.** Probabilidades estimada de la clasificación de un patrón factor/covariable en la categoría pronosticada. Esta probabilidad también es el máximo de las probabilidades estimadas para el patrón de factor/covariable.
- **Probabilidad de la categoría real.** Probabilidad estimada de la clasificación de un patrón factor/covariable en la categoría real.

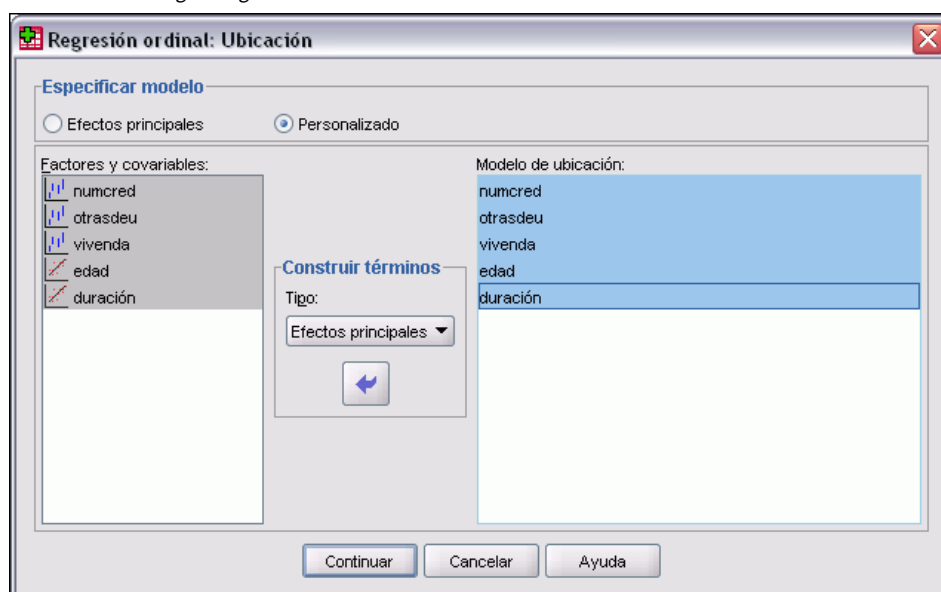
**Imprimir log-verosimilitud.** Controla la representación del logaritmo de la verosimilitud. Incluir constante multinomial ofrece el valor completo de la verosimilitud. Para comparar los resultados con los productos que no incluyan la constante, puede seleccionar la opción de excluirla.

## Modelo de ubicación de la regresión ordinal

El cuadro de diálogo Ubicación le permite especificar el modelo de ubicación para el análisis.

Figura 27-4

Cuadro de diálogo Regresión ordinal: Ubicación



**Especificar modelo.** Un modelo de efectos principales contiene los efectos principales de las covariables y los factores, pero no contiene efectos de interacción. Puede crear un modelo personalizado para especificar subconjuntos de interacciones entre los factores o bien interacciones entre las covariables.

**Factores y covariables.** Muestra una lista de los factores y las covariables.

**Modelo de ubicación.** El modelo depende de los efectos principales y de los de interacción que seleccione.

## Construir términos

Para las covariables y los factores seleccionados:

**Interacción.** Crea el término de interacción de mayor nivel con todas las variables seleccionadas. Este es el método por defecto.

**Efectos principales.** Crea un término de efectos principales para cada variable seleccionada.

**Todas de 2.** Crea todas las interacciones dobles posibles de las variables seleccionadas.

**Todas de 3.** Crea todas las interacciones triples posibles de las variables seleccionadas.

**Todas de 4.** Crea todas las interacciones cuádruples posibles de las variables seleccionadas.

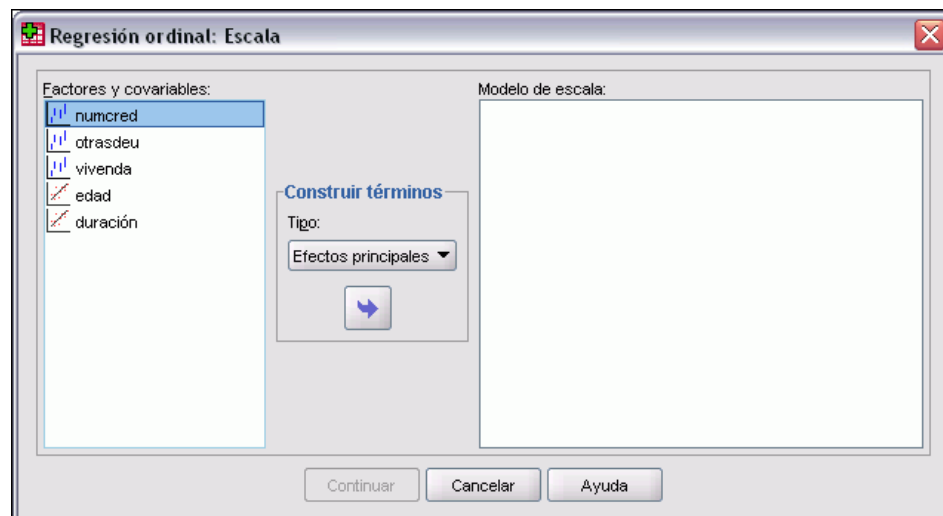
**Todas de 5.** Crea todas las interacciones quíntuples posibles de las variables seleccionadas.

## Modelo de escala de la regresión ordinal

El cuadro de diálogo Escala le permite especificar el modelo de escala para el análisis.

Figura 27-5

Cuadro de diálogo Regresión ordinal: Escala



**Factores y covariables.** Muestra una lista de los factores y las covariables.

**Modelo de escala.** El modelo depende de los efectos principales y de los de interacción que seleccione.

## **Construir términos**

Para las covariables y los factores seleccionados:

**Interacción.** Crea el término de interacción de mayor nivel con todas las variables seleccionadas. Este es el método por defecto.

**Efectos principales.** Crea un término de efectos principales para cada variable seleccionada.

**Todas de 2.** Crea todas las interacciones dobles posibles de las variables seleccionadas.

**Todas de 3.** Crea todas las interacciones triples posibles de las variables seleccionadas.

**Todas de 4.** Crea todas las interacciones cuádruples posibles de las variables seleccionadas.

**Todas de 5.** Crea todas las interacciones quíntuples posibles de las variables seleccionadas.

## **Funciones adicionales del comando PLUM**

Se puede personalizar la regresión ordinal si se pegan las selecciones en una ventana de sintaxis y se edita la sintaxis del comando PLUM resultante. Con el lenguaje de sintaxis de comandos también podrá:

- Crear contrastes de hipótesis personalizados especificando las hipótesis nulas como combinaciones lineales de los parámetros.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

## ***Estimación curvilínea***

El procedimiento Estimación Curvilínea genera estadísticos de estimación curvilínea por regresión y gráficos relacionados para 11 modelos diferentes de estimación curvilínea por regresión. Se produce un modelo diferente para cada variable dependiente. También se pueden guardar valores pronosticados, residuos e intervalos pronosticados como nuevas variables.

**Ejemplo.** Un proveedor de servicios de Internet realiza un seguimiento del porcentaje de tráfico de correo electrónico infectado de virus en la red a lo largo del tiempo. Un diagrama de dispersión revela que la relación es no lineal. Se puede ajustar un modelo lineal a los datos y comprobar la validez de los supuestos y la bondad de ajuste del modelo.

**Estadísticos.** Para cada modelo: coeficientes de regresión,  $R$  múltiple,  $R^2$  cuadrado,  $R^2$  cuadrado corregida, error típico de la estimación, tabla de análisis de varianza, valores pronosticados, residuos e intervalos de pronóstico. Modelos: lineal, logarítmico, inverso, cuadrático, cúbico, de potencia, compuesto, curva-S, logístico, de crecimiento y exponencial.

**Datos.** Las variables dependiente e independientes deben ser cuantitativas. Si selecciona Tiempo del conjunto de datos activo como variable independiente (en lugar de una variable), el procedimiento Estimación curvilínea generará una variable de tiempo en la que la distancia temporal entre los casos es uniforme. Si se selecciona Tiempo, la variable dependiente debe ser una medida de serie temporal. El análisis de series temporales requiere una estructura particular para los archivos de datos, de manera que cada caso (cada fila) represente un conjunto de observaciones en un momento determinado del tiempo y que la distancia temporal entre los casos sea uniforme.

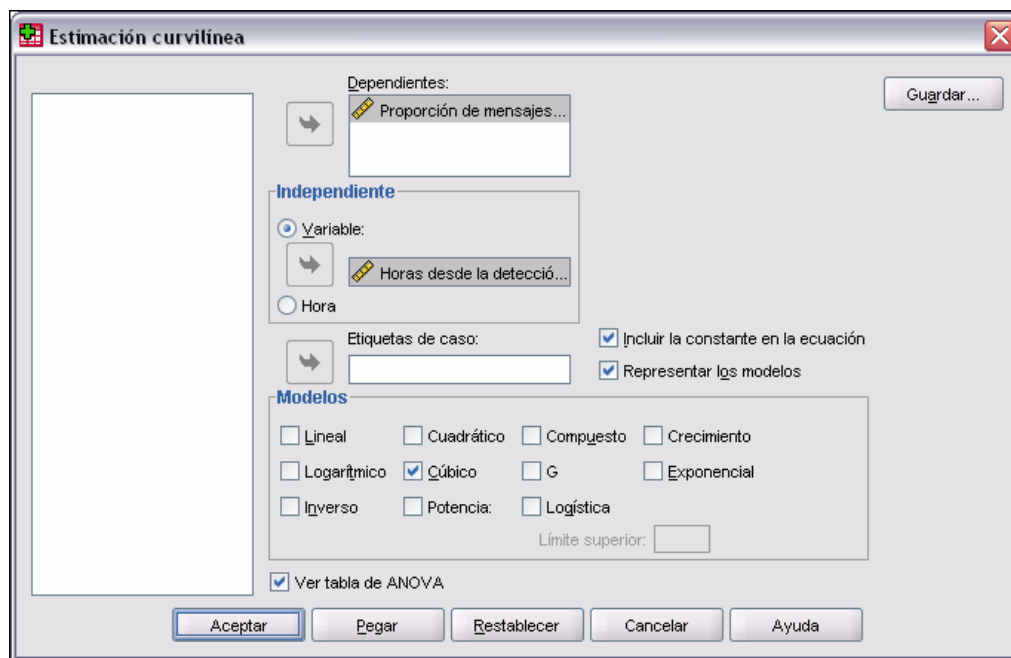
**Supuestos.** Represente los datos gráficamente para determinar cómo se relacionan las variables dependientes e independiente (linealmente, exponencialmente, etc.). Los residuos de un buen modelo deben distribuirse de forma aleatoria y normal. Si se utiliza un modelo lineal, se deben cumplir los siguientes supuestos: Para cada valor de la variable independiente, la distribución de la variable dependiente debe ser normal. La varianza de distribución de la variable dependiente debe ser constante para todos los valores de la variable independiente. La relación entre la variable dependiente y la variable independiente debe ser lineal y todas las observaciones deben ser independientes.

### ***Para obtener una estimación curvilínea***

- ▶ Elija en los menús:
  - Analizar
  - Regresión
  - Estimación curvilínea...



Figura 28-1  
Cuadro de diálogo Estimación curvilínea



- ▶ Seleccione una o más variables dependientes. Se produce un modelo diferente para cada variable dependiente.
- ▶ Seleccione una variable independiente (seleccione una variable del conjunto de datos activo o Tiempo).
- ▶ Si lo desea:
  - Seleccionar una variable para etiquetar los casos en los diagramas de dispersión. Para cada punto en el diagrama de dispersión, se puede utilizar la herramienta de Identificación de puntos para mostrar el valor de la variable utilizada en Etiquetas de caso.
  - Pulsar en Guardar para guardar los valores pronosticados, los residuos y los intervalos de pronóstico como nuevas variables.

También se encuentran disponibles las siguientes opciones:

- **Incluir la constante en la ecuación.** Estima un término constante en la ecuación de regresión. La constante se incluye por defecto.
- **Representar los modelos.** Representa los valores de la variable dependiente y cada modelo seleccionado frente a la variable independiente. Se genera un gráfico distinto para cada variable dependiente.
- **Ver tabla de ANOVA.** Muestra una tabla de análisis de varianza de resumen para cada modelo seleccionado.

## Modelos del procedimiento Estimación curvilínea

Se puede seleccionar uno o más modelos de estimación curvilínea por regresión. Para determinar qué modelo utilizar, represente los datos. Si las variables parecen estar relacionadas linealmente, utilice un modelo de regresión lineal simple. Cuando las variables no estén relacionadas linealmente, intente transformar los datos. Cuando la transformación no resulte útil, puede necesitar un modelo más complicado. Inspeccione un diagrama de dispersión de los datos; si el diagrama se parece a una función matemática reconocible, ajuste los datos a ese tipo de modelo. Por ejemplo, si los datos se parecen a una función exponencial, utilice un modelo exponencial.

**Lineal.** Modelo cuya ecuación es  $Y = b_0 + (b_1 * t)$ . Los valores de la serie se modelan como una función lineal del tiempo.

**Logarítmica.** Modelo cuya ecuación es  $Y = b_0 + (b_1 * \ln(t))$ .

**Inversa.** Modelo cuya ecuación es  $Y = b_0 + (b_1 / t)$ .

**Cuadrático.** Modelo cuya ecuación es  $Y = b_0 + (b_1 * t) + (b_2 * t^{**2})$ . El modelo cuadrático puede utilizarse para modelar una serie que "despega" o una serie que se amortigua.

**Cúbico.** Modelo definido por la ecuación  $Y = b_0 + (b_1 * t) + (b_2 * t^{**2}) + (b_3 * t^{**3})$ .

**Potencia.** Modelo cuya ecuación es  $Y = b_0 * (t^{**b_1})$  ó  $\ln(Y) = \ln(b_0) + (b_1 * \ln(t))$ .

**Compuesto.** Modelo cuya ecuación es  $Y = b_0 * (b_1^{**t})$  ó  $\ln(Y) = \ln(b_0) + (\ln(b_1) * t)$ .

**curva-S.** Modelo cuya ecuación es  $Y = e^{**}(b_0 + (b_1/t))$  ó  $\ln(Y) = b_0 + (b_1/t)$ .

**Logística.** Modelo cuya ecuación es  $Y = 1 / (1/u + (b_0 * (b_1^{**t})))$  o  $\ln(1/Y - 1/u) = \ln(b_0 + (\ln(b_1) * t))$  donde u es el valor del límite superior. Tras seleccionar Logístico, especifique un valor para el límite superior que será usado en la ecuación de regresión. El valor debe ser un número positivo mayor que el valor máximo de la variable dependiente.

**Crecimiento.** Modelo cuya ecuación es  $Y = e^{**}(b_0 + (b_1 * t))$  ó  $\ln(Y) = b_0 + (b_1 * t)$ .

**Exponencial.** Modelo cuya ecuación es  $Y = b_0 * (e^{**}(b_1 * t))$  ó  $\ln(Y) = \ln(b_0) + (b_1 * t)$ .

## Estimación curvilínea: Guardar

Figura 28-2

Cuadro de diálogo Estimación curvilínea: Guardar

**Guardar variables.** Para cada modelo seleccionado se pueden guardar los valores pronosticados, los residuos (el valor observado de la variable dependiente menos el valor pronosticado por el modelo) y los intervalos de pronóstico (sus límites superior e inferior). En la ventana de resultados, se muestran en una tabla los nombres de las nuevas variables y las etiquetas descriptivas.

**Pronosticar casos.** En el conjunto de datos activo, si se selecciona Tiempo como variable independiente en lugar de una variable, se puede especificar un período de predicción que vaya más allá del final de la serie temporal. Puede elegir una de las siguientes alternativas:

- **Desde el período de estimación hasta el último caso.** Pronostica los valores para todos los casos del archivo, basándose en los casos del período de estimación. El período de estimación, que se muestra en la parte inferior del cuadro de diálogo, se define con el subcuadro de diálogo Rango de la opción Seleccionar casos en el menú Datos. Si no se ha definido un período de estimación, se utilizan todos los casos para pronosticar los valores.
- **Predecir hasta.** Predice los valores hasta la fecha especificada, hora o número de observación, basándose en los casos del período de estimación. Esta característica se puede utilizar para predecir valores más allá del último caso de la serie temporal. Las variables definidas actualmente determinan los cuadros de texto disponibles para especificar el final del período de predicción. Si no existen variables de fecha definidas, se puede especificar el número de la observación (caso) final.

Utilice la opción de Definir fechas en el menú Datos para crear las variables de fecha.

# *Regresión de mínimos cuadrados parciales*

El procedimiento Regresión de mínimos cuadrados parciales estima los modelos de regresión de los mínimos cuadrados parciales (PLS, también denominados “proyección a la estructura latente”). La PLS es una técnica de predicción alternativa a la regresión de mínimos cuadrados ordinarios (OLS), a la correlación canónica o al modelado de ecuaciones estructurales, y resulta particularmente útil cuando las variables predictoras están muy correlacionadas o cuando el número de predictores es superior al número de casos.

La PLS combina las características del análisis de componentes principales y la regresión múltiple. En primer lugar, extrae un conjunto de factores latentes que explica en la mayor medida posible la covarianza entre las variables dependientes e independientes. A continuación, un paso de regresión pronostica los valores de las variables dependientes mediante la descomposición de las variables independientes.

**Disponibilidad.** La PLS es un comando de extensión que necesita que el módulo de extensión Python se haya instalado en el sistema donde se vaya a ejecutar PLS. El módulo de extensión PLS se debe instalar por separado y el instalador se puede descargar de <http://www.spss.com/devcentral>.

*Nota:* el módulo de extensión PLS depende del software Python. SPSS no es el propietario ni el licenciario del software Python. Todos los usuarios de Python deben aceptar los términos del acuerdo de licencia de Python que se indican en el sitio Web de Python. SPSS no realiza ninguna afirmación sobre la calidad del programa Python. SPSS rechaza completamente todas las responsabilidades asociadas al uso del programa Python.

**Tablas.** La proporción de la varianza explicada (por factor latente), las ponderaciones y las saturaciones de los factores latentes, la importancia de la variable independiente en proyección (VIP) y las estimaciones de los parámetros de la regresión (por variable dependiente) se generan por defecto.

**Gráficos.** La variable independiente en proyección (VIP), las puntuaciones factoriales, las ponderaciones factoriales de los tres primeros factores latentes y la distancia al modelo se generan desde la pestaña [Opciones](#).

**Nivel de medida.** Las variables (predictoras) dependientes e independientes pueden ser de escala, nominales u ordinales. El procedimiento supone que se ha asignado el nivel de medida adecuado a todas las variables, aunque puede cambiar temporalmente el nivel de medida de una variable pulsando el botón derecho la variable en la lista de variables de origen y seleccionando un nivel de medida en el menú contextual. El procedimiento trata por igual las variables categóricas (nominales u ordinales).

**Codificación de la variable categórica.** El procedimiento recodifica temporalmente las variables dependientes categóricas utilizando la codificación “una de  $c$ ” durante el procedimiento. Si hay  $c$  categorías de una variable, la variable se almacena como  $c$  vectores, con la primera categoría denotada como  $(1,0,\dots,0)$ , la siguiente categoría  $(0,1,0,\dots,0)$ , ... y la última categoría  $(0,0,\dots,0,1)$ . Las variables dependientes categóricas se representan mediante una variable auxiliar o dummy; es decir, omitiendo simplemente el indicador correspondiente a la categoría de referencia.

**Ponderaciones de frecuencia.** Los valores de ponderación se redondean al número entero más cercano antes de utilizarlos. Los casos con ponderaciones perdidas o ponderaciones inferiores a 0,5 no se emplearán en los análisis.

**Valores perdidos.** Los valores definidos como perdidos por el usuario o el sistema se consideran no válidos.

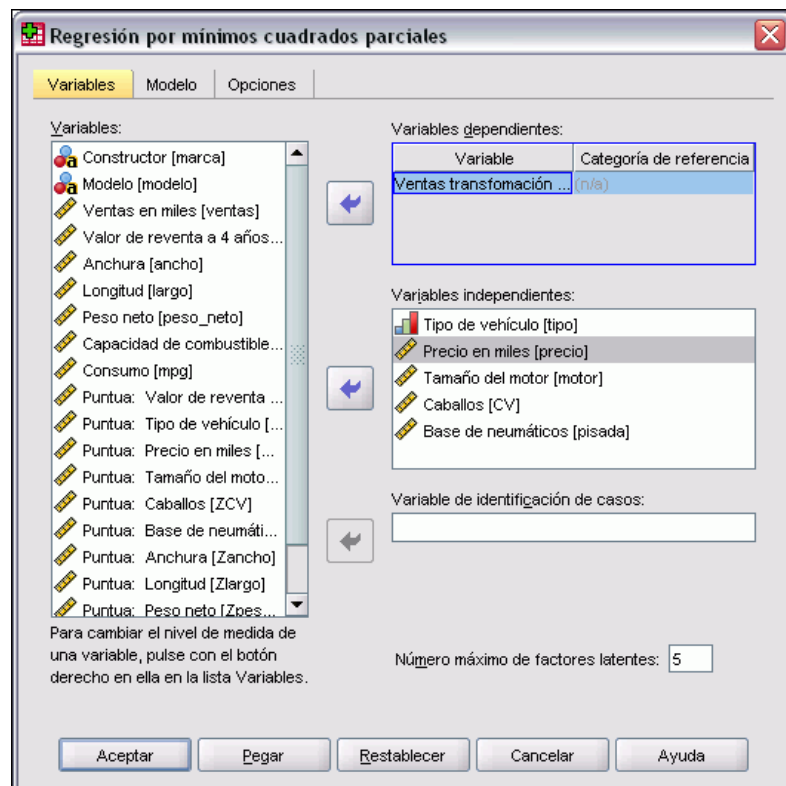
**Cambio de escala.** Todas las variables del modelo se centran y tipifican, incluidas las variables indicador que representan variables categóricas.

### Para obtener la regresión de mínimos cuadrados parciales

Elija en los menús:

Analizar  
Regresión  
Mínimos cuadrados parciales...

Figura 29-1  
Regresión de mínimos cuadrados parciales, pestaña Variables



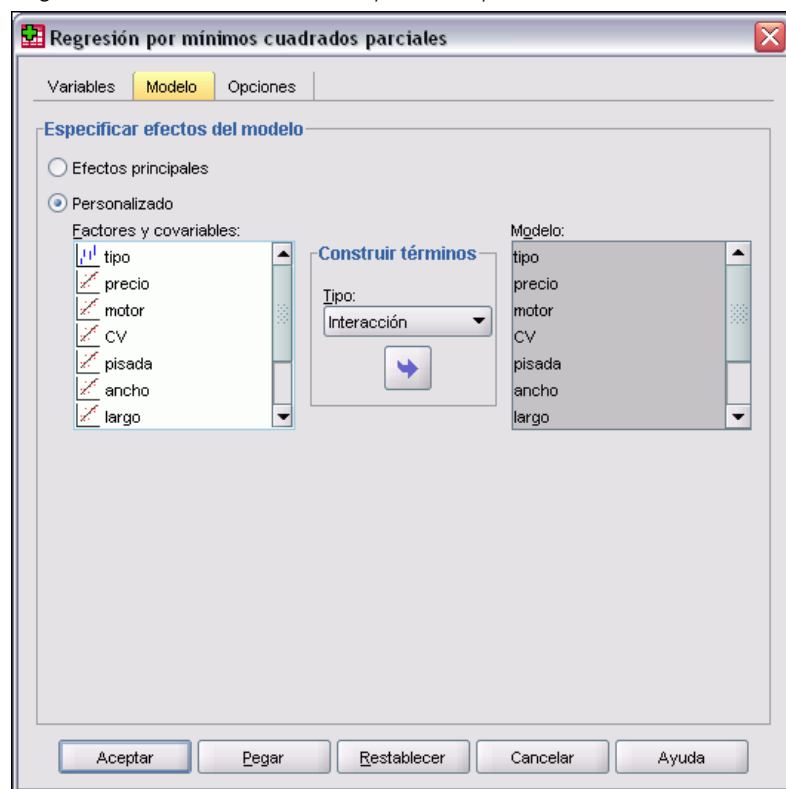
- ▶ Seleccione al menos una variable dependiente.
- ▶ Seleccione al menos una variable independiente.

Si lo desea, puede:

- Especificar una categoría de referencia para las variables dependientes categóricas (nominales u ordinales).
- Especificar la variable que se utilizará como identificador único para los resultados por casos y los conjuntos de datos guardados.
- Especificar un límite máximo para el número de factores latentes que se extraerán.

## Modelo

Figura 29-2  
Regresión de mínimos cuadrados parciales, pestaña Modelo



**Especificar efectos del modelo.** El modelo de efectos principales contiene todos los efectos principales de los factores y de las covariables. Seleccione Personalizado para especificar las interacciones. Indique todos los términos que desee incluir en el modelo.

**Factores y Covariables.** Muestra una lista de los factores y las covariables.

**Modelo.** El modelo depende de la naturaleza de los datos. Después de seleccionar Personalizado, puede elegir los efectos principales y las interacciones que sean de interés para el análisis.

### Construir términos

Para las covariables y los factores seleccionados:

**Interacción.** Crea el término de interacción de mayor nivel con todas las variables seleccionadas. Este es el método por defecto.

**Efectos principales.** Crea un término de efectos principales para cada variable seleccionada.

**Todas de 2.** Crea todas las interacciones dobles posibles de las variables seleccionadas.

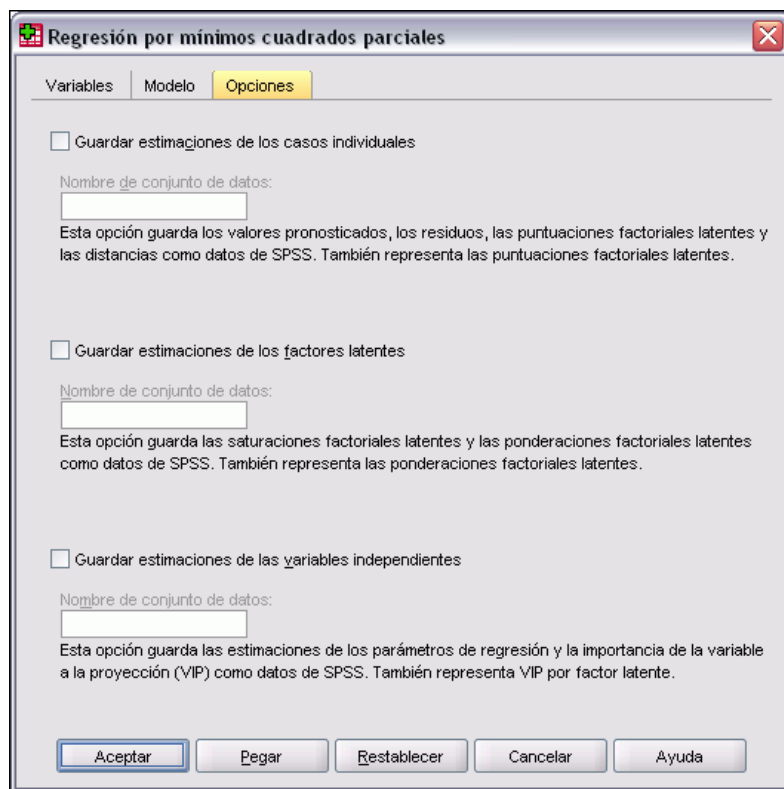
**Todas de 3.** Crea todas las interacciones triples posibles de las variables seleccionadas.

**Todas de 4.** Crea todas las interacciones cuádruples posibles de las variables seleccionadas.

**Todas de 5.** Crea todas las interacciones quíntuples posibles de las variables seleccionadas.

## Opciones

Figura 29-3  
Regresión de mínimos cuadrados parciales, pestaña Opciones



La pestaña Opciones permite al usuario guardar y representar las estimaciones de los modelos para los determinados casos, factores latentes y predictores.

Para cada tipo de datos, especifique el nombre del conjunto de datos de SPSS. Los nombres de los conjuntos de datos deben ser únicos. Si introduce el nombre de un conjunto de datos ya existente, se reemplazarán los contenidos. En otro caso, se creará un nuevo conjunto de datos.

- **Guardar estimaciones para casos individuales.** Guarda las siguientes estimaciones de modelos por casos: valores pronosticados, residuos, distancia respecto al modelo del factor latente y puntuaciones de los factores latentes. También representa las puntuaciones de los factores latentes.
- **Guardar estimaciones para factores latentes.** Guarda las saturaciones y las ponderaciones de los factores latentes. También representa las ponderaciones de factores latentes.
- **Guardar estimaciones para variables independientes.** Guarda las estimaciones de los parámetros de regresión y la importancia de la variable en la proyección (VIP). También representa la VIP por factor latente.



## **Análisis discriminante**

El análisis discriminante crea un modelo predictivo para la pertenencia al grupo. El modelo está compuesto por una función discriminante (o, para más de dos grupos, un conjunto de funciones discriminantes) basada en combinaciones lineales de las variables predictoras que proporcionan la mejor discriminación posible entre los grupos. Las funciones se generan a partir de una muestra de casos para los que se conoce el grupo de pertenencia; posteriormente, las funciones pueden ser aplicadas a nuevos casos que dispongan de medidas para las variables predictoras pero de los que se desconozca el grupo de pertenencia.

*Nota:* La variable de agrupación puede tener más de dos valores. Los códigos de la variable de agrupación han de ser números enteros y es necesario especificar sus valores máximo y mínimo. Los casos con valores fuera de estos límites se excluyen del análisis.

**Ejemplo.** Por término medio, las personas de los países de zonas templadas consumen más calorías por día que las de los trópicos, y una proporción mayor de la población de las zonas templadas vive en núcleos urbanos. Un investigador desea combinar esta información en una función para determinar cómo de bien un individuo es capaz de discriminar entre los dos grupos de países. El investigador considera además que el tamaño de la población y la información económica también pueden ser importantes. El análisis discriminante permite estimar los coeficientes de la función discriminante lineal, que tiene el aspecto de la parte derecha de una ecuación de regresión lineal múltiple. Es decir, utilizando los coeficientes  $a$ ,  $b$ ,  $c$  y  $d$ , la función es:

$$D = a * \text{clima} + b * \text{urbanos} + c * \text{población} + d * \text{producto interior bruto per capita}$$

Si estas variables resultan útiles para discriminar entre las dos zonas climáticas, los valores de  $D$  serán diferentes para los países templados y para los tropicales. Si se utiliza un método de selección de variables por pasos, quizás no se necesite incluir las cuatro variables en la función.

**Estadísticos.** Para cada variable: medias, desviaciones típicas, ANOVA univariado. Para cada análisis:  $M$  de Box, matriz de correlaciones intra-grupos, matriz de covarianzas intra-grupos, matriz de covarianzas de los grupos separados, matriz de covarianzas total. Para cada función discriminante canónica: autovalores, porcentaje de varianza, correlación canónica, lambda de Wilks, chi-cuadrado. Para cada paso: probabilidades previas, coeficientes de la función de Fisher, coeficientes de función no tipificados, lambda de Wilks para cada función canónica.

**Datos.** La variable de agrupación debe tener un número limitado de categorías distintas, codificadas como números enteros. Las variables independientes que sean nominales deben ser recodificadas a variables dummy o de contraste.

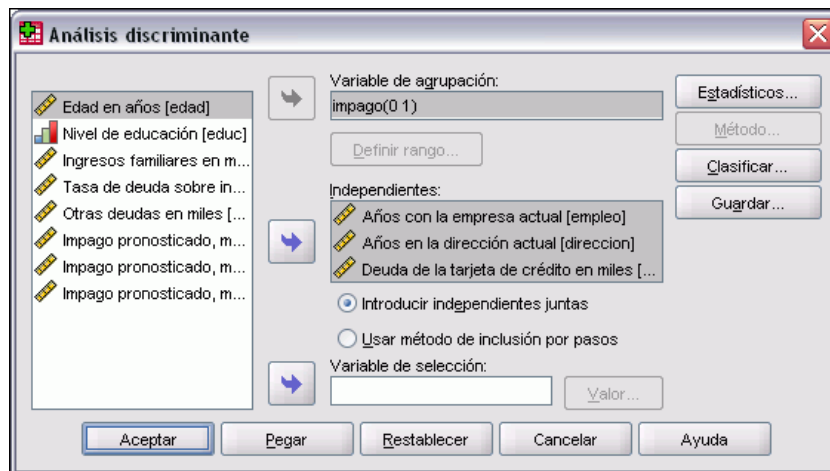
**Supuestos.** Los casos deben ser independientes. Las variables predictoras deben tener una distribución normal multivariada y las matrices de varianzas-covarianzas intra-grupos deben ser iguales en todos los grupos. Se asume que la pertenencia al grupo es mutuamente exclusiva (es decir, ningún caso pertenece a más de un grupo) y exhaustiva de modo colectivo (es decir, todos

los casos son miembros de un grupo). El procedimiento es más efectivo cuando la pertenencia al grupo es una variable verdaderamente categórica; si la pertenencia al grupo se basa en los valores de una variable continua (por ejemplo, un cociente de inteligencia alto respecto a uno bajo), considere el uso de la regresión lineal para aprovechar la información más rica ofrecida por la propia variable continua.

### Para obtener un análisis discriminante

- ▶ Elija en los menús:
  - Analizar
  - Clasificar
  - Discriminante...

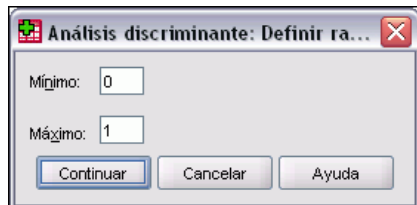
Figura 30-1  
Cuadro de diálogo Análisis discriminante



- ▶ Seleccione una variable de agrupación con valores enteros y pulse en Definir rango para especificar las categorías de interés.
- ▶ Seleccione las variables independientes o predictoras. (Si la variable de agrupación no tiene valores enteros, la opción Recodificación automática en el menú Transformar creará una variable que los tenga).
- ▶ Seleccione el método de introducción de las variables independientes.
  - **Introducir independientes juntas.** Introducir simultáneamente todas las variables independientes que satisfacen el criterio de tolerancia.
  - **Usar método de inclusión por pasos.** Utiliza el análisis por pasos para controlar la entrada y la salida de variables.
- ▶ Si lo desea, seleccione casos mediante una variable de selección.

## Análisis discriminante: Definir rango

Figura 30-2  
Cuadro de diálogo Análisis discriminante: Definir rango



Especifique los valores mínimo y máximo de la variable de agrupación para el análisis. Los casos con valores fuera de este rango no se utilizan en el análisis discriminante, pero sí se clasifican en uno de los grupos existentes a partir de los resultados que obtengan en el análisis. Los valores mínimo y máximo deben ser números enteros.

## Análisis discriminante: Seleccionar casos

Figura 30-3  
Cuadro de diálogo Análisis discriminante: Establecer valor



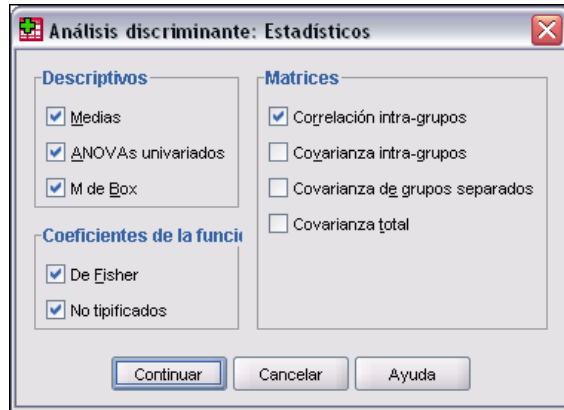
Para seleccionar casos para el análisis:

- ▶ En el cuadro de diálogo Análisis discriminante, seleccione una variable de selección.
- ▶ Pulse en Valor para introducir un número entero como valor de selección.

Sólo se utilizan los casos con el valor especificado en la variable de selección para derivar las funciones discriminantes. Tanto para los casos seleccionados como para los no seleccionados se generan resultados de clasificaciones y estadísticos. Este proceso ofrece un mecanismo para clasificar casos nuevos basados en datos previos o para dividir los datos en subconjuntos de contraste y comprobación para realizar procedimientos de validación en el modelo generado.

## Análisis discriminante: Estadísticos

Figura 30-4  
Cuadro de diálogo Análisis discriminante: Estadísticos



**Descriptivos.** Las opciones disponibles son: Medias (que incluye las desviaciones típicas), ANOVAs univariados y prueba  $M$  de Box.

- **Medias.** Muestra la media y desviación típica totales y las medias y desviaciones típicas de grupo, para las variables independientes.
- **ANOVAs univariados.** Realiza un análisis de varianza de un factor sobre la igualdad de las medias de grupo para cada variable independiente.
- **M de Box.** Contraste sobre la igualdad de las matrices de covarianza de los grupos. Para tamaños de muestra suficientemente grandes, un valor de  $p$  no significativo quiere decir que no hay evidencia suficiente de que las matrices difieran. Esta prueba es sensible a las desviaciones de la normalidad multivariada.

**Coeficientes de la función.** Las opciones disponibles son: Coeficientes de clasificación de Fisher y Coeficientes no tipificados.

- **De Fisher.** Muestra los coeficientes de la función de clasificación de Fisher que pueden utilizarse directamente para la clasificación. Se obtiene un conjunto de coeficientes para cada grupo, y se asigna un caso al grupo para el que tiene una mayor puntuación discriminante.
- **No tipificados.** Muestra los coeficientes de la función discriminante sin estandarizar (los coeficientes brutos).

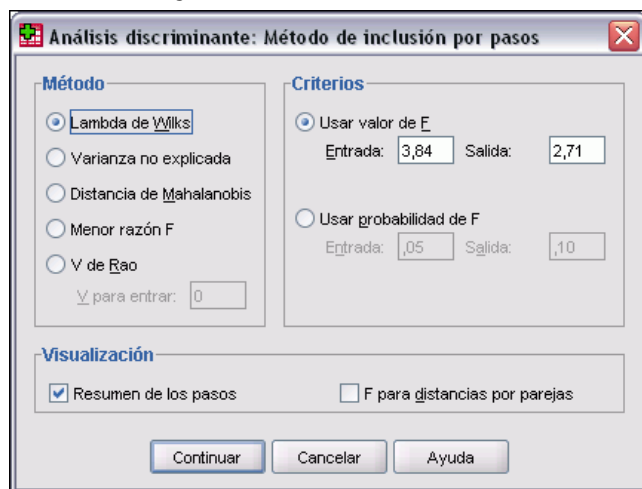
**Matrices.** Las matrices de coeficientes disponibles para las variables independientes son las de: Correlación intra-grupos, Covarianza intra-grupos, Covarianza de grupos separados y Covarianza total.

- **Correlación intra-grupos.** Muestra la matriz de correlaciones intra-grupos combinada, que se obtiene de promediar las matrices de covarianza individuales para todos los grupos antes de calcular las correlaciones.
- **Covarianza intra-grupos.** Muestra la matriz de covarianza intra-grupos combinada, la cual puede diferir de la matriz de covarianza total. La matriz se obtiene de promediar, para todos los grupos, las matrices de covarianza individuales.

- **Covarianza de grupos separados.** Muestra las matrices de covarianza de cada grupo por separado.
- **Covarianza total.** Muestra la matriz de covarianza para todos los casos, como si fueran una única muestra.

## Análisis discriminante: Método de inclusión por pasos

Figura 30-5  
Cuadro de diálogo Análisis discriminante: Método de inclusión por pasos



**Método.** Seleccione el estadístico que se va a utilizar para introducir o eliminar nuevas variables. Las alternativas disponibles son: Lambda de Wilks, Varianza no explicada, Distancia de Mahalanobis, Menor razón  $F$  y  $V$  de Rao. Con la  $V$  de Rao se puede especificar el incremento mínimo de  $V$  para introducir una variable.

- **lambda de Wilks.** Método para la selección de variables por pasos del análisis discriminante que escoge las variables para su introducción en la ecuación basándose en cuánto contribuyen a disminuir la lambda de Wilks. En cada paso se introduce la variable que minimiza la lambda de Wilks global.
- **Varianza no explicada.** En cada paso se introduce la variable que minimiza la suma de la variación no explicada entre los grupos.
- **Distancia de Mahalanobis.** Medida de cuánto difieren los valores de un caso en las variables independientes respecto al promedio para todos los casos. Una distancia de Mahalanobis grande identificará a un caso que tenga valores extremos en una o más de las variables independientes.
- **Menor razón  $F$ .** Método para la selección de variables en los análisis por pasos que se basa en maximizar la razón  $F$ , calculada a partir de la distancia de Mahalanobis entre los grupos.
- **$V$  de Rao.** Medida de las diferencias entre las medias de los grupos. También llamada traza de Lawley-Hotelling. En cada paso, se introduce la variable que maximiza el incremento en la  $V$  de Rao. Después de seleccionar esta opción, introduzca el valor mínimo que debe tener una variable para entrar en el modelo.

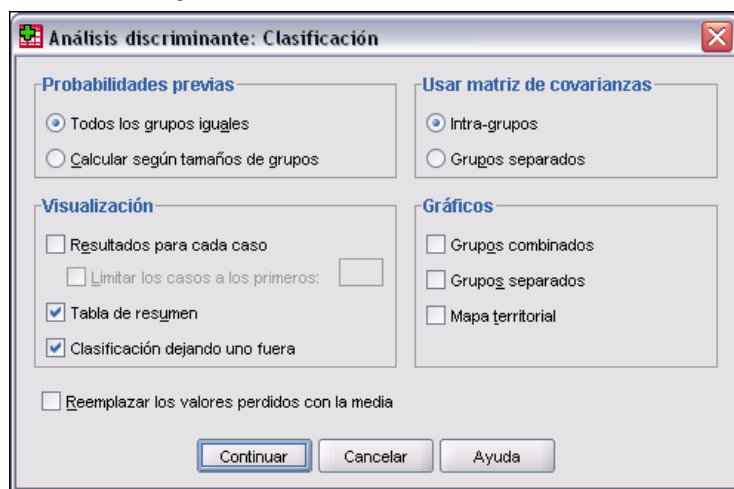
**Criterios.** Las opciones disponibles son Usar valor de F y Usar probabilidad de F. Introduzca valores para introducir y eliminar variables.

- **Usar valor de F.** Una variable se introduce en el modelo si su valor de F es mayor que el valor de entrada, y se elimina si su valor de F es menor que el valor de salida. La entrada debe ser mayor que la salida y ambos valores deben ser positivos. Para introducir más variables en el modelo, disminuya el valor de entrada. Para eliminar más variables del modelo, eleve el valor de salida.
- **Usar probabilidad de F.** Una variable se introduce en el modelo si el nivel de significación de su valor de F es menor que el valor de entrada, y se elimina si el nivel de significación de su valor de F es mayor que el valor de salida. La entrada debe ser menor que la salida y ambos valores deben ser positivos. Para introducir más variables en el modelo, eleve el valor de entrada. Para eliminar más variables del modelo, disminuya el valor de salida.

**Mostrar.** Resumen de los pasos muestra los estadísticos para todas las variables después de cada paso; F para distancias por parejas muestra una matriz de razones  $F$  por parejas para cada pareja de grupos.

## Análisis discriminante: Clasificar

Figura 30-6  
Cuadro de diálogo Análisis discriminante: Clasificar



**Probabilidades previas.** Esta opción determina si se corrigen los coeficientes de clasificación teniendo en cuenta la información previa sobre la pertenencia a los grupos.

- **Todos los grupos iguales.** Se suponen probabilidades previas iguales para todos los grupos. Esta opción no tiene ningún efecto sobre los coeficientes.
- **Calcular según tamaños de grupos.** Los tamaños de los grupos observados de la muestra determinan las probabilidades previas de la pertenencia a los grupos. Por ejemplo, si el 50% de las observaciones incluidas en el análisis corresponden al primer grupo, el 25% al segundo y el 25% al tercero, se corregirán los coeficientes de clasificación para aumentar la probabilidad de la pertenencia al primer grupo respecto a los otros dos.

**Mostrar.** Las opciones de presentación disponibles son: Resultados por casos, Tabla de resumen y Clasificación dejando uno fuera.

- **Resultados para cada caso.** Se muestran, para cada caso, los códigos del grupo real de pertenencia, el grupo pronosticado, las probabilidades posteriores y las puntuaciones discriminantes.
- **Tabla de resumen.** Número de casos correcta e incorrectamente asignados a cada uno de los grupos, basándose en el análisis discriminante. En ocasiones recibe el nombre de la "Matriz de Confusión".
- **Clasificación dejando uno fuera.** Se clasifica cada caso del análisis mediante la función derivada a partir de todos los casos, excepto el propio caso. También conocido como método-U.

**Reemplazar los valores perdidos con la media.** Seleccione esta opción para sustituir la media de una variable independiente para un valor perdido sólo durante la fase de clasificación.

**Usar matriz de covarianza.** Existe la opción de clasificar los casos utilizando una matriz de covarianza intra-grupos o una matriz de covarianza de los grupos separados.

- **Intra-grupos.** Se utiliza la matriz de covarianza intra-grupos combinada para clasificar los casos.
- **Grupos separados.** Para la clasificación se utilizan las matrices de covarianza de los grupos separados. Dado que la clasificación se basa en las funciones discriminantes y no en las variables originales, esta opción no siempre es equivalente a la discriminación cuadrática.

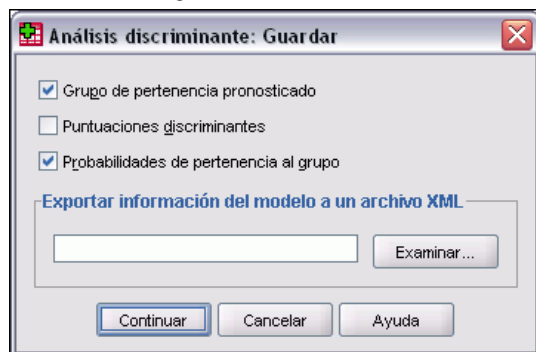
**Diagramas.** Las opciones de gráficos disponibles son: Grupos combinados, Grupos separados y Mapa territorial.

- **Grupos combinados.** Crea un diagrama de dispersión, con todos los grupos, de los valores en las dos primeras funciones discriminantes. Si sólo hay una función, en su lugar se muestra un histograma.
- **Grupos separados.** Crea diagramas de dispersión, de los grupos por separado, para los valores en las dos primeras funciones discriminantes. Si sólo hay una función se muestra un histograma en su lugar.
- **Mapa territorial.** Gráfico de las fronteras utilizadas para clasificar los casos en grupos a partir de los valores en las funciones. Los números corresponden a los grupos en los que se clasifican los casos. La media de cada grupo se indica mediante un asterisco situado dentro de sus fronteras. No se mostrará el mapa si sólo hay una función discriminante.

## Análisis discriminante: Guardar

Figura 30-7

Cuadro de diálogo Análisis discriminante: Guardar nuevas variables



Es posible añadir variables nuevas al archivo de datos activo. Las opciones disponibles son las de grupo de pertenencia pronosticado (una única variable), puntuaciones discriminantes (una variable para cada función discriminante en la solución) y probabilidades de pertenencia al grupo según las puntuaciones discriminantes (una variable para cada grupo).

También se puede exportar información del modelo al archivo especificado en formato XML (PMML). SmartScore y servidor de SPSS (un producto independiente) pueden utilizar este archivo del modelo para aplicar la información del modelo en otros archivos de datos con fines de puntuación.

## Funciones adicionales del comando DISCRIMINANT

Con el lenguaje de sintaxis de comandos también podrá:

- Realizar varios análisis discriminantes (con un comando) y controlar el orden en el que se introducen las variables (mediante el subcomando `ANALYSIS`).
- Especificar probabilidades previas para la clasificación (mediante el subcomando `PRIORS`).
- Mostrar matrices de estructura y de configuración rotadas (mediante el subcomando `ROTATE`).
- Limitar el número de funciones discriminantes extraídas (mediante el subcomando `FUNCTIONS`).
- Restringir la clasificación a los casos que están seleccionados (o no seleccionados) para el análisis (mediante el subcomando `SELECT`).
- Leer y analizar una matriz de correlaciones (mediante el subcomando `MATRIX`).
- Escribir una matriz de correlaciones para su análisis posterior (mediante el subcomando `MATRIX`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).



# Análisis factorial

El análisis factorial intenta identificar variables subyacentes, o **factores**, que expliquen la configuración de las correlaciones dentro de un conjunto de variables observadas. El análisis factorial se suele utilizar en la reducción de los datos para identificar un pequeño número de factores que explique la mayoría de la varianza observada en un número mayor de variables manifiestas. También puede utilizarse para generar hipótesis relacionadas con los mecanismos causales o para inspeccionar las variables para análisis subsiguientes (por ejemplo, para identificar la colinealidad antes de realizar un análisis de regresión lineal).

El procedimiento de análisis factorial ofrece un alto grado de flexibilidad:

- Existen siete métodos de extracción factorial disponibles.
- Existen cinco métodos de rotación disponibles, entre ellos el oblimin directo y el promax para rotaciones no ortogonales.
- Existen tres métodos disponibles para calcular las puntuaciones factoriales; y las puntuaciones pueden guardarse como variables para análisis adicionales.

**Ejemplo.** ¿Qué actitudes subyacentes hacen que las personas respondan a las preguntas de una encuesta política de la manera en que lo hacen? Examinando las correlaciones entre los elementos de la encuesta se deduce que hay una superposición significativa entre los diversos subgrupos de elementos (las preguntas sobre los impuestos tienden a estar correlacionadas entre sí, las preguntas sobre temas militares también están correlacionadas entre sí, y así sucesivamente). Con el análisis factorial, se puede investigar el número de factores subyacentes y, en muchos casos, identificar lo que los factores representan conceptualmente. Adicionalmente, se pueden calcular las puntuaciones factoriales para cada encuestado, que pueden utilizarse en análisis subsiguientes. Por ejemplo, es posible construir un modelo de regresión logística para predecir el comportamiento de voto basándose en las puntuaciones factoriales.

**Estadísticos.** Para cada variable: número de casos válidos, media y desviación típica. Para cada análisis factorial: matriz de correlaciones de variables, incluidos niveles de significación, determinante, inversa; matriz de correlaciones reproducida, que incluye anti-imagen; solución inicial (comunalidades, autovalores y porcentaje de varianza explicada); KMO (medida de la adecuación muestral de Kaiser-Meyer-Olkin) y prueba de esfericidad de Bartlett; solución sin rotar, que incluye saturaciones factoriales, comunalidades y autovalores; y solución rotada, que incluye la matriz de configuración rotada y la matriz de transformación. Para rotaciones oblicuas: las matrices de estructura y de configuración rotadas; matriz de coeficientes para el cálculo de las puntuaciones factoriales y matriz de covarianzas entre los factores. Gráficos: gráfico de sedimentación y gráfico de las saturaciones de los dos o tres primeros factores.

**Datos.** Las variables deben ser cuantitativas a nivel de **para los intervalos** o **razón**. Los datos categóricos (como la religión o el país de origen) no son adecuados para el análisis factorial. Los datos para los cuales razonablemente se pueden calcular los coeficientes de correlación de Pearson, deberían ser adecuados para el análisis factorial.

**Supuestos.** Los datos deben tener una distribución normal bivariada para cada pareja de variables y las observaciones deben ser independientes. El modelo de análisis factorial especifica que las variables vienen determinadas por los factores comunes (los factores estimados por el modelo) y por factores únicos (los cuales no se superponen entre las distintas variables observadas); las estimaciones calculadas se basan en el supuesto de que ningún factor único está correlacionado con los demás, ni con los factores comunes.

### **Para obtener un análisis factorial**

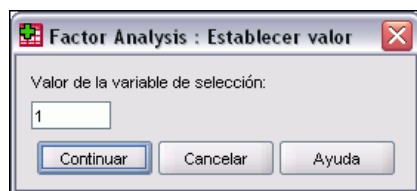
- ▶ Elija en los menús:  
   Analizar  
     Reducción de datos  
       Análisis factorial...
- ▶ Seleccione las variables para el análisis factorial.

Figura 31-1  
Cuadro de diálogo *Análisis factorial*



## **Selección de casos en el análisis factorial**

Figura 31-2  
Cuadro de diálogo *Análisis factorial: Establecer valor*



Para seleccionar casos para el análisis:

- ▶ Seleccione una variable de selección.
- ▶ Pulse en Valor para introducir un número entero como valor de selección.

En el análisis factorial, sólo se usarán los casos con ese valor para la variable de selección.

## Análisis factorial: Descriptivos

Figura 31-3  
Cuadro de diálogo Análisis factorial: Descriptivos



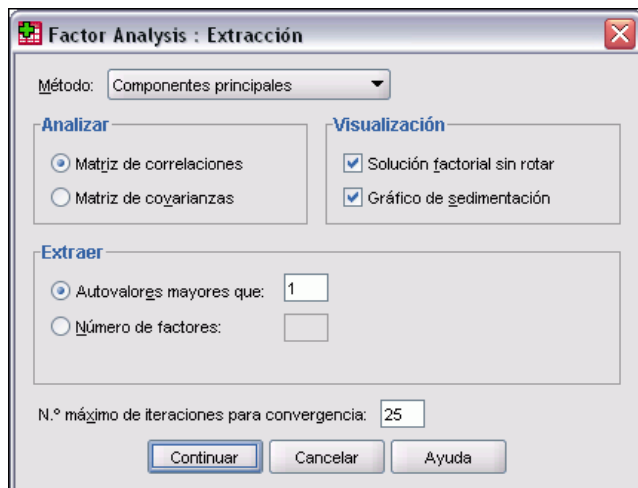
**Estadísticos.** Los descriptivos univariados incluyen la media, la desviación típica y el número de casos válidos para cada variable. La solución inicial muestra las comunalidades iniciales, los autovalores y el porcentaje de varianza explicada.

**Matriz de correlaciones.** Las opciones disponibles son: coeficientes, niveles de significación, determinante, inversa, reproducida, anti-imagen y KMO y prueba de esfericidad de Bartlett.

- **KMO y prueba de esfericidad de Bartlett.** La medida de la adecuación muestral de Kaiser-Meyer-Olkin contrasta si las correlaciones parciales entre las variables son pequeñas. La prueba de esfericidad de Bartlett contrasta si la matriz de correlaciones es una matriz identidad, que indicaría que el modelo factorial es inadecuado.
- **Reproducida.** Matriz de correlaciones estimada a partir de la solución factorial. También se muestran las correlaciones residuales (la diferencia entre la correlación observada y la reproducida).
- **Anti-imagen.** La matriz de correlaciones anti-imagen contiene los negativos de los coeficientes de correlación parcial y la matriz de covarianza anti-imagen contiene los negativos de las covarianzas parciales. En un buen modelo factorial la mayoría de los elementos no diagonales deben ser pequeños. En la diagonal de la matriz de correlaciones anti-imagen se muestra la medida de adecuación muestral para esa variable.

## Análisis factorial: Extracción

Figura 31-4  
Cuadro de diálogo Análisis factorial: Extracción



**Método.** Permite especificar el método de extracción factorial. Los métodos disponibles son: Componentes principales, Mínimos cuadrados no ponderados, Mínimos cuadrados generalizados, Máxima verosimilitud, factorización de Ejes principales, factorización Alfa y factorización Imagen.

- **Análisis de componentes principales.** Método para la extracción de factores utilizada para formar combinaciones lineales independientes de las variables observadas. La primera componente tiene la varianza máxima. Las componentes sucesivas explican progresivamente proporciones menores de la varianza y no están correlacionadas las unas con las otras. El análisis de componentes principales se utiliza para obtener la solución factorial inicial. Puede utilizarse cuando una matriz de correlaciones es singular.
- **Método de mínimos cuadrados no ponderados.** Método de extracción factorial que minimiza la suma de los cuadrados de las diferencias entre las matrices de correlaciones observada y reproducida, ignorando las diagonales.
- **Método de mínimos cuadrados generalizados.** Método de extracción de factores que minimiza la suma de los cuadrados de las diferencias entre las matrices de correlación observada y reproducida. Las correlaciones se ponderan por el inverso de su unicidad, de manera que las variables que tengan un valor alto de unicidad reciban un peso menor que aquéllas que tengan un valor bajo de unicidad.
- **Método de máxima verosimilitud.** Método de extracción factorial que proporciona las estimaciones de los parámetros que con mayor probabilidad han producido la matriz de correlaciones observada, si la muestra procede de una distribución normal multivariada. Las correlaciones se ponderan por el inverso de la unicidad de las variables, y se emplea un algoritmo iterativo.
- **Factorización de ejes principales.** Método para la extracción de factores que parte de la matriz de correlaciones original con los cuadrados de los coeficientes de correlación múltiple insertados en la diagonal principal como estimaciones iniciales de las comunalidades. Las saturaciones factoriales resultantes se utilizan para estimar de nuevo las comunalidades y

reemplazan a las estimaciones previas en la diagonal de la matriz. Las iteraciones continúan hasta que el cambio en las comunalidades, de una iteración a la siguiente, satisfaga el criterio de convergencia para la extracción.

- **Alfa.** Método de extracción factorial que considera a las variables incluidas en el análisis como una muestra del universo de las variables posibles. Este método maximiza el Alfa de Cronbach para los factores.
- **Análisis Imagen.** Método para la extracción de factores, desarrollado por Guttman y basado en la teoría de las imágenes. La parte común de una variable, llamada la imagen parcial, se define como su regresión lineal sobre las restantes variables, en lugar de ser una función de los factores hipotéticos.

**Analizar.** Permite especificar o una matriz de correlaciones o una matriz de covarianzas.

- **Matriz de correlaciones.** Es útil si las variables de su análisis se miden sobre escalas distintas.
- **Matriz de covarianzas.** Es útil si se desea aplicar el análisis factorial a varios grupos con distintas varianzas para cada variable.

**Extraer.** Se pueden retener todos los factores cuyos autovalores excedan un valor especificado o retener un número específico de factores.

**Mostrar.** Permite solicitar la solución factorial sin rotar y el gráfico de sedimentación de los autovalores.

- **Solución factorial sin rotar.** Muestra las saturaciones factoriales sin rotar (la matriz de configuración factorial), las comunalidades y los autovalores de la solución factorial.
- **Gráfico de sedimentación.** Gráfico de la varianza asociada a cada factor. Se utiliza para determinar cuántos factores deben retenerse. Típicamente el gráfico muestra la clara ruptura entre la pronunciada pendiente de los factores más importantes y el descenso gradual de los restantes (los sedimentos).

**Nº máximo de iteraciones para convergencia.** Permite especificar el número máximo de pasos que el algoritmo puede seguir para estimar la solución.

## Análisis factorial: Rotación

Figura 31-5

Cuadro de diálogo Análisis factorial: Rotación



**Método.** Permite seleccionar el método de rotación factorial. Los métodos disponibles son: varimax, equamax, quartimax, oblimin directo y promax.

- **Método varimax.** Método de rotación ortogonal que minimiza el número de variables que tienen saturaciones altas en cada factor. Simplifica la interpretación de los factores.
- **Criterio Oblimin directo.** Método para la rotación oblicua (no ortogonal). Cuando delta es igual a cero (el valor por defecto) las soluciones son las más oblicuas. A medida que delta se va haciendo más negativo, los factores son menos oblicuos. Para anular el valor por defecto 0 para delta, introduzca un número menor o igual que 0,8.
- **Método quartimax.** Método de rotación que minimiza el número de factores necesarios para explicar cada variable. Simplifica la interpretación de las variables observadas.
- **Método equamax.** Método de rotación que es combinación del método varimax, que simplifica los factores, y el método quartimax, que simplifica las variables. Se minimiza tanto el número de variables que saturan alto en un factor como el número de factores necesarios para explicar una variable.
- **Rotación Promax.** Rotación oblicua que permite que los factores estén correlacionados. Esta rotación puede calcularse más rápidamente que una rotación oblimin directa, por lo que es útil para conjuntos de datos grandes.

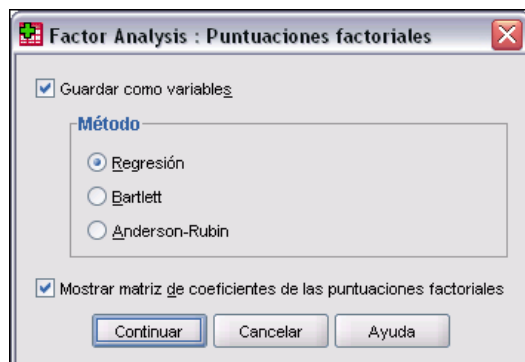
**Mostrar.** Permite incluir los resultados de la solución rotada, así como los gráficos de las saturaciones para los dos o tres primeros factores.

- **Solución rotada.** Debe seleccionarse un método de rotación para obtener la solución rotada. Para las rotaciones ortogonales, se muestran la matriz de configuración rotada y la matriz de transformación. Para las rotaciones oblicuas, se muestran las matrices de estructura, de configuración y de correlaciones entre los factores.
- **Diagrama de las saturaciones factoriales.** Representación tridimensional de las saturaciones factoriales para los tres primeros factores. Para una solución de dos factores, se representa un diagrama bidimensional. No se muestra el gráfico si sólo se extrae un factor. Si se solicita la rotación, los diagramas representan las soluciones rotadas.

**Nº máximo de iteraciones para convergencia.** Permite especificar el número máximo de pasos que el algoritmo puede seguir para llevar a cabo la rotación.

## **Análisis factorial: Puntuaciones factoriales**

Figura 31-6  
Cuadro de diálogo Análisis factorial: Puntuaciones factoriales



**Guardar como variables.** Crea una nueva variable para cada factor en la solución final.

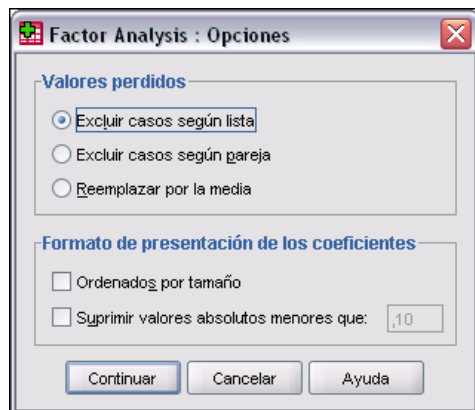
**Método.** Los métodos alternativos para calcular las puntuaciones factoriales son: regresión, Bartlett, y Anderson-Rubin.

- **Método de regresión.** Método para estimar los coeficientes de las puntuaciones factoriales. Las puntuaciones resultantes tienen de media 0 y varianza igual al cuadrado de la correlación múltiple entre las puntuaciones factoriales estimadas y los valores factoriales verdaderos. Las puntuaciones pueden estar correlacionadas incluso cuando los factores son ortogonales.
- **Puntuaciones de Bartlett.** Método para estimar los coeficientes de las puntuaciones factoriales. Las puntuaciones resultantes tienen una media de 0. Se minimiza la suma de cuadrados de los factores únicos sobre el rango de las variables.
- **Método de Anderson-Rubin.** Método para calcular los coeficientes para las puntuaciones factoriales; es una modificación del método de Bartlett, que asegura la ortogonalidad de los factores estimados. Las puntuaciones resultantes tienen una media 0, una desviación típica de 1 y no correlacionan entre sí.

**Mostrar matriz de coeficientes de las puntuaciones factoriales.** Muestra los coeficientes por los cuales se multiplican las variables para obtener puntuaciones factoriales. También muestra las correlaciones entre las puntuaciones factoriales.

## Análisis factorial: Opciones

Figura 31-7  
Cuadro de diálogo Análisis factorial: Opciones



**Valores perdidos.** Permite especificar el tratamiento que reciben los valores perdidos. Las opciones disponibles son: excluir casos **según lista**, excluir casos **según pareja** y reemplazar por la media.

**Formato de presentación de los coeficientes.** Permite controlar aspectos de las matrices de resultados. Los coeficientes se ordenan por tamaño y se suprimen aquellos cuyos valores absolutos sean menores que el valor especificado.

## Funciones adicionales del comando FACTOR

Con el lenguaje de sintaxis de comandos también podrá:

- Especificar los criterios de convergencia para la iteración durante la extracción y la rotación.

- Especificar gráficos factoriales rotados individuales.
- Especificar el número de puntuaciones factoriales que se van a guardar.
- Especificar valores diagonales para el método de factorización del eje principal.
- Escribir matrices de correlación o matrices de carga factorial en el disco para su análisis posterior.
- Leer y analizar matrices de correlación o matrices de carga factorial.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).



# ***Selección de procedimientos para la conglomeración***

Los análisis de conglomerados se pueden realizar mediante los procedimientos de análisis de conglomerados en dos fases, jerárquico o de K-medias. Cada uno de estos procedimientos emplea un algoritmo distinto en la creación de conglomerados y contiene opciones que no están disponibles en los otros.

**Análisis de conglomerados en dos fases.** En algunas aplicaciones, se puede seleccionar como método el procedimiento Análisis de conglomerados en dos fases. Ofrece una serie de funciones únicas que se detallan a continuación:

- Selección automática del número más apropiado de conglomerados y medidas para la selección de los distintos modelos de conglomerado.
- Posibilidad de crear modelos de conglomerado basados al mismo tiempo en variables categóricas y continuas.
- Posibilidad de guardar el modelo de conglomerados en un archivo XML externo y, a continuación, leer el archivo y actualizar el modelo de conglomerados con datos más recientes.

Asimismo, el procedimiento Análisis de conglomerados en dos fases puede analizar archivos de datos grandes.

**Análisis de conglomerados jerárquico.** El uso del procedimiento Análisis de conglomerados jerárquico se limita a archivos de datos más pequeños (cientos de objetos por conglomerar) y ofrece una serie de funciones únicas que se detallan a continuación:

- Posibilidad de conglomerar casos o variables.
- Posibilidad de calcular un rango de soluciones posibles y guardar los conglomerados de pertenencia para cada una de dichas soluciones.
- Distintos métodos de formación de conglomerados, transformación de variables y medida de disimilaridad entre conglomerados.

Siempre que todas las variables sean del mismo tipo, el procedimiento Análisis de conglomerados jerárquico podrá analizar variables de intervalo (continuas), de recuento o binarias.

**Análisis de conglomerados de K-medias.** El uso del procedimiento Análisis de conglomerados de K-medias se limita a datos continuos y requiere que el usuario especifique previamente el número de conglomerados y ofrece una serie de funciones únicas que se detallan a continuación:

- Posibilidad de guardar las distancias desde los centros de los conglomerados hasta los distintos objetos.
- Posibilidad de leer los centros de los conglomerados iniciales y guardar los centros de los conglomerados finales desde un archivo SPSS externo.

Asimismo, el procedimiento Análisis de conglomerados de K-medias puede analizar archivos de datos grandes.

# *Análisis de conglomerados en dos fases*

El procedimiento Análisis de conglomerados en dos fases es una herramienta de exploración diseñada para descubrir las agrupaciones naturales (o conglomerados) de un conjunto de datos que, de otra manera, no sería posible detectar. El algoritmo que emplea este procedimiento incluye varias atractivas funciones que lo hacen diferente de las técnicas de conglomeración tradicionales:

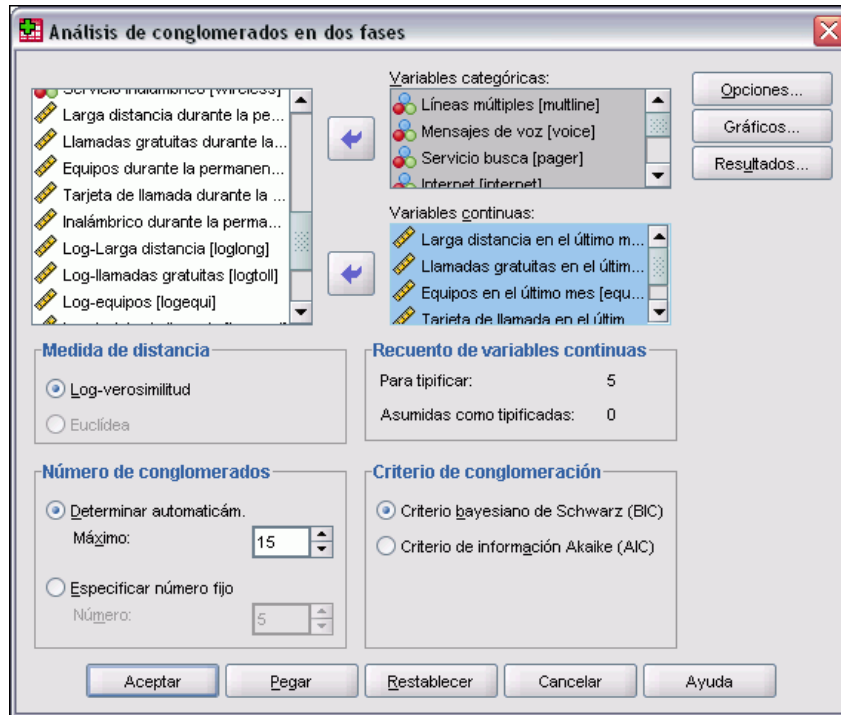
- **Tratamiento de variables categóricas y continuas.** Al suponer que las variables son independientes, es posible aplicar una distribución normal multinomial conjunta en las variables continuas y categóricas.
- **Selección automática del número de conglomerados.** Mediante la comparación de los valores de un criterio de selección del modelo para diferentes soluciones de conglomeración, el procedimiento puede determinar automáticamente el número óptimo de conglomerados.
- **Escalabilidad.** Mediante la construcción de un árbol de características de conglomerados (CF) que resume los registros, el algoritmo en dos fases puede analizar archivos de datos de gran tamaño.

**Ejemplo.** Las empresas minoristas y de venta de productos para el consumidor suelen aplicar técnicas de conglomeración a los datos que describen los hábitos de consumo, sexo, edad, nivel de ingresos, etc. de los clientes. Estas empresas adaptan sus estrategias de desarrollo de productos y de marketing en función de cada grupo de consumidores para aumentar las ventas y el nivel de fidelidad a la marca.

**Estadísticos.** El procedimiento genera criterios de información (AIC o BIC) según el número de conglomerados de la solución, las frecuencias de los conglomerados para la conglomeración final y los estadísticos descriptivos por conglomerado para la conglomeración final.

**Diagramas.** El procedimiento genera gráficos de barras para las frecuencias de los conglomerados, gráficos de sectores para las frecuencias de los conglomerados y gráficos de la importancia de las variables.

Figura 33-1  
Cuadro de diálogo *Análisis de conglomerados en dos fases*



**Medida de distancia.** Esta opción determina cómo se calcula la similaridad entre dos conglomerados.

- **Log-verosimilitud.** La medida de la verosimilitud realiza una distribución de probabilidad entre las variables. Las variables continuas se supone que tienen una distribución normal, mientras que las variables categóricas se supone que son multinomiales. Se supone que todas las variables son independientes.
- **Euclídea.** La medida euclídea es la distancia según una “línea recta” entre dos conglomerados. Sólo se puede utilizar cuando todas las variables son continuas.

**Número de conglomerados.** Esta opción permite especificar cómo se va a determinar el número de conglomerados.

- **Determinar automáticamente.** El procedimiento determinará automáticamente el número “óptimo” de conglomerados, utilizando el criterio especificado en el grupo Criterio de conglomeración. Si lo desea, introduzca un entero positivo para especificar el número máximo de conglomerados que el procedimiento debe tener en cuenta.
- **Especificar número fijo.** Permite fijar el número de conglomerados de la solución. Introduzca un número entero positivo.

**Recuento de variables continuas.** Este grupo proporciona un resumen de las especificaciones acerca de la tipificación de variables continuas realizadas en el cuadro de diálogo Opciones. Si desea obtener más información, consulte [Opciones del análisis de conglomerados en dos fases](#) en p. 390.

**Criterio de conglomeración.** Esta opción determina cómo el algoritmo de conglomeración determina el número de conglomerados. Se puede especificar tanto el criterio de información bayesiano (BIC) como el criterio de información de Akaike (AIC).

**Datos.** Este procedimiento trabaja tanto con variables continuas como categóricas. Los casos representan los objetos que se van a conglomerar y las variables representan los atributos en los que se va a basar la conglomeración.

**Orden de casos.** Observe que el árbol de características de conglomerados y la solución final pueden depender del orden de los casos. Para minimizar los efectos del orden, ordene los casos aleatoriamente. Puede que desee obtener varias soluciones distintas con los casos ordenados en distintos órdenes aleatorios para comprobar la estabilidad de una solución determinada. En situaciones en que esto resulta difícil debido a unos tamaños de archivo demasiado grandes, se pueden sustituir varias ejecuciones por una muestra de casos ordenados con distintos órdenes aleatorios.

**Supuestos.** La medida de la distancia de la verosimilitud supone que las variables del modelo de conglomerados son independientes. Además, se supone que cada variable continua tiene una distribución normal (de Gauss) y que cada variable categórica tiene una distribución multinomial. Las comprobaciones empíricas internas indican que este procedimiento es bastante robusto frente a las violaciones tanto del supuesto de independencia como de las distribuciones, pero aún así es preciso tener en cuenta hasta qué punto se cumplen estos supuestos.

Utilice el procedimiento [Correlaciones bivariadas](#) para comprobar la independencia de dos variables continuas. Utilice el procedimiento [Tablas de contingencia](#) para comprobar la independencia de dos variables categóricas. Utilice el procedimiento [Medias](#) para comprobar la independencia existente entre una variable continua y otra categórica. Utilice el procedimiento [Explorar](#) para comprobar la normalidad de una variable continua. Utilice el procedimiento [Prueba de chi-cuadrado](#) para comprobar si una variable categórica tiene una determinada distribución multinomial.

#### ***Para obtener un análisis de conglomerados en dos fases***

- ▶ Elija en los menús:

- Analizar
  - Clasificar
  - Conglomerado en dos fases...

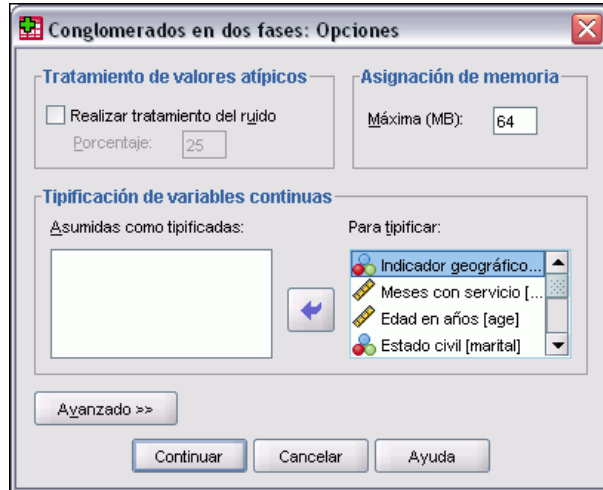
- ▶ Seleccione una o varias variables categóricas o continuas.

Si lo desea, puede:

- Ajustar los criterios utilizados para generar los conglomerados.
- Seleccionar los ajustes para el tratamiento del ruido, la asignación de memoria, la tipificación de las variables y la entrada del modelo de conglomerados.
- Solicitar gráficos y tablas opcionales.
- Guardar los resultados del modelo en el archivo de trabajo o en un archivo XML externo.

## Opciones del análisis de conglomerados en dos fases

Figura 33-2  
Cuadro de diálogo Opciones de conglomerados en dos fases



**Tratamiento de valores atípicos.** Este grupo permite tratar los valores atípicos de manera especial durante la conglomeración si se llena el árbol de características de los conglomerados (CF). El árbol CF se considera lleno si no puede aceptar ningún caso más en un nodo hoja y no hay ningún nodo hoja que se pueda dividir.

- Si selecciona el tratamiento del ruido y el árbol CF se llena, se hará volver a crecer después de colocar los casos existentes en hojas poco densas en una hoja de “ruido”. Se considera que una hoja es poco densa si contiene un número de casos inferior a un determinado porcentaje de casos del máximo tamaño de hoja. Tras volver a hacer crecer el árbol, los valores atípicos se colocarán en el árbol CF en caso de que sea posible. Si no es así, se descartarán los valores atípicos.
- Si no selecciona el tratamiento del ruido y el árbol CF se llena, se hará volver a crecer utilizando un umbral del cambio en distancia mayor. Tras la conglomeración final, los valores que no se puedan asignar a un conglomerado se considerarán como valores atípicos. Al conglomerado de valores atípicos se le asigna un número de identificación de -1 y no se incluirá en el recuento del número de conglomerados.

**Asignación de memoria.** Este grupo permite especificar la cantidad máxima de memoria en megabytes (MB) que puede utilizar el algoritmo de conglomeración. Si el procedimiento supera este máximo, utilizará el disco para almacenar la información que no se pueda colocar en la memoria. Especifique un número mayor o igual que 4.

- Consulte con el administrador del sistema si desea conocer el valor máximo que puede especificar en su sistema.
- Si este valor es demasiado bajo, es posible que el algoritmo no consiga obtener el número correcto o deseado de conglomerados.

**Tipificación de variables.** El algoritmo de conglomeración trabaja con variables continuas tipificadas. Todas las variables continuas que no estén tipificadas deben dejarse como variables en la lista Para tipificar. Para ahorrar algún tiempo y trabajo para el ordenador, puede seleccionar

todas las variables continuas que ya haya tipificado como variables en la lista Asumidas como tipificadas.

### **Opciones avanzadas**

**Criterios de ajuste del árbol CF.** Los siguientes ajustes del algoritmo de conglomeración se aplican específicamente al árbol de características de conglomerados (CF) y deberán cambiarse con cuidado:

- **Umbral del cambio en distancia inicial.** Éste es el umbral inicial que se utiliza para hacer crecer el árbol CF. Si se ha insertado una determinada hoja en el árbol CF que produciría una densidad inferior al umbral, la hoja no se dividirá. Si la densidad supera el umbral, se dividirá la hoja.
- **Nº máximo de ramas (por nodo hoja).** Número máximo de nodos filiales que puede tener una hoja.
- **Máxima profundidad de árbol.** Número máximo de niveles que puede tener un árbol CF.
- **Máximo número posible de nodos.** Indica el número máximo de nodos del árbol CF que puede generar potencialmente el procedimiento, de acuerdo con la función  $(b^{d+1} - 1) / (b - 1)$ , donde  $b$  es el número máximo de ramas y  $d$  es la profundidad máxima del árbol. Tenga en cuenta que un árbol CF excesivamente grande puede agotar los recursos del sistema y afectar negativamente al rendimiento del procedimiento. Como mínimo, cada nodo requiere 16 bytes.

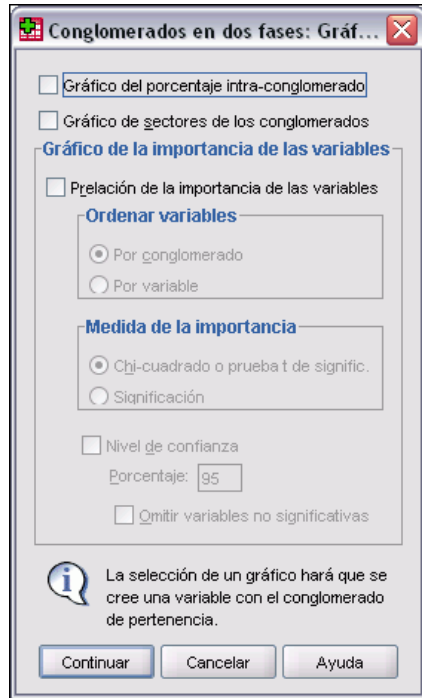
**Actualización del modelo de conglomerados.** Este grupo permite importar y actualizar un modelo de conglomerados generado en un análisis anterior. El archivo de entrada contiene el árbol CF en formato XML. A continuación, se actualizará el modelo con los datos existentes en el archivo activo. Debe seleccionar los nombres de variable en el cuadro de diálogo principal en el mismo orden en que se especificaron en el análisis anterior. El archivo XML permanecerá inalterado, a no ser que escriba específicamente la nueva información del modelo en el mismo nombre de archivo. Si desea obtener más información, consulte [Resultados de análisis de conglomerados en dos fases](#) en p. 393.

Si se ha especificado una actualización del modelo de conglomerados, se utilizarán las opciones pertenecientes a la generación del árbol CF que se especificaron para el modelo original. Concretamente, se utilizarán los ajustes del modelo guardado acerca de la medida de distancia, el tratamiento del ruido, la asignación de memoria y los criterios de ajuste del árbol CF, por lo que se ignorarán todos los ajustes de estas opciones que se hayan especificado en los cuadros de diálogo.

*Nota:* Al realizar una actualización del modelo de conglomerados, el procedimiento supone que ninguno de los casos seleccionados en el conjunto de datos activo se utilizó para crear el modelo de conglomerados original. El procedimiento también supone que los casos utilizados en la actualización del modelo proceden de la misma población que los casos utilizados para crear el modelo; es decir, se supone que las medias y las varianzas de las variables continuas y los niveles de las variables categóricas son los mismos en ambos conjuntos de casos. Si los conjuntos de casos “nuevo” y “antiguo” proceden de poblaciones heterogéneas, deberá ejecutar el procedimiento Análisis de conglomerados en dos fases para los conjuntos combinados de casos para obtener los resultados óptimos.

## Gráficos de análisis de conglomerados en dos fases

Figura 33-3  
Cuadro de diálogo Conglomerados en dos fases: Gráficos



**Gráfico del porcentaje intra-conglomerado.** Muestra los gráficos que indican la variación dentro del conglomerado de cada variable. Para cada variable categórica, se genera un gráfico de barras agrupado, mostrando la frecuencia de la categoría por identificador de conglomerado. Para cada variable continua, se genera un gráfico de barras de error, indicando las barras de error por identificador de conglomerado.

**Gráfico de sectores de los conglomerados.** Muestra un gráfico de sectores que muestra el porcentaje y las frecuencias de observaciones correspondientes a cada conglomerado.

**Gráfico de la importancia de las variables.** Muestra varios gráficos diferentes que indican la importancia de cada variable dentro de cada conglomerado. Los resultados se ordenan según el nivel de importancia de cada variable.

- **Ordenar variables.** Esta opción determina si los gráficos que se crearán para cada conglomerado (Por variable) o para cada variable (Por conglomerado).
- **Medida de la importancia.** Esta opción le permite seleccionar la medida de la importancia de las variables que se va a representar en el gráfico. Chi-cuadrado o prueba t de significación muestra un estadístico chi-cuadrado de Pearson como la importancia de una variable categórica y un estadístico  $t$  como importancia de una variable continua. Significación muestra uno menos el valor  $p$  para la prueba de igualdad de las medias para una variable continua y la frecuencia esperada del conjunto de datos global para una variable categórica.
- **Nivel de confianza.** Esta opción permite establecer el nivel de confianza para la prueba de igualdad de la distribución de una variable dentro de un conglomerado frente a la distribución global de la variable. Especifique un número inferior a 100 y superior o igual a 50. El valor



del nivel de confianza se muestra como una línea vertical en los gráficos de la importancia de las variables, si los gráficos se crean por variable o si se representa gráficamente la medida de la significación.

- **Omitir variables no significativas.** Las variables que no son significativas para el nivel de confianza especificado no aparecen en los gráficos de la importancia de las variables.

## Resultados de análisis de conglomerados en dos fases

Figura 33-4

Cuadro de diálogo Conglomerados en dos fases: Resultados



**Estadísticos.** Este grupo proporciona opciones para la presentación de tablas con los resultados de la conglomeración. Se generan los estadísticos descriptivos y las frecuencias de los conglomerados para el modelo de conglomerados final, mientras que la tabla de criterio de información muestra los resultados correspondientes a varias soluciones de conglomerados.

- **Descriptivos por conglomerado.** Muestra dos tablas que describen las variables de cada conglomerado. En una tabla, se informa de las medias y las desviaciones típicas para las variables continuas por conglomerado. La otra tabla informa de las frecuencias de las variables categóricas por conglomerado.
- **Frecuencias de los conglomerados.** Muestra una tabla que informa del número de observaciones existentes en cada conglomerado.
- **Criterio de información (AIC o BIC).** Muestra una tabla que contiene los valores de AIC o BIC, dependiendo del criterio elegido en el cuadro de diálogo principal, para diferentes números de conglomerados. Sólo se proporciona esta tabla cuando el número de conglomerados se determina automáticamente. Si el número de conglomerados es fijo, se ignora este ajuste y no se proporciona la tabla.

**Archivo de datos de trabajo.** Este grupo permite guardar las variables en el conjunto de datos activo.

- **Crear variable del conglomerado de pertenencia.** Esta variable contiene un número de identificación de conglomerado para cada caso. El nombre de esta variable es *tsc\_n*, donde *n* es un número entero positivo que indica el ordinal de la operación de almacenamiento del conjunto de datos activo realizada por este procedimiento en una determinada sesión.

**Archivos XML.** El modelo de conglomerados final y el árbol CF son dos tipos de archivos de resultados que se pueden exportar en formato XML.

- **Exportar modelo final.** También se puede exportar el modelo de conglomerado final al archivo especificado en formato XML (PMML). SmartScore y servidor de SPSS (un producto independiente) pueden utilizar este archivo del modelo para aplicar la información del modelo en otros archivos de datos con fines de puntuación.
- **Exportar árbol CF.** Esta opción permite guardar el estado actual del árbol de conglomerados y actualizarlo más tarde utilizando nuevos datos.

## ***Análisis de conglomerados jerárquico***

Este procedimiento intenta identificar grupos relativamente homogéneos de casos (o de variables) basándose en las características seleccionadas, mediante un algoritmo que comienza con cada caso (o cada variable) en un conglomerado diferente y combina los conglomerados hasta que sólo queda uno. Es posible analizar las variables brutas o elegir de entre una variedad de transformaciones de estandarización. Las medidas de distancia o similitud se generan mediante el procedimiento Proximidades. Los estadísticos se muestran en cada etapa para ayudar a seleccionar la mejor solución.

**Ejemplo.** ¿Existen grupos identificables de programas televisivos que atraigan a audiencias similares dentro de cada grupo? Con el análisis de conglomerados jerárquico, podría agrupar los programas de TV (los casos) en grupos homogéneos basados en las características del espectador. Esto se puede utilizar para identificar segmentos de mercado. También puede agrupar ciudades (los casos) en grupos homogéneos, de manera que se puedan seleccionar ciudades comparables para probar diversas estrategias de marketing.

**Estadísticos.** Historial de conglomeración, matriz de distancias (o similitudes) y pertenencia a los conglomerados para una solución única o una serie de soluciones. Gráficos: dendrogramas y diagramas de témpanos.

**Datos.** Las variables pueden ser cuantitativas, binarias o datos de recuento (frecuencias). El escalamiento de las variables es un aspecto importante, ya que las diferencias en el escalamiento pueden afectar a las soluciones de conglomeración. Si las variables muestran grandes diferencias en el escalamiento (por ejemplo, una variable se mide en dólares y la otra se mide en años), debería considerar la posibilidad de estandarizarlas (esto puede llevarse a cabo automáticamente mediante el propio procedimiento Análisis de conglomerados jerárquico).

**Orden de casos.** Si hay distancias empatadas o similitudes en los datos de entrada o si éstas se producen entre los conglomerados actualizados durante la unión, la solución de conglomerado resultante puede depender del orden de los casos del archivo. Puede que desee obtener varias soluciones distintas con los casos ordenados en distintos órdenes aleatorios para comprobar la estabilidad de una solución determinada.

**Supuestos.** Las medidas de distancia o similitud empleadas deben ser adecuadas para los datos analizados (véase el procedimiento Proximidades para obtener más información sobre la elección de las medidas de distancia y similitud). Asimismo, debe incluir todas las variables relevantes en el análisis. Si se omiten variables de interés la solución obtenida puede ser equívoca. Debido a que el análisis de conglomerados jerárquico es un método exploratorio, los resultados deben considerarse provisionales hasta que sean confirmados mediante otra muestra independiente.

**Para obtener un análisis de conglomerados jerárquico**

- ▶ Elija en los menús:
  - Analizar
  - Clasificar
  - Conglomerados jerárquicos...

Figura 34-1

*Cuadro de diálogo Análisis de conglomerados jerárquico*

- ▶ Si está aglomerando casos, seleccione al menos una variable numérica. Si está aglomerando variables, seleccione al menos tres variables numéricas.

Si lo desea, puede seleccionar una variable de identificación para etiquetar los casos.

## Análisis de conglomerados jerárquico: Método

Figura 34-2

Cuadro de diálogo Análisis de conglomerados jerárquico: Método

**Método de conglomeración.** Las opciones disponibles son: Vinculación inter-grupos, Vinculación intra-grupos, Vecino más próximo, Vecino más lejano, Agrupación de centroides, Agrupación de medianas y Método de Ward.

**Medida.** Permite especificar la medida de distancia o similitud que será empleada en la aglomeración. Seleccione el tipo de datos y la medida de distancia o similitud adecuada:

- **Intervalo.** Distancia euclídea, Distancia euclídea al cuadrado, Coseno, Correlación de Pearson, Chebychev, Bloque, Minkowski y Personalizada.
- **Recuentos.** Las opciones disponibles son: Medida de chi-cuadrado y Medida de phi-cuadrado.
- **Binary.** Las opciones disponibles son: Distancia euclídea, Distancia euclídea al cuadrado, Diferencia de tamaño, Diferencia de configuración, Varianza, Dispersión, Forma, Concordancia simple, Correlación phi de 4 puntos, Lambda,  $D$  de Anderberg, Dice, Hamann, Jaccard, Kulczynski 1, Kulczynski 2, Lance y Williams, Ochiai, Rogers y Tanimoto, Russel y Rao, Sokal y Sneath 1, Sokal y Sneath 2, Sokal y Sneath 3, Sokal y Sneath 4, Sokal y Sneath 5,  $Y$  de Yule y  $Q$  de Yule.

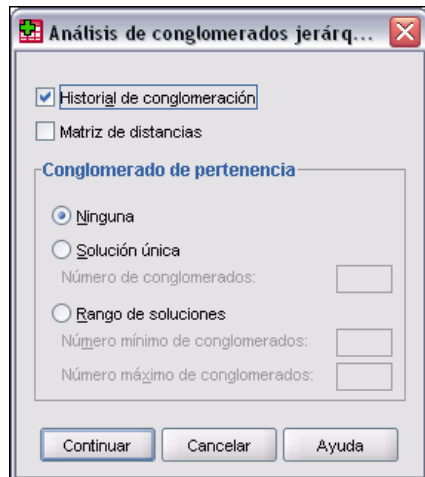
**Transformar valores.** Permite estandarizar los valores de los datos, para los casos o las variables, antes de calcular las proximidades (no está disponible para datos binarios). Los métodos disponibles de estandarización son: Puntuaciones  $z$ , Rango -1 a -1, Rango 0 a 1, Magnitud máxima de 1, Media de 1 y Desviación típica 1.

**Transformar medidas.** Permite transformar los valores generados por la medida de distancia. Se aplican después de calcular la medida de distancia. Las alternativas disponibles son valores absolutos, cambiar el signo y cambiar la escala al rango 0-1.

## Análisis de conglomerados jerárquico: Estadísticos

Figura 34-3

Cuadro de diálogo Análisis de conglomerados jerárquico:



**Historial de conglomeración.** Muestra los casos o conglomerados combinados en cada etapa, las distancias entre los casos o los conglomerados que se combinan, así como el último nivel del proceso de aglomeración en el que cada caso (o variable) se unió a su conglomerado correspondiente.

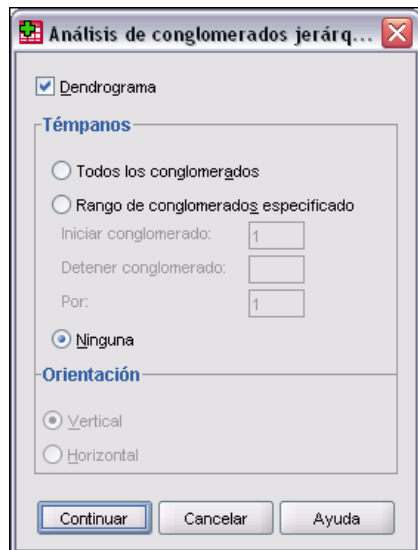
**Matriz de distancias.** Proporciona las distancias o similitudes entre los elementos.

**Conglomerado de pertenencia.** Muestra el conglomerado al cual se asigna cada caso en una o varias etapas de la combinación de los conglomerados. Las opciones disponibles son: Solución única y Rango de soluciones.

## Análisis de conglomerados jerárquico: Gráficos

Figura 34-4

Cuadro de diálogo Análisis de conglomerados jerárquico: Gráficos



**Dendrograma.** Muestra un **dendograma**. Los dendogramas pueden emplearse para evaluar la cohesión de los conglomerados que se han formado y proporcionar información sobre el número adecuado de conglomerados que deben conservarse.

**Témpanos.** Muestra un **diagrama de témpanos**, que incluye todos los conglomerados o un rango especificado de conglomerados. Los diagramas de témpanos muestran información sobre cómo se combinan los casos en los conglomerados, en cada iteración del análisis. La orientación permite seleccionar un diagrama vertical u horizontal.

## Análisis de conglomerados jerárquico: Guardar variables nuevas

Figura 34-5

Cuadro de diálogo Análisis de conglomerados jerárquico: Guardar variables nuevas



**Conglomerado de pertenencia.** Permite guardar los conglomerados de pertenencia para una solución única o un rango de soluciones. Las variables guardadas pueden emplearse en análisis posteriores para explorar otras diferencias entre los grupos.

## ***Funciones adicionales de la sintaxis de comandos CLUSTER***

El procedimiento Conglomerado jerárquico utiliza la sintaxis de comandos CLUSTER. Con el lenguaje de sintaxis de comandos también podrá:

- Utilizar varios métodos de agrupación en un único análisis.
- Leer y analizar una matriz de proximidades.
- Escribir una matriz de distancias para su análisis posterior.
- Especificar cualquier valor para la potencia y la raíz en la medida de distancia personalizada (potencia).
- Especificar nombres para variables guardadas.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).



# **Análisis de conglomerados de *K*-medias**

Este procedimiento intenta identificar grupos de casos relativamente homogéneos basándose en las características seleccionadas y utilizando un algoritmo que puede gestionar un gran número de casos. Sin embargo, el algoritmo requiere que el usuario especifique el número de conglomerados. Puede especificar los centros iniciales de los conglomerados si conoce de antemano dicha información. Puede elegir uno de los dos métodos disponibles para clasificar los casos: la actualización de los centros de los conglomerados de forma iterativa o sólo la clasificación. Asimismo, puede guardar la pertenencia a los conglomerados, información de la distancia y los centros de los conglomerados finales. Si lo desea, puede especificar una variable cuyos valores sean utilizados para etiquetar los resultados por casos. También puede solicitar los estadísticos  $F$  de los análisis de varianza. Aunque estos estadísticos son oportunistas (ya que el procedimiento trata de formar grupos que de hecho difieran), el tamaño relativo de los estadísticos proporciona información acerca de la contribución de cada variable a la separación de los grupos.

**Ejemplo.** ¿Cuáles son los grupos identificables de programas de televisión que atraen audiencias parecidas dentro de cada grupo? Con el análisis de conglomerados de  $k$ -medias, podría agrupar los programas de televisión (los casos) en  $k$  grupos homogéneos, basados en las características del televidente. Este proceso se puede utilizar para identificar segmentos de mercado. También puede agrupar ciudades (los casos) en grupos homogéneos, de manera que se puedan seleccionar ciudades comparables para probar diversas estrategias de marketing.

**Estadísticos.** Para la solución completa: centros iniciales de los conglomerados, tabla de ANOVA. Para cada caso: información del conglomerado, distancia desde el centro del conglomerado.

**Datos.** Las variables deben ser cuantitativas en el nivel de intervalo o de razón. Si las variables son binarias o recuentos, utilice el procedimiento Análisis de conglomerados jerárquicos.

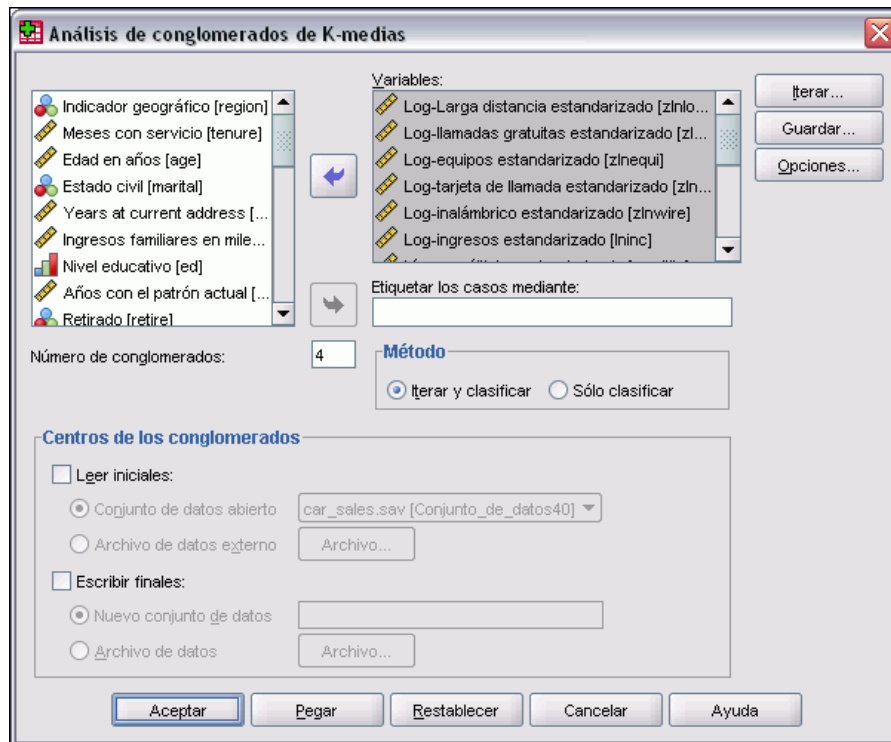
**Orden de casos y centro de conglomerados iniciales.** El algoritmo por defecto para elegir centros de conglomerados iniciales no es invariable con respecto a la ordenación de casos. La opción Usar medias actualizadas del cuadro de diálogo Iterar hace que la solución resultante dependa potencialmente del orden de casos con independencia de cómo se eligen los centros de conglomerados iniciales. Si va a utilizar alguno de estos métodos, puede que desee obtener varias soluciones distintas con los casos ordenados en distintos órdenes aleatorios para comprobar la estabilidad de una solución determinada. La especificación de los centros de conglomerados iniciales y la no utilización de la opción Usar medias actualizadas evita los problemas relacionados con el orden de casos. No obstante, la ordenación de los centros de conglomerados iniciales puede afectar a la solución en caso de haber distancias empatadas desde los casos a los centros de conglomerados. Para evaluar la estabilidad de una solución determinada, puede comparar los resultados de los análisis con las distintas permutaciones de los valores de centros iniciales.

**Supuestos.** Las distancias se calculan utilizando la distancia euclídea simple. Si desea utilizar otra medida de distancia o de similitud, utilice el procedimiento Análisis de conglomerados jerárquicos. El escalamiento de variables es una consideración importante. Si sus variables utilizan diferentes escalas (por ejemplo, una variable se expresa en dólares y otra, en años), los resultados podrían ser equívocos. En estos casos, debería considerar la estandarización de las variables antes de realizar el análisis de conglomerados de  $k$ -medias (esta tarea se puede hacer en el procedimiento Descriptivos). Este procedimiento supone que ha seleccionado el número apropiado de conglomerados y que ha incluido todas las variables relevantes. Si ha seleccionado un número inapropiado de conglomerados o ha omitido variables relevantes, los resultados podrían ser equívocos.

### Para obtener un análisis de conglomerados de K-medias

- Elija en los menús:
  - Analizar
  - Clasificar
  - Conglomerado de K medias...

Figura 35-1  
Cuadro de diálogo Análisis de conglomerados de K-medias



- Seleccione las variables que se van a utilizar en el análisis de conglomerados.
- Especifique el número de conglomerados. (Este número no debe ser inferior a 2 ni superior al número de casos del archivo de datos.)
- Seleccione Iterar y clasificar o Sólo clasificar.

- Si lo desea, seleccione una variable de identificación para etiquetar los casos.

## Eficacia del análisis de conglomerados de K-medias

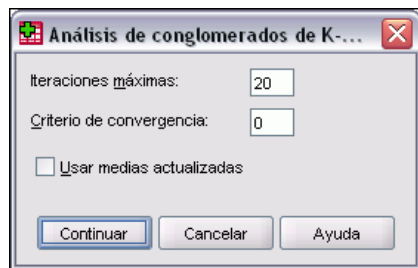
El comando de análisis de conglomerados de  $k$ -medias es eficaz principalmente porque no calcula las distancias entre todos los pares de casos, como hacen muchos algoritmos de conglomeración, como el utilizado por el comando de conglomeración jerárquica.

Para conseguir la máxima eficacia, tome una muestra de los casos y seleccione el método Iterar y clasificar para determinar los centros de los conglomerados. Seleccione Escribir finales en. A continuación, restaure el archivo de datos completo, seleccione el método Sólo clasificar y seleccione Leer iniciales de para clasificar el archivo completo utilizando los centros estimados a partir de la muestra. Se puede escribir y leer desde un archivo o conjunto de datos. Los conjuntos de datos están disponibles para su uso posterior durante la misma sesión, pero no se guardarán como archivos a menos que se hayan guardado explícitamente antes de que finalice la sesión. El nombre de un conjunto de datos debe cumplir las normas de denominación de variables. Si desea obtener más información, consulte [Nombres de variable](#) en Capítulo 5 en p. 82.

## Análisis de conglomerados de K-medias: Iterar

Figura 35-2

Cuadro de diálogo Análisis de conglomerados de K-medias: Iterar



*Nota:* estas opciones sólo están disponibles si se selecciona el método Iterar y clasificar en el cuadro de diálogo Análisis de conglomerados de K-medias.

**Nº máximo de iteraciones.** Limita el número de iteraciones en el algoritmo  $k$ -medias. La iteración se detiene después de este número de iteraciones, incluso si no se ha satisfecho el criterio de convergencia. Este número debe estar entre el 1 y el 999.

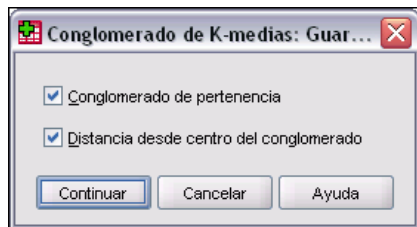
Para reproducir el algoritmo utilizado por el comando Quick Cluster en las versiones previas a la 5.0, establezca Máximo de iteraciones en 1.

**Criterio de convergencia.** Determina cuándo cesa la iteración. Representa una proporción de la distancia mínima entre los centros iniciales de los conglomerados, por lo que debe ser mayor que 0 pero no mayor que 1. Por ejemplo, si el criterio es igual a 0,02, la iteración cesará si una iteración completa no mueve ninguno de los centros de los conglomerados en una distancia superior al dos por ciento de la distancia menor entre cualquiera de los centros iniciales.

**Usar medias actualizadas.** Permite solicitar la actualización de los centros de los conglomerados tras la asignación de cada caso. Si no selecciona esta opción, los nuevos centros de los conglomerados se calcularán después de la asignación de todos los casos.

## Análisis de conglomerados de K-medias: Guardar

Figura 35-3  
Cuadro de diálogo Análisis de conglomerados de K-medias: Guardar



Puede guardar información sobre la solución como nuevas variables para que puedan ser utilizadas en análisis subsiguientes:

**Conglomerado de pertenencia.** Crea una nueva variable que indica el conglomerado final al que pertenece cada caso. Los valores de la nueva variable van desde el 1 hasta el número de conglomerados.

**Distancia desde centro del conglomerado.** Crea una nueva variable que indica la distancia euclídea entre cada caso y su centro de clasificación.

## Análisis de conglomerados de K-medias: Opciones

Figura 35-4  
Cuadro de diálogo Análisis de conglomerados de K-medias: Opciones



**Estadísticos.** Puede seleccionar los siguientes estadísticos: Centros de conglomerados iniciales, Tabla de ANOVA e Información del conglomerado para cada caso.

- **Centros de conglomerados iniciales.** Primera estimación de las medias de las variables para cada uno de los conglomerados. Por defecto se selecciona entre los datos un número de casos debidamente espaciados igual al número de conglomerados. Los centros iniciales de los conglomerados se utilizan como criterio para una primera clasificación y, a partir de ahí, se van actualizando.

- **Tabla de ANOVA.** Muestra una tabla de análisis de varianza que incluye las pruebas F univariadas para cada variable de aglomeración. Las pruebas F son sólo descriptivas y las probabilidades resultantes no deben ser interpretadas. La tabla de ANOVA no se mostrará si se asignan todos los casos a un único conglomerado.
- **Información del conglomerado para cada caso.** Muestra, para cada caso, el conglomerado final asignado y la distancia euclídea entre el caso y el centro del conglomerado utilizado para clasificar el caso. También muestra la distancia euclídea entre los centros de los conglomerados finales.

**Valores perdidos.** Las opciones disponibles son: Excluir casos según lista o Excluir casos según pareja.

- **Excluir casos según lista.** Excluye los casos con valores perdidos para cualquier variable de agrupación del análisis.
- **Excluir casos según pareja.** Asigna casos a los conglomerados en función de las distancias que se calculan desde todas las variables con valores no perdidos.

## ***Funciones adicionales del comando QUICK CLUSTER***

El procedimiento de conglomerados de K-medias utiliza la sintaxis de comandos QUICK CLUSTER. Con el lenguaje de sintaxis de comandos también podrá:

- Aceptar los primeros casos  $k$  como primeros centros de conglomerados iniciales y, por lo tanto, evitar la lectura de datos que normalmente se utiliza para calcularlos.
- Especificar los centros de conglomerados iniciales directamente como parte de la sintaxis de comandos.
- Especificar nombres para variables guardadas.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

## ***Pruebas no paramétricas***

El procedimiento Pruebas no paramétricas proporciona varias pruebas que no necesitan supuestos sobre la forma de la distribución subyacente.

**Prueba de chi-cuadrado.** Tabula una variable en categorías y calcula un estadístico de chi-cuadrado basándose en las diferencias entre las frecuencias observadas y las esperadas.

**Prueba binomial.** Compara la frecuencia observada en cada categoría de una variable dicotómica con las frecuencias esperadas en la distribución binomial.

**Prueba de rachas.** Comprueba si el orden de aparición de dos valores de una variable es aleatorio.

**Prueba de Kolmogorov-Smirnov para una muestra.** Compara la función de distribución acumulativa observada de una variable con una distribución teórica especificada, que puede ser normal, uniforme, exponencial o de Poisson.

**Pruebas para dos muestras independientes.** Compara dos grupos de casos en una variable. Se encuentran disponibles la prueba  $U$  de Mann-Whitney, la prueba de Kolmogorov-Smirnov para dos muestras, la prueba de Moses de reacciones extremas y la prueba de rachas de Wald-Wolfowitz.

**Pruebas para dos muestras relacionadas.** Compara las distribuciones de dos variables. La prueba de Wilcoxon de los rangos con signo, la prueba de signos y la prueba de McNemar.

**Pruebas para varias muestras independientes.** Compara dos o más grupos de casos en una variable. Se encuentran disponibles la prueba de Kruskal-Wallis, la prueba de la mediana y la prueba de Jonckheere-Terpstra.

**Pruebas para varias muestras relacionadas.** Compara las distribuciones de dos o más variables. Se encuentran disponibles la prueba de Friedman, la  $W$  de Kendall y la  $Q$  de Cochran.

Además, para todas las pruebas citadas anteriormente están disponibles los cuartiles y la media, la desviación típica, el mínimo, el máximo y, por último, el número de casos no perdidos.

### ***Prueba de chi-cuadrado***

El procedimiento Prueba de chi-cuadrado tabula una variable en categorías y calcula un estadístico de chi-cuadrado. Esta prueba de bondad de ajuste compara las frecuencias observadas y esperadas en cada categoría para contrastar que todas las categorías contengan la misma proporción de valores o que cada categoría contenga una proporción de valores especificada por el usuario.

**Ejemplos.** La prueba de chi-cuadrado podría utilizarse para determinar si una bolsa de caramelos contiene en igualdad de proporción caramelos de color azul, marrón, verde, naranja, rojo y amarillo. También podría utilizarse para ver si una bolsa de caramelos contiene un 5% de color

azul, un 30% de color marrón, un 10% de color verde, un 20% de color naranja, un 15% de color rojo y un 15% de color amarillo.

**Estadísticos.** Media, desviación típica, mínimo, máximo y cuartiles. Número y porcentaje de casos perdidos y no perdidos; número de casos observados y esperados de cada categoría; residuos y estadístico de chi-cuadrado.

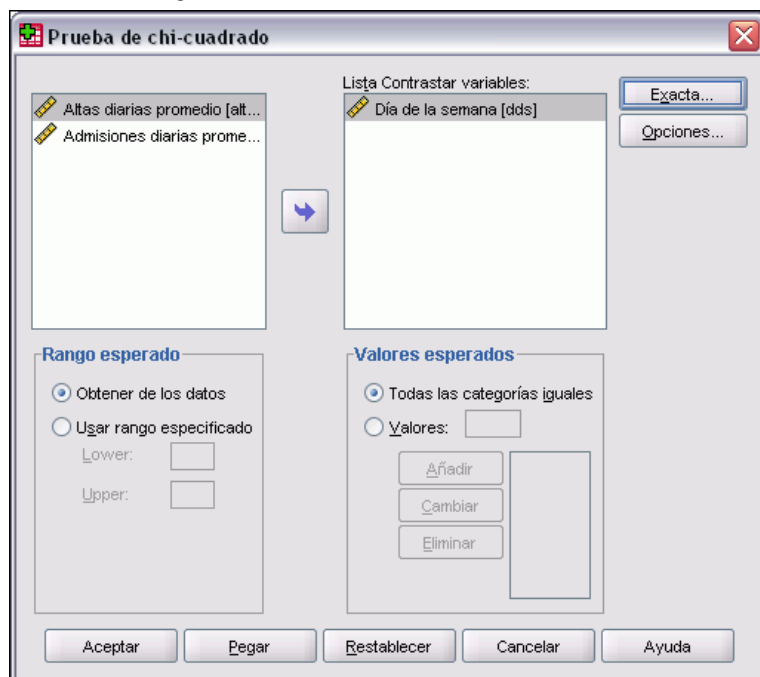
**Datos.** Use variables categóricas numéricas ordenadas o no ordenadas (niveles de medida ordinal o nominal). Para convertir las variables de cadena en variables numéricas, utilice el procedimiento Recodificación automática, disponible en el menú Transformar.

**Supuestos.** Las pruebas no paramétricas no requieren supuestos sobre la forma de la distribución subyacente. Se asume que los datos son una muestra aleatoria. Las frecuencias esperadas para cada categoría deberán ser 1 como mínimo. No más de un 20% de las categorías deberán tener frecuencias esperadas menores que 5.

### **Para obtener una prueba de chi-cuadrado**

- Elija en los menús:
  - Analizar
  - Pruebas no paramétricas
  - Chi-cuadrado...

Figura 36-1  
Cuadro de diálogo Prueba de chi-cuadrado



- Seleccione una o más variables de contraste. Cada variable genera una prueba independiente.
- Si lo desea, puede pulsar en Opciones para obtener estadísticos descriptivos, cuartiles y controlar el tratamiento de los datos perdidos.

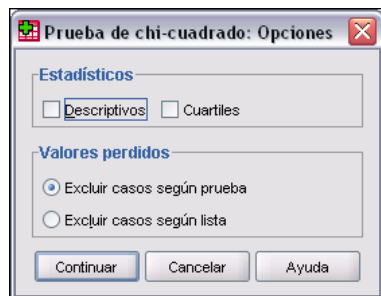
## Prueba de chi-cuadrado: Rango y valores esperados

**Rango esperado.** Por defecto, cada valor distinto de la variable se define como una categoría. Para establecer categorías dentro de un rango específico, seleccione Usarrango especificado e introduzca valores enteros para los límites inferior y superior. Se establecerán categorías para cada valor entero dentro del rango inclusivo y los casos con valores fuera de los límites se excluirán. Por ejemplo, si se especifica 1 como límite inferior y 4 como límite superior, únicamente se utilizarán los valores enteros entre 1 y 4 para la prueba de chi-cuadrado.

**Valores esperados.** Por defecto, todas las categorías tienen valores esperados iguales. Las categorías pueden tener proporciones esperadas especificadas por el usuario. Seleccione Valores, introduzca un valor mayor que 0 para cada categoría de la variable de contraste y, a continuación, pulse en Añadir. Cada vez que se añade un valor, éste aparece al final de la lista de valores. El orden de los valores es importante; corresponde al orden ascendente de los valores de categoría de la variable de contraste. El primer valor de la lista corresponde al valor de grupo mínimo de la variable de contraste y el último valor corresponde al valor máximo. Los elementos de la lista de valores se suman y, a continuación, cada valor se divide por esta suma para calcular la proporción de casos esperados en la categoría correspondiente. Por ejemplo, una lista de valores de 3, 4, 5, 4 especifica unas proporciones esperadas de 3/16, 4/16, 5/16 y 4/16.

## Prueba de chi-cuadrado: Opciones

Figura 36-2  
Cuadro de diálogo Prueba de chi-cuadrado: Opciones



**Estadísticos.** Puede elegir uno o los dos estadísticos de resumen.

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

**Valores perdidos.** Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable se excluyen de todos los análisis.



## **Funciones adicionales del comando NPAR TESTS (prueba de chi-cuadrado)**

Con el lenguaje de sintaxis de comandos también podrá:

- Especificar diferentes valores mínimo y máximo o las frecuencias esperadas para diferentes variables (mediante el subcomando `CHISQUARE`).
- Contrastar la misma variable respecto a las diferentes frecuencias esperadas, o bien usar diferentes rangos (mediante el subcomando `EXPECTED`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

## **Prueba binomial**

El procedimiento Prueba binomial compara las frecuencias observadas de las dos categorías de una variable dicotómica con las frecuencias esperadas en una distribución binomial con un parámetro de probabilidad especificado. Por defecto, el parámetro de probabilidad para ambos grupos es 0,5. Para cambiar las probabilidades, puede introducirse una proporción de prueba para el primer grupo. La probabilidad del segundo grupo será 1 menos la probabilidad especificada para el primer grupo.

**Ejemplo.** Si se lanza una moneda al aire, la probabilidad de que salga cara es 1/2. Basándose en esta hipótesis, se lanza una moneda al aire 40 veces y se anotan los resultados (cara o cruz). De la prueba binomial, podría deducir que en 3/4 de los lanzamientos salió cara y que el nivel crítico es pequeño (0,0027). Estos resultados indican que no es verosímil que la probabilidad de que salga cara sea 1/2; probablemente la moneda presenta una tendencia a caer por un sentido determinado.

**Estadísticos.** Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles.

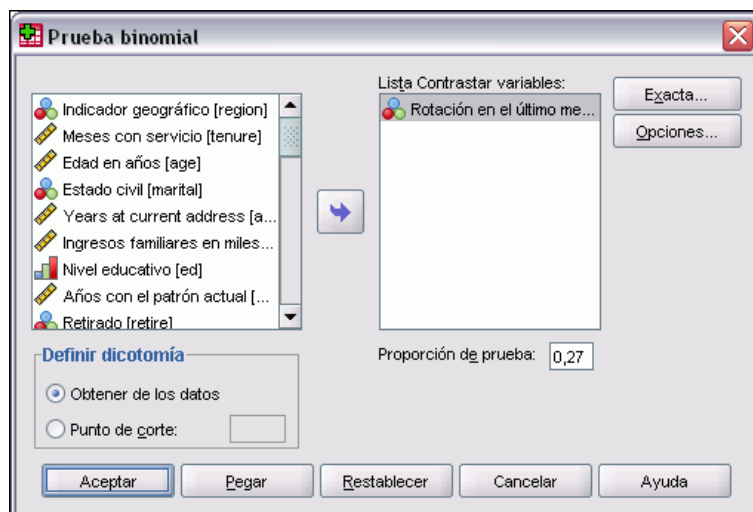
**Datos.** Las variables de contraste deben ser numéricas y dicotómicas. Para convertir las variables de cadena en variables numéricas, utilice el procedimiento Recodificación automática, disponible en el menú Transformar. Una **variable dicotómica** es una variable que sólo puede tomar dos valores posibles: *sí* o *no*, *verdadero* o *falso*, 0 ó 1, etc. Si las variables no son dicotómicas, debe especificar un punto de corte. El punto de corte asigna a un grupo los casos con valores mayores que el punto de corte y el resto de los casos a otro grupo.

**Supuestos.** Las pruebas no paramétricas no requieren supuestos sobre la forma de la distribución subyacente. Se asume que los datos son una muestra aleatoria.

### **Para obtener una prueba binomial**

- ▶ Elija en los menús:
  - Analizar
  - Pruebas no paramétricas
  - Binomial...

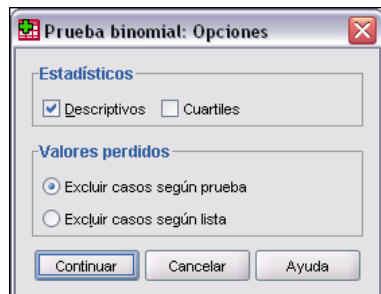
Figura 36-3  
Cuadro de diálogo Prueba binomial



- ▶ Seleccione una o más variables de contraste numéricas.
- ▶ Si lo desea, puede pulsar en Opciones para obtener estadísticos descriptivos, cuartiles y controlar el tratamiento de los datos perdidos.

### Prueba binomial: Opciones

Figura 36-4  
Cuadro de diálogo Prueba binomial: Opciones



**Estadísticos.** Puede elegir uno o los dos estadísticos de resumen.

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

**Valores perdidos.** Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Se excluirán de todos los análisis los casos con valores perdidos de cualquier variable.

## **Funciones adicionales del comando NPAR TESTS (Prueba binomial)**

Con el lenguaje de sintaxis de comandos también podrá:

- Seleccionar grupos específicos (y excluir otros grupos) si una variable tiene más de dos categorías (mediante el subcomando BINOMIAL).
- Especificar diferentes probabilidades o puntos de corte para diferentes variables (mediante el subcomando BINOMIAL).
- Contrastar la misma variable respecto a diferentes probabilidades o puntos de corte (mediante el subcomando EXPECTED).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

## **Prueba de rachas**

El procedimiento Prueba de rachas contrasta si es aleatorio el orden de aparición de dos valores de una variable. Una racha es una secuencia de observaciones similares. Una muestra con un número excesivamente grande o excesivamente pequeño de rachas sugiere que la muestra no es aleatoria.

**Ejemplos.** Suponga que se realiza una encuesta a 20 personas para saber si comprarían un producto. Si todas estas personas fueran del mismo sexo, se pondría seriamente en duda la supuesta aleatoriedad de la muestra. La prueba de rachas se puede utilizar para determinar si la muestra fue extraída de manera aleatoria.

**Estadísticos.** Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles.

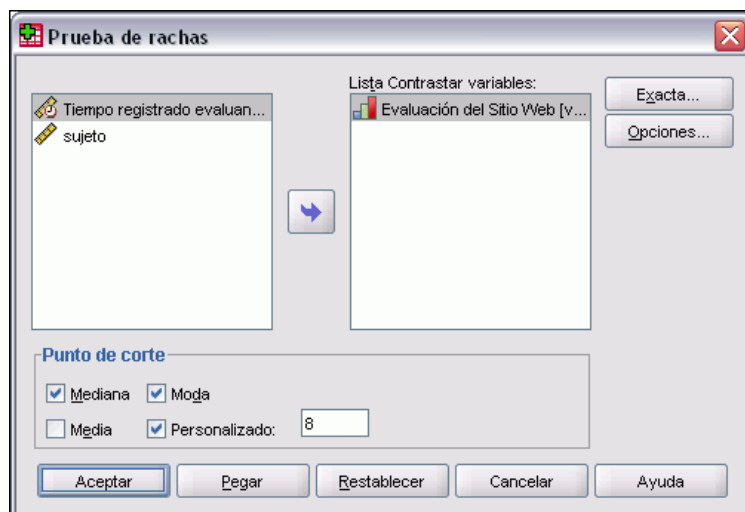
**Datos.** Las variables deben ser numéricas. Para convertir las variables de cadena en variables numéricas, utilice el procedimiento Recodificación automática, disponible en el menú Transformar.

**Supuestos.** Las pruebas no paramétricas no requieren supuestos sobre la forma de la distribución subyacente. Utilice muestras de distribuciones de probabilidad continua.

### **Para obtener una prueba de rachas**

- ▶ Elija en los menús:
  - Analizar
  - Pruebas no paramétricas
  - Rachas...

Figura 36-5  
Adición de un punto de corte personalizado



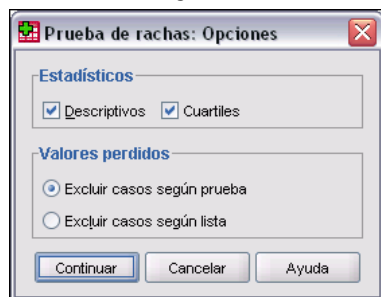
- ▶ Seleccione una o más variables de contraste numéricas.
- ▶ Si lo desea, puede pulsar en Opciones para obtener estadísticos descriptivos, cuartiles y controlar el tratamiento de los datos perdidos.

### ***Prueba de rachas: punto de corte***

**Punto de corte.** Especifica un punto de corte para dicotomizar las variables seleccionadas. Puede utilizar como punto de corte los valores observados para la media, la mediana o la moda, o bien un valor especificado. Los casos con valores menores que el punto de corte se asignarán a un grupo y los casos con valores mayores o iguales que el punto de corte se asignarán a otro grupo. Se lleva a cabo una prueba para cada punto de corte seleccionado.

### ***Prueba de rachas: Opciones***

Figura 36-6  
Cuadro de diálogo Prueba de rachas: Opciones



**Estadísticos.** Puede elegir uno o los dos estadísticos de resumen.

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

**Valores perdidos.** Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable se excluyen de todos los análisis.

### ***Funciones adicionales del comando NPAR TESTS (Prueba de rachas)***

Con el lenguaje de sintaxis de comandos también podrá:

- Especificar puntos de corte diferentes para las distintas variables (con el subcomando RUNS).
- Contrastar la misma variable con distintos puntos de corte personalizados (con el subcomando RUNS).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

### ***Prueba de Kolmogorov-Smirnov para una muestra***

El procedimiento Prueba de Kolmogorov-Smirnov para una muestra compara la función de distribución acumulada observada de una variable con una distribución teórica determinada, que puede ser la normal, la uniforme, la de Poisson o la exponencial. La Z de Kolmogorov-Smirnov se calcula a partir de la diferencia mayor (en valor absoluto) entre las funciones de distribución acumuladas teórica y observada. Esta prueba de bondad de ajuste contrasta si las observaciones podrían razonablemente proceder de la distribución especificada.

**Ejemplo.** Muchas pruebas paramétricas requieren que las variables se distribuyan de forma normal. La prueba de Kolmogorov-Smirnov para una muestra se puede utilizar para comprobar que una variable (por ejemplo *ingresos*) se distribuye normalmente.

**Estadísticos.** Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles.

**Datos.** Utilice variables cuantitativas (a nivel de medida de razón o de intervalo).

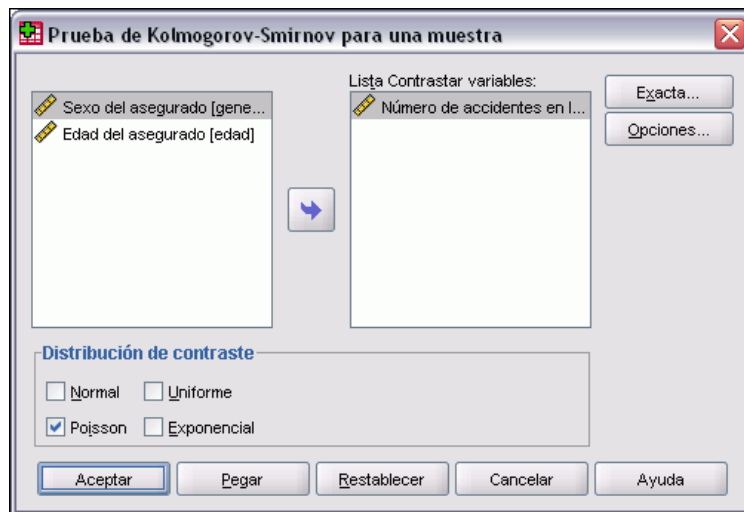
**Supuestos.** La prueba de Kolmogorov-Smirnov asume que los parámetros de la distribución de contraste se han especificado previamente. Este procedimiento estima los parámetros a partir de la muestra. La media y la desviación típica de la muestra son los parámetros de una distribución normal, los valores mínimo y máximo de la muestra definen el rango de la distribución uniforme, la media muestral es el parámetro de la distribución de Poisson y la media muestral es el parámetro de la distribución exponencial. La capacidad de la prueba para detectar desviaciones a partir de la distribución hipotetizada puede disminuir gravemente. Para contrastarla con una distribución normal con parámetros estimados, considere la posibilidad de utilizar la prueba de K-S Lilliefors (disponible en el procedimiento Explorar).

### **Para obtener una prueba de Kolmogorov-Smirnov para una muestra**

- ▶ Elija en los menús:
  - Analizar
  - Pruebas no paramétricas
  - K-S de 1 muestra...

Figura 36-7

Cuadro de diálogo Prueba de Kolmogorov-Smirnov para una muestra

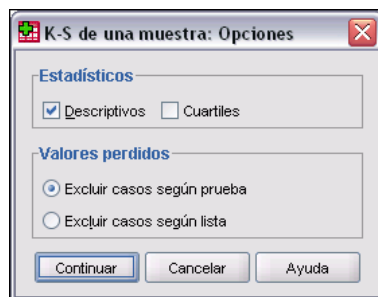


- ▶ Seleccione una o más variables de contraste numéricas. Cada variable genera una prueba independiente.
- ▶ Si lo desea, puede pulsar en Opciones para obtener estadísticos descriptivos, cuartiles y controlar el tratamiento de los datos perdidos.

### **Prueba de Kolmogorov-Smirnov para una muestra: Opciones**

Figura 36-8

Cuadro de diálogo Prueba de Kolmogorov-Smirnov para una muestra: Opciones



**Estadísticos.** Puede elegir uno o los dos estadísticos de resumen.

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

**Valores perdidos.** Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable se excluyen de todos los análisis.

### ***Funciones adicionales del comando NPAR TESTS (Prueba de Kolmogorov-Smirnov para una muestra)***

El lenguaje de sintaxis de comandos también permite especificar los parámetros de la distribución de contraste (con el subcomando  $K-S$ ).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

## ***Pruebas para dos muestras independientes***

El procedimiento Pruebas para dos muestras independientes compara dos grupos de casos existentes en una variable.

**Ejemplo.** Se han desarrollado nuevos correctores dentales diseñados para que sean más cómodos y estéticos, así como para facilitar un progreso más rápido en la realineación de la dentadura. Para averiguar si el nuevo corrector debe llevarse tanto tiempo como el modelo antiguo, se eligen 10 niños al azar para que lleven este último y otros 10 niños para que usen el nuevo. Mediante la prueba  $U$  de Mann-Whitney podría descubrir que, de media, los niños que llevaban el nuevo corrector tenían que llevarlo puesto menos tiempo que los que llevaban el antiguo.

**Estadísticos.** Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles. Pruebas:  $U$  de Mann-Whitney, reacciones extremas de Moses,  $Z$  de Kolmogorov-Smirnov, rachas de Wald-Wolfowitz.

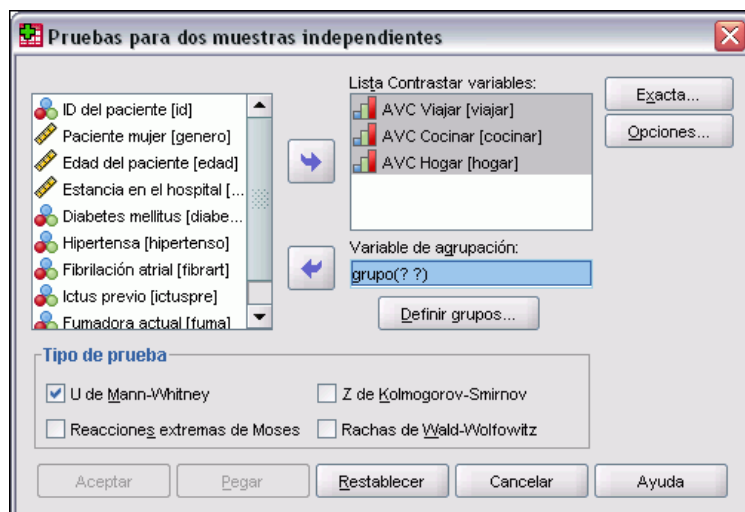
**Datos.** Utilice variables numéricas que puedan ser ordenables.

**Supuestos.** Utilice muestras independientes y aleatorias. La prueba  $U$  de Mann-Whitney requiere que las dos muestras probadas sean similares en la forma.

### ***Para obtener pruebas para dos muestras independientes***

- ▶ Elija en los menús:
  - Analizar
  - Pruebas no paramétricas
  - 2 muestras independientes...

Figura 36-9  
Cuadro de diálogo Pruebas para dos muestras independientes



- ▶ Seleccione una o más variables numéricas.
- ▶ Seleccione una variable de agrupación y pulse en Definir grupos para segmentar el archivo en dos grupos o muestras.

### ***Dos muestras independientes: Tipos de pruebas***

**Tipo de prueba.** Hay cuatro pruebas disponibles para contrastar si dos muestras (grupos) independientes proceden de una misma población.

La **prueba U de Mann-Whitney** es la más conocida de las pruebas para dos muestras independientes. Es equivalente a la prueba de la suma de rangos de Wilcoxon y a la prueba de Kruskal-Wallis para dos grupos. La prueba de Mann-Whitney contrasta si dos poblaciones muestreadas son equivalentes en su posición. Las observaciones de ambos grupos se combinan y clasifican, asignándose el rango promedio en caso de producirse empates. El número de empates debe ser pequeño en relación con el número total de observaciones. Si la posición de las poblaciones es idéntica, los rangos deberían mezclarse aleatoriamente entre las dos muestras. La prueba calcula el número de veces que una puntuación del grupo 1 precede a una puntuación del grupo 2 y el número de veces que una puntuación del grupo 2 precede a una puntuación del grupo 1. El estadístico  $U$  de Mann-Whitney es el menor de estos dos números. El estadístico  $W$  de la suma de rangos de Wilcoxon, que también se muestra, representa la suma más pequeña de los dos rangos. Si ambas muestras tienen el mismo número de observaciones,  $W$  es la suma de los rangos del grupo denominado primero en el cuadro de diálogo Dos muestras independientes: Definir grupos.

La **prueba Z de Kolmogorov-Smirnov** y la **prueba de rachas de Wald-Wolfowitz** son pruebas más generales que detectan las diferencias entre las posiciones y las formas de las distribuciones. La prueba de Kolmogorov-Smirnov se basa en la diferencia máxima absoluta entre las funciones de distribución acumulada observadas para ambas muestras. Cuando esta diferencia es significativamente grande, se consideran diferentes las dos distribuciones. La prueba de rachas de Wald-Wolfowitz combina y ordena las observaciones de ambos grupos. Si las dos

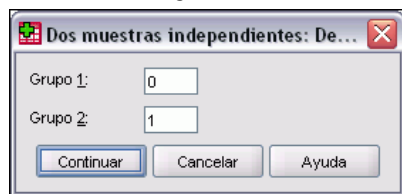


muestras proceden de una misma población, los dos grupos deben dispersarse aleatoriamente en la ordenación de los rangos.

La **prueba de reacciones extremas de Moses** presupone que la variable experimental afectará a algunos sujetos en una dirección y a otros sujetos en la dirección opuesta. La prueba contrasta las respuestas extremas comparándolas con un grupo de control. Esta prueba se centra en la amplitud del grupo de control y supone una medida de la influencia de los valores extremos del grupo experimental en la amplitud al combinarse con el grupo de control. El grupo de control se define en el cuadro Grupo 1 del cuadro de diálogo Dos muestras independientes: Definir grupos. Las observaciones de ambos grupos se combinan y ordenan. La amplitud del grupo de control se calcula como la diferencia entre los rangos de los valores mayor y menor del grupo de control más 1. Debido a que los valores atípicos ocasionales pueden distorsionar fácilmente el rango de la amplitud, de manera automática se recorta de cada extremo un 5% de los casos de control.

### Pruebas para dos muestras independientes: Definir grupos

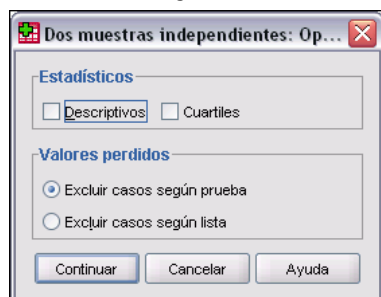
Figura 36-10  
Cuadro de diálogo Dos muestras independientes: Definir grupos



Para segmentar el archivo en dos grupos o muestras, introduzca un valor entero para el Grupo 1 y otro valor para el Grupo 2. Los casos con otros valores se excluyen del análisis.

### Pruebas para dos muestras independientes: Opciones

Figura 36-11  
Cuadro de diálogo Dos muestras independientes: Opciones



**Estadísticos.** Puede elegir uno o los dos estadísticos de resumen.

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

**Valores perdidos.** Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable se excluyen de todos los análisis.

### ***Funciones adicionales del comando NPAR TESTS (Dos muestras independientes)***

El lenguaje de sintaxis de comandos permite especificar el número de casos que se recortarán en la prueba de Moses (con el subcomando `MOSES`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

## ***Pruebas para dos muestras relacionadas***

El procedimiento Pruebas para dos muestras relacionadas compara las distribuciones de dos variables.

**Ejemplo.** En general, cuando una familia vende su casa ¿logra obtener la cantidad que pide inicialmente? Si aplica la prueba de Wilcoxon de los rangos con signo a 10 casas, podría descubrir que siete familias reciben menos cantidad de la solicitada, una recibe más y dos familias reciben el precio solicitado.

**Estadísticos.** Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles. Pruebas: Wilcoxon de los rangos con signo, signo, McNemar. Si se ha instalado la opción Pruebas exactas (disponible sólo en los sistemas operativos Windows), la prueba de homogeneidad marginal también estará disponible.

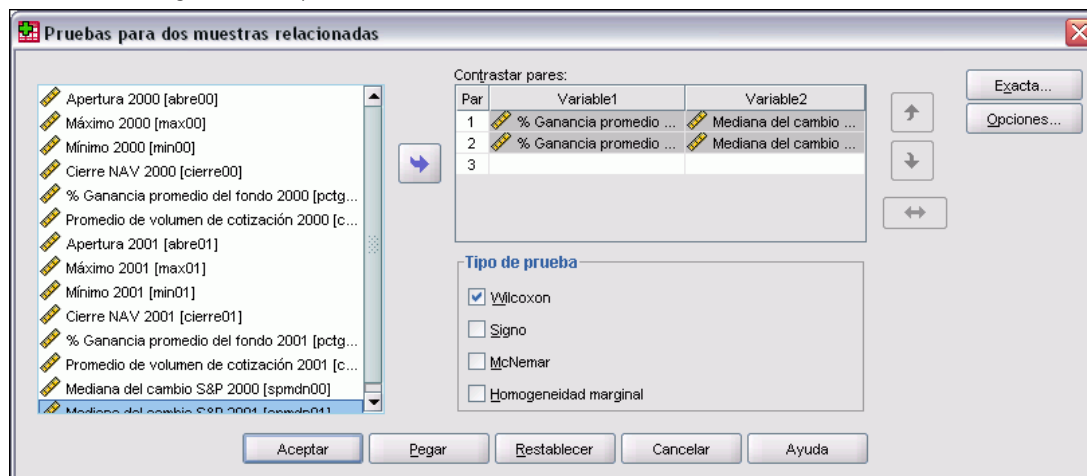
**Datos.** Utilice variables numéricas que puedan ser ordenables.

**Supuestos.** Aunque no se suponen distribuciones en particular para las dos variables, se supone que la distribución de población de las diferencias emparejadas es simétrica.

### ***Para obtener pruebas para dos muestras relacionadas***

- ▶ Elija en los menús:
  - Analizar
  - Pruebas no paramétricas
  - 2 muestras relacionadas...

Figura 36-12  
Cuadro de diálogo Pruebas para dos muestras relacionadas



- Seleccione uno o más pares de variables.

### ***Dos muestras relacionadas: Tipos de pruebas***

Las pruebas de esta sección comparan las distribuciones de dos variables relacionadas. La prueba apropiada depende del tipo de datos.

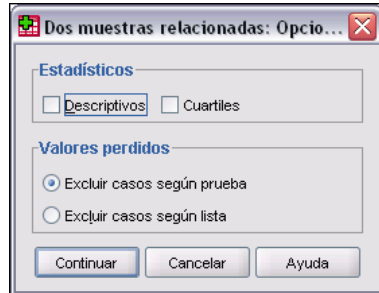
Si los datos son continuos, use la prueba de los signos o la prueba de Wilcoxon de los rangos con signo. La **prueba de los signos** calcula las diferencias entre las dos variables para todos los casos y clasifica las diferencias como positivas, negativas o empatadas. Si las dos variables tienen una distribución similar, el número de diferencias positivas y negativas no difiere de forma significativa. La **prueba de Wilcoxon de los rangos con signo** tiene en cuenta la información del signo de las diferencias y de la magnitud de las diferencias entre los pares. Dado que la prueba de Wilcoxon de los rangos con signo incorpora más información acerca de los datos, es más potente que la prueba de los signos.

Si los datos son binarios, use la **prueba de McNemar**. Esta prueba se utiliza normalmente en una situación de medidas repetidas, en la que la respuesta de cada sujeto se obtiene dos veces, una antes y otra después de que ocurra un evento especificado. La prueba de McNemar determina si la tasa de respuesta inicial (antes del evento) es igual a la tasa de respuesta final (después del evento). Esta prueba es útil para detectar cambios en las respuestas causadas por la intervención experimental en los diseños del tipo antes-después.

Si los datos son categóricos, use la **prueba de homogeneidad marginal**. Se trata de una extensión de la prueba de McNemar a partir de la respuesta binaria a la respuesta multinomial. Contrasta los cambios de respuesta, utilizando la distribución de chi-cuadrado, y es útil para detectar cambios de respuesta causados por intervención experimental en diseños antes-después. La prueba de homogeneidad marginal sólo está disponible si se ha instalado Pruebas exactas.

## Pruebas para dos muestras relacionadas: Opciones

Figura 36-13  
Cuadro de diálogo *Dos muestras relacionadas: Opciones*



**Estadísticos.** Puede elegir uno o los dos estadísticos de resumen.

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

**Valores perdidos.** Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable se excluyen de todos los análisis.

## Funciones adicionales del comando *NPAR TESTS (Dos muestras relacionadas)*

El lenguaje de sintaxis de comandos también permite contrastar una variable con cada variable de la lista.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

## Pruebas para varias muestras independientes

El procedimiento Pruebas para varias muestras independientes compara dos o más grupos de casos respecto a una variable.

**Ejemplo.** ¿Es diferente el tiempo medio en que se fundirán las bombillas de 100 vatios de tres marcas distintas? A partir del análisis de varianza de un factor de Kruskal-Wallis, puede comprobar que las tres marcas sí se diferencian en su vida media.

**Estadísticos.** Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles. Pruebas:  $H$  de Kruskal-Wallis, de la mediana.

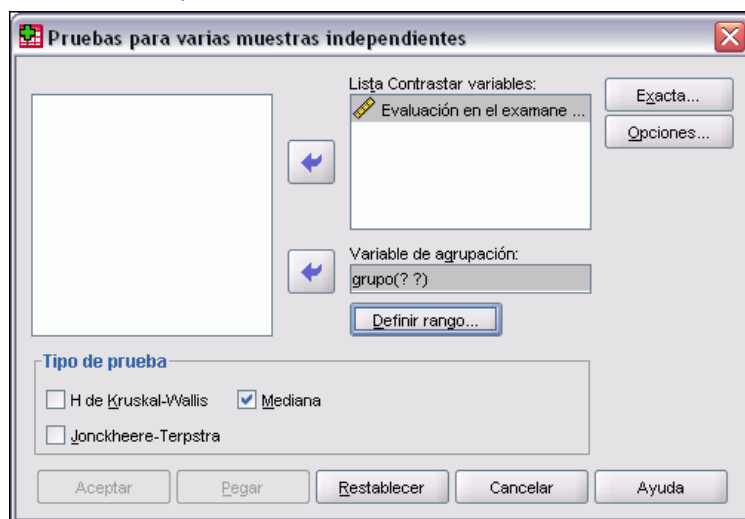
**Datos.** Utilice variables numéricas que puedan ser ordenables.

**Supuestos.** Utilice muestras independientes y aleatorias. La prueba  $H$  de Kruskal-Wallis requiere que las muestras comparadas tengan formas similares.

### Para obtener pruebas para varias muestras independientes

- ▶ Elija en los menús:
  - Analizar
  - Pruebas no paramétricas
  - K muestras independientes...

Figura 36-14  
Definición de la prueba de la mediana



- ▶ Seleccione una o más variables numéricas.
- ▶ Seleccione una variable de agrupación y pulse en Definir rango para especificar los valores enteros máximo y mínimo para la variable de agrupación.

### Tipos de pruebas para varias muestras independientes

Se hallan disponibles tres pruebas para determinar si varias muestras independientes proceden de la misma población. La prueba *H* de Kruskal-Wallis, la prueba de la mediana y la prueba de Jonckheere-Terpstra contrastan si varias muestras independientes proceden de la misma población.

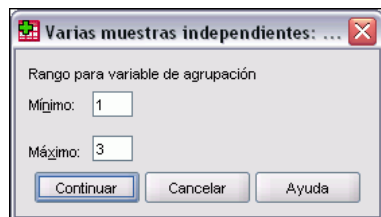
La prueba **H de Kruskal-Wallis**, una extensión de la prueba *U* de Mann-Whitney, es el análogo no paramétrico del análisis de varianza de un factor y detecta las diferencias en la localización de las distribuciones. La **prueba de la mediana**, que es una prueba más general pero no tan potente, detecta diferencias distribucionales en la localización y en la forma. La prueba *H* de Kruskal-Wallis y la prueba de la mediana suponen que no existe una ordenación *a priori* de las poblaciones *k* de las cuales se extraen las muestras.

Cuando *existe* una ordenación natural *a priori* (ascendente o descendente) de las poblaciones *k*, la prueba **Jonckheere-Terpstra** es más potente. Por ejemplo, las *k* poblaciones pueden representar *k* temperaturas ascendentes. Se contrasta la hipótesis de que diferentes temperaturas producen la misma distribución de respuesta, con la hipótesis alternativa de que cuando la temperatura aumenta, la magnitud de la respuesta aumenta. La hipótesis alternativa se encuentra aquí ordenada; por tanto, la prueba de Jonckheere-Terpstra es la prueba más apropiada. La prueba de Jonckheere-Terpstra estará disponible sólo si ha instalado el módulo adicional Pruebas exactas.

### Varias muestras independientes: Definir rango

Figura 36-15

Cuadro de diálogo Varias muestras independientes: Definir rango

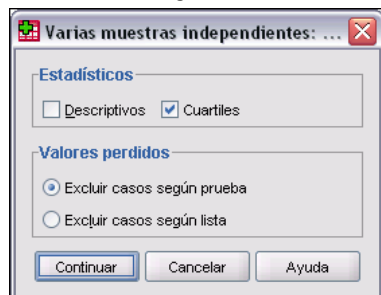


Para definir el rango, introduzca valores enteros para el mínimo y el máximo que se correspondan con las categorías mayor y menor de la variable de agrupación. Se excluyen los casos con valores fuera de los límites. Por ejemplo, si indica un valor mínimo de 1 y un valor máximo de 3, únicamente se utilizarán los valores enteros entre 1 y 3. Debe indicar ambos valores y el valor mínimo ha de ser menor que el máximo.

### Varias muestras independientes: Opciones

Figura 36-16

Cuadro de diálogo Varias muestras independientes: Opciones



**Estadísticos.** Puede elegir uno o los dos estadísticos de resumen.

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

**Valores perdidos.** Controla el tratamiento de los valores perdidos.

- **Excluir casos según prueba.** Cuando se especifican varias pruebas, cada una se evalúa separadamente respecto a los valores perdidos.
- **Excluir casos según lista.** Los casos con valores perdidos para cualquier variable se excluyen de todos los análisis.

### Funciones adicionales del comando NPAR TESTS (K muestras independientes)

El lenguaje de sintaxis de comandos permite especificar un valor distinto al de la mediana observada para la prueba de la mediana (mediante el subcomando MEDIAN).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

## Pruebas para varias muestras relacionadas

El procedimiento Pruebas para varias muestras relacionadas compara las distribuciones de dos o más variables.

**Ejemplo.** ¿Asocia la gente diferentes niveles de prestigio a doctores, abogados, policías y profesores? Se pide a diez personas que ordenen estas cuatro profesiones por orden de prestigio. La prueba de Friedman indica que la gente asocia diferentes niveles de prestigio con estas cuatro profesiones.

**Estadísticos.** Media, desviación típica, mínimo, máximo, número de casos no perdidos y cuartiles. Pruebas: Friedman,  $W$  de Kendall y  $Q$  de Cochran.

**Datos.** Utilice variables numéricas que puedan ser ordenables.

**Supuestos.** Las pruebas no paramétricas no requieren supuestos sobre la forma de la distribución subyacente. Utilice muestras aleatorias y dependientes.

### Para obtener pruebas para varias muestras relacionadas

- Elija en los menús:

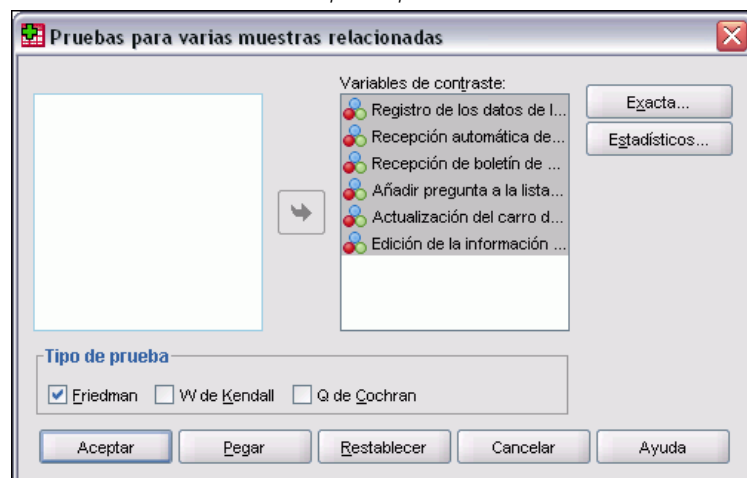
Analizar

Pruebas no paramétricas

K muestras relacionadas...

Figura 36-17

Selección de Cochran como tipo de prueba



- Seleccione dos o más variables de contraste numéricas.

### Tipos de prueba en el procedimiento Pruebas para varias muestras relacionadas

Hay tres pruebas disponibles para comparar las distribuciones de diversas variables relacionadas.

La **prueba de Friedman** es el equivalente no paramétrico de un diseño de medidas repetidas para una muestra o un análisis de varianza de dos factores con una observación por casilla. Friedman contrasta la hipótesis nula de que las  $k$  variables relacionadas procedan de la misma población. En cada caso, las  $k$  variables se ordenan en un rango de 1 a  $k$ . El estadístico de contraste se basa en estos rangos.

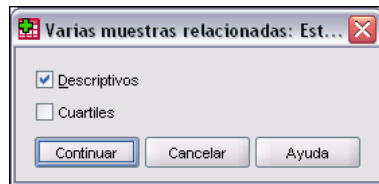
La **W de Kendall** es una normalización del estadístico de Friedman. La prueba  $W$  de Kendall se puede interpretar como el coeficiente de concordancia, que es una medida de acuerdo entre evaluadores. Cada caso es un juez o evaluador y cada variable es un elemento o persona que está siendo evaluada. Para cada variable, se calcula la suma de rangos. La  $W$  de Kendall varía entre 0 (no hay acuerdo) y 1 (acuerdo completo).

La prueba **Q de Cochran** es idéntica a la prueba de Friedman pero se puede aplicar cuando todas las respuestas son binarias. Esta prueba es una extensión de la prueba de McNemar para la situación de  $k$  muestras. La  $Q$  de Cochran contrasta la hipótesis de que diversas variables dicotómicas relacionadas tienen la misma media. Las variables se miden al mismo individuo o a individuos emparejados.

### **Pruebas para varias muestras relacionadas: Estadísticos**

Figura 36-18

Cuadro de diálogo Varias pruebas relacionadas: Estadísticos



Puede elegir estadísticos.

- **Descriptivos.** Muestra la media, la desviación típica, el mínimo, el máximo y el número de casos no perdidos.
- **Cuartiles.** Muestra los valores correspondientes a los percentiles 25, 50 y 75.

### **Funciones adicionales del comando NPAR TESTS (K pruebas relacionadas)**

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).



## ***Análisis de respuestas múltiples***

Se ofrecen dos procedimientos para analizar los conjuntos de categorías múltiples y de dicotomías múltiples. El procedimiento Frecuencias de respuestas múltiples muestra tablas de frecuencias. El procedimiento Tablas de contingencia de respuestas múltiples muestra tablas de contingencia de dos y tres dimensiones. Antes de utilizar cualquiera de estos procedimientos, deberá definir conjuntos de respuestas múltiples.

**Ejemplo.** Este ejemplo ilustra el uso de elementos de respuestas múltiples en un estudio de investigación de mercado. Los datos son ficticios y no deben interpretarse como reales. Una línea aérea podría hacer una encuesta a los pasajeros que realicen una determinada ruta para evaluar las líneas aéreas de la competencia. En este ejemplo, American Airlines desea conocer el uso que hacen sus pasajeros de otras líneas aéreas en la ruta Chicago-Nueva York y la importancia relativa del horario y el servicio a la hora de seleccionar una línea aérea. El encargado del vuelo proporciona a cada pasajero un breve cuestionario durante el embarque. La primera pregunta dice: Rodee con un círculo todas las líneas aéreas con la que haya volado al menos una vez en los últimos seis meses en este mismo trayecto: American, United, TWA, USAir, Otras. Se trata de una pregunta de respuestas múltiples, ya que el pasajero puede marcar más de una respuesta. Sin embargo, la pregunta no se puede codificar directamente, ya que una variable sólo puede tener un valor para cada caso. Deberá utilizar distintas variables para asignar respuestas a cada pregunta. Existen dos formas de hacerlo. Una consiste en definir una variable para cada una de las opciones (por ejemplo, American, United, TWA, USAir y Otras). Si el pasajero marca United, a la variable *united* se le asignará el código 1; en caso contrario se le asignará 0. Éste es un **método de dicotomías múltiples** de asignación de variables. La otra forma de asignar respuestas es el **método de categorías múltiples**, en el que se estima el número máximo de posibles respuestas a la pregunta y se configura el mismo número de variables, con códigos para especificar la línea aérea utilizada. Examinando una muestra de cuestionarios, podría observarse que ningún usuario ha volado en más de tres líneas aéreas diferentes en esta ruta durante los últimos seis meses. Aún más, se observará que debido a la liberalización de las líneas aéreas, aparecen otras 10 en la categoría Otras. Con el método de respuestas múltiples, definiría tres variables, cada una codificada como 1 = *american*, 2 = *united*, 3 = *twa*, 4 = *usair*, 5 = *delta* y así sucesivamente. Si un pasajero determinado marca American y TWA, la primera variable tendrá el código 1, la segunda el 3 y la tercera un código de valor perdido. Otro pasajero podría haber marcado American e introducido Delta. Así, la primera variable tendrá el código 1, la segunda el 5 y la tercera un código de valor perdido. Por el contrario, si utiliza el método de dicotomías múltiples, terminará con 14 variables independientes. Aunque cualquiera de los métodos de asignación anteriores es viable para este estudio, el método seleccionado dependerá de la distribución de respuestas.

## **Definir conjuntos de respuestas múltiples**

El procedimiento Definir conjuntos de respuestas múltiples agrupa variables elementales en conjuntos de categorías múltiples y de dicotomías múltiples, para los que se pueden obtener tablas de frecuencias y tablas de contingencia. Se pueden definir hasta 20 conjuntos de respuestas múltiples. Cada conjunto debe tener un nombre exclusivo. Para eliminar un conjunto, resáltelo en la lista de conjuntos de respuestas múltiples y pulse en Borrar. Para cambiar un conjunto, resáltelo en la lista, modifique cualquier característica de la definición del conjunto y pulse en Cambiar.

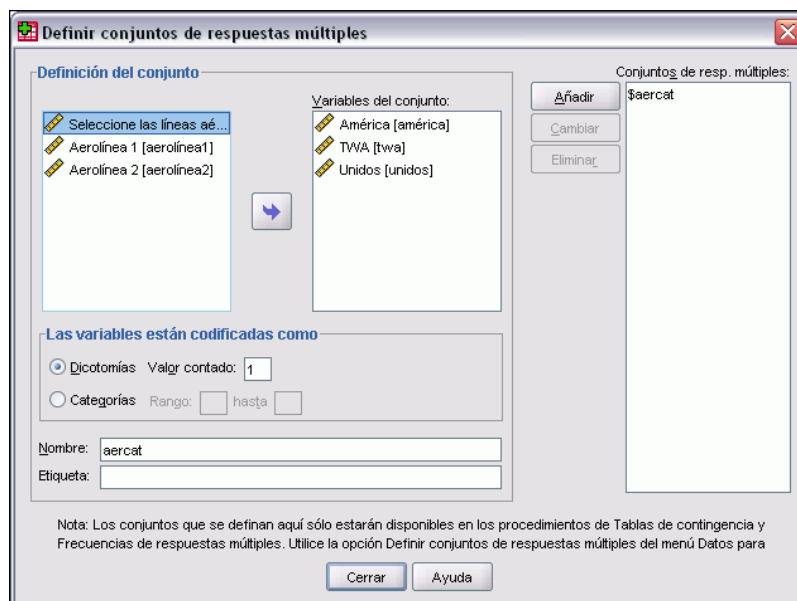
Las variables elementales se pueden codificar como dicotomías o categorías. Para utilizar variables dicotómicas, seleccione Dicotomías para crear un conjunto de dicotomías múltiples. Introduzca un valor entero en Valor contado. Cada variable que tenga al menos una aparición del valor contado se convierte en una categoría del conjunto de dicotomías múltiples. Seleccione Categorías para crear un conjunto de categorías múltiples con el mismo rango de valores que las variables que lo componen. Introduzca valores enteros para los valores máximo y mínimo del rango para las categorías del conjunto de categorías múltiples. El procedimiento suma cada valor entero distinto en el rango inclusivo para todas las variables que lo componen. Las categorías vacías no se tabulan.

A cada conjunto de respuestas múltiples se le debe asignar un nombre exclusivo de hasta siete caracteres. El procedimiento coloca delante del nombre asignado un signo dólar (\$). No se pueden utilizar los siguientes nombres reservados: *casenum*, *sysmis*, *jdate*, *date*, *time*, *length* y *width*. El nombre del conjunto de respuestas múltiples sólo se encuentra disponible para su uso en los procedimientos de respuestas múltiples. No se puede hacer referencia a nombres de conjuntos de respuestas múltiples en otros procedimientos. Si lo desea, puede introducir una etiqueta de variable descriptiva para el conjunto de respuestas múltiples. La etiqueta puede tener hasta 40 caracteres.

### **Para definir conjuntos de respuestas múltiples**

- ▶ Elija en los menús:
  - Analizar
  - Respuesta múltiple
  - Definir conjuntos...

Figura 37-1  
Cuadro de diálogo Definir conjuntos de respuestas múltiples



- ▶ Seleccione dos o más variables.
- ▶ Si las variables están codificadas como dicotomías, indique qué valor desea contar. Si las variables están codificadas como categorías, defina el rango de las categorías.
- ▶ Escriba un nombre distintivo para cada conjunto de respuestas múltiples.
- ▶ Pulse Añadir para añadir el conjunto de respuestas múltiples a la lista de conjuntos definidos.

## ***Frecuencias de respuestas múltiples***

El procedimiento Frecuencias de respuestas múltiples produce tablas de frecuencias para conjuntos de respuestas múltiples. En primer lugar es necesario definir uno o más conjuntos de respuestas múltiples (véase “Definir conjuntos de respuestas múltiples”).

Para los conjuntos de dicotomías múltiples, los nombres de categorías que se muestran en los resultados proceden de etiquetas de variable definidas para variables elementales del grupo. Si las etiquetas de variable no están definidas, los nombres de las variables se utilizarán como etiquetas. Para los conjuntos de categorías múltiples, las etiquetas de categoría proceden de las etiquetas de valor de la primera variable del grupo. Si las categorías perdidas para la primera variable están presentes para otras variables del grupo, defina una etiqueta de valor para las categorías perdidas.

**Valores perdidos.** Los casos con valores perdidos se excluyen en base a tabla por tabla. Si lo desea, puede seleccionar una de las opciones siguientes o ambas:

- **Excluir los casos según lista dentro de las dicotomías.** Excluye los casos con valores perdidos en cualquier variable de la tabulación del conjunto de dicotomías múltiples. Esto sólo se aplica a conjuntos de respuestas múltiples definidos como conjuntos de dicotomías. Por defecto, un caso se considera perdido para un conjunto de dicotomías múltiples si ninguna de sus variables que lo componen contiene el valor contado. Los casos con valores perdidos en

algunas variables, pero no en todas, se incluyen en las tabulaciones del grupo si al menos una variable contiene el valor contado.

- **Excluir los casos según lista dentro de las categorías.** Excluye los casos con valores perdidos en cualquier variable de la tabulación del conjunto de categorías múltiples. Esto sólo se aplica a conjuntos de respuestas múltiples definidos como conjuntos de categorías. Por defecto, un caso se considera perdido para un conjunto de categorías múltiples sólo si ninguno de sus componentes tiene valores válidos dentro del rango definido.

**Ejemplo.** Cada variable creada a partir de una pregunta de una encuesta es una variable elemental. Para analizar un elemento de respuestas múltiples, deberá combinar las variables en uno o dos tipos de conjuntos de respuestas múltiples: un conjunto de dicotomías múltiples o un conjunto de categorías múltiples. Por ejemplo, si una encuesta sobre líneas aéreas preguntara al encuestado cuál de las tres líneas (American, United, TWA) ha utilizado durante los seis últimos meses y usted utilizara variables dicotómicas y definiera un **conjunto de dicotomías múltiples**, cada una de las tres variables del conjunto se convertiría en una categoría de la variable de grupo. Las frecuencias y los porcentajes de las tres líneas aéreas se muestran en una tabla de frecuencias. Si observa que ningún encuestado ha mencionado más de dos líneas aéreas, podría crear dos variables, cada una con tres códigos, uno para cada línea aérea. Si define un **conjunto de categorías múltiples**, los valores se tabulan añadiendo los mismos códigos en las variables elementales juntas. El conjunto de valores resultantes es igual a los de cada una de las variables elementales. Por ejemplo, 30 respuestas para United son la suma de las cinco respuestas de United para la línea aérea 1 y las 25 respuestas de United para la línea aérea 2. Las frecuencias y los porcentajes de las tres líneas aéreas se muestran en una tabla de frecuencias.

**Estadísticos.** Tablas de frecuencias que muestran recuentos, porcentajes de respuestas, porcentajes de casos, número de casos válidos y número de casos perdidos.

**Datos.** Utilice conjuntos de respuestas múltiples.

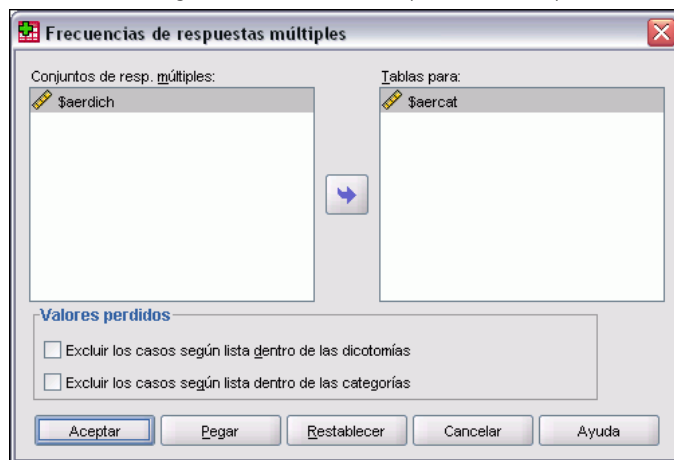
**Supuestos.** Las frecuencias y los porcentajes proporcionan una descripción útil de los datos de cualquier distribución.

**Procedimientos relacionados.** El procedimiento Definir conjuntos de respuestas múltiples permite definir este tipo de conjuntos.

### ***Para obtener frecuencias de respuestas múltiples***

- ▶ Elija en los menús:
  - Analizar
  - Respuesta múltiple
  - Frecuencias...

Figura 37-2  
Cuadro de diálogo Frecuencias de respuestas múltiples



- Seleccione uno o más conjuntos de respuestas múltiples.

## Tablas de contingencia de respuestas múltiples

El procedimiento Tablas de contingencia de respuestas múltiples presenta en forma de tabla de contingencia conjuntos de respuestas múltiples, variables elementales o una combinación. También puede obtener porcentajes de casilla basados en casos o respuestas, modificar la gestión de los valores perdidos u obtener tablas de contingencia emparejadas. Antes debe definir uno o varios conjuntos de respuestas múltiples (véase “Para definir conjuntos de respuestas múltiples”).

Para los conjuntos de dicotomías múltiples, los nombres de categorías que se muestran en los resultados proceden de etiquetas de variable definidas para variables elementales del grupo. Si las etiquetas de variable no están definidas, los nombres de las variables se utilizarán como etiquetas. Para los conjuntos de categorías múltiples, las etiquetas de categoría proceden de las etiquetas de valor de la primera variable del grupo. Si las categorías perdidas para la primera variable están presentes para otras variables del grupo, defina una etiqueta de valor para las categorías perdidas. El procedimiento muestra las etiquetas de categoría por columnas en tres líneas, con un máximo de ocho caracteres por línea. Para evitar la división de palabras, puede invertir los elementos de las filas y las columnas o volver a definir las etiquetas.

**Ejemplo.** Tanto los conjuntos de categorías múltiples como los conjuntos de dicotomías múltiples se pueden presentar en forma de tabla de contingencia con otras variables de este procedimiento. Un estudio sobre pasajeros de líneas aéreas solicita a éstos la siguiente información: Marque las líneas aéreas con las que ha volado al menos una vez en los seis últimos meses (American, United, TWA). ¿Qué considera más importante a la hora de seleccionar un vuelo, el horario o el servicio? Seleccione sólo uno. Después de introducir los datos como dicotomías o categorías múltiples y combinarlos en un conjunto, puede presentar en forma de tabla de contingencia las selecciones de línea aérea con la pregunta relativa al servicio o al horario.

**Estadísticos.** Tabla de contingencia con recuentos de casilla, fila, columna y totales, así como porcentajes de casilla, fila, columna y totales. Los porcentajes de casilla se basan en casos o respuestas.

**Datos.** Utilice conjuntos de respuestas múltiples o variables categóricas numéricas.

**Supuestos.** Las frecuencias y los porcentajes proporcionan una útil descripción de los datos de cualquier distribución.

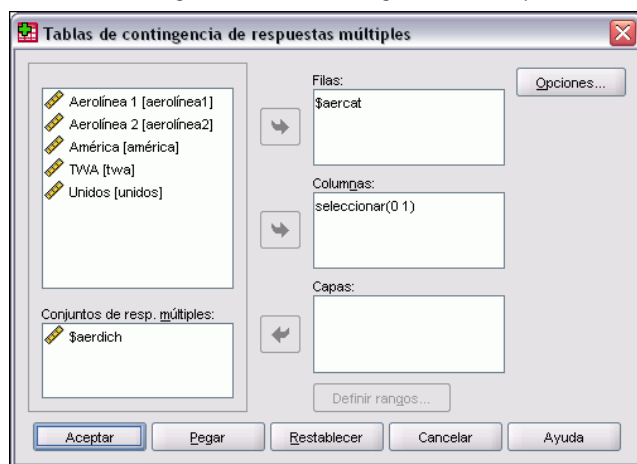
**Procedimientos relacionados.** El procedimiento Definir conjuntos de respuestas múltiples permite definir este tipo de conjuntos.

### **Para obtener tablas de contingencia de respuestas múltiples**

- ▶ Elija en los menús:
  - Analizar
  - Respuesta múltiple
  - Tablas de contingencia...

Figura 37-3

Cuadro de diálogo Tablas de contingencia de respuestas múltiples



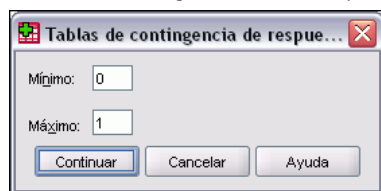
- ▶ Seleccione una o más variables numéricas o conjuntos de respuestas múltiples para cada dimensión de la tabla de contingencia.
- ▶ Defina el rango de cada variable elemental.

Si lo desea, puede obtener una tabla de contingencia de doble clasificación para cada categoría de una variable de control o conjunto de respuestas múltiples. Seleccione uno o varios elementos para la lista Capas.

## **Tablas de respuestas múltiples: Definir rangos de las variables**

Figura 37-4

Cuadro de diálogo Tablas de respuestas múltiples: Definir rangos de las variables

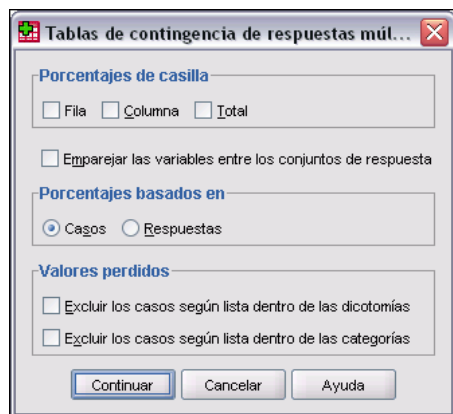


Los rangos de valores deben definirse para cualquier variable elemental de la tabla de contingencia. Introduzca los valores enteros de categoría máximos y mínimos que desee tabular. Las categorías que estén fuera del rango se excluyen del análisis. Se entiende que los valores que estén dentro del rango inclusivo son enteros (los no enteros quedan truncados).

## Tablas de contingencia de respuestas múltiples: Opciones

Figura 37-5

Cuadro de diálogo Tablas de contingencia de respuestas múltiples: Opciones



**Porcentajes de casilla.** Las frecuencias de la casilla siempre se muestran. Puede elegir entre mostrar los porcentajes de fila, los de columna o los de tabla de doble clasificación (totales).

**Porcentajes basados en.** Los porcentajes de casilla pueden basarse en casos (o encuestados). Esta opción no estará disponible si selecciona la concordancia de variables en conjuntos de categorías múltiples. También se pueden basar en las respuestas. Para los conjuntos de dicotomías múltiples, el número de respuestas es igual al número de valores contados por los casos. Para los conjuntos de categorías múltiples, el número de respuestas es el número de valores del rango definido.

**Valores perdidos.** Puede elegir una o ambas de las siguientes opciones:

- **Excluir los casos según lista dentro de las dicotomías.** Excluye los casos con valores perdidos en cualquier variable de la tabulación del conjunto de dicotomías múltiples. Esto sólo se aplica a conjuntos de respuestas múltiples definidos como conjuntos de dicotomías. Por defecto, un caso se considera perdido para un conjunto de dicotomías múltiples si ninguna de sus variables que lo componen contiene el valor contado. Los casos con valores perdidos para algunas variables, pero no todas, se incluyen en las tabulaciones del grupo si al menos una variable contiene el valor contado.
- **Excluir los casos según lista dentro de las categorías.** Excluye los casos con valores perdidos en cualquier variable de la tabulación del conjunto de categorías múltiples. Esto sólo se aplica a conjuntos de respuestas múltiples definidos como conjuntos de categorías. Por defecto, un caso se considera perdido para un conjunto de categorías múltiples sólo si ninguno de sus componentes tiene valores válidos dentro del rango definido.

Por defecto, cuando se presentan dos conjuntos de categorías múltiples en forma de tabla de contingencia, el procedimiento tabula cada variable del primer grupo con cada variable del segundo y suma las frecuencias de cada casilla; de esta forma, algunas respuestas pueden aparecer más de una vez en una tabla. Puede seleccionar la opción siguiente:

**Emparejar las variables entre los conjuntos de respuesta.** Empareja la primera variable del primer grupo con la primera variable del segundo, y así sucesivamente. Si selecciona esta opción, el procedimiento basará los porcentajes de casilla en las respuestas en lugar de hacerlo en los encuestados. El emparejamiento no está disponible para conjuntos de dicotomías múltiples o variables elementales.

## ***Funciones adicionales del comando MULT RESPONSE***

Con el lenguaje de sintaxis de comandos también podrá:

- Obtener tablas de contingencia de hasta cinco dimensiones (con el subcomando `BY`).
- Cambiar las opciones de formato de los resultados, incluyendo la supresión de etiquetas de valor (con el subcomando `FORMAT`).

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).



## ***Informes de los resultados***

Los listados de casos y los estadísticos descriptivos son herramientas básicas para estudiar y presentar los datos. Puede obtener listados de casos con el Editor de datos o el procedimiento Resumir, frecuencias y estadísticos descriptivos con el procedimiento Frecuencias, y estadísticos de subpoblación con el procedimiento Medias. Cada uno utiliza un formato diseñado para que la información sea clara. Si desea ver la información con otro formato, las opciones Informe de estadísticos en filas e Informe de estadísticos en columnas le ofrecen el control que precisa para presentar los datos.

### ***Informe de estadísticos en filas***

Informe de estadísticos en filas genera informes en los cuales se presentan distintos estadísticos de resumen en filas. También se encuentran disponibles listados de los casos, con o sin estadísticos de resumen.

**Ejemplo.** Una empresa con una cadena de tiendas registra los datos de sus empleados, incluyendo el salario, el cargo, la tienda y la sección en la que trabaja cada uno. Se podría generar un informe que proporcione los datos individuales de cada empleado (listado) desglosados por tienda y sección (variables de ruptura), con estadísticos de resumen (por ejemplo, el salario medio) por tienda, sección y sección dentro de cada tienda.

**Columnas de datos.** Muestra una lista de las variables del informe para las que desea obtener el listado de los casos o los estadísticos de resumen y controla el formato de presentación de las columnas de datos.

**Romper columnas por.** Muestra una lista de las variables de ruptura opcionales que dividen el informe en grupos y controla los estadísticos de resumen y los formatos de presentación de las columnas de ruptura. Si hay varias variables de ruptura, se creará un grupo distinto para cada una de las categorías de las variables de ruptura dentro de las categorías de la variable de ruptura anterior en la lista. Las variables de ruptura deben ser variables categóricas discretas que dividan los casos en un número limitado de categorías con sentido. Los valores individuales de cada variable de ruptura aparecen ordenados en una columna distinta situada a la izquierda de todas las columnas de datos.

**Informe.** Controla las características globales del informe, incluyendo los estadísticos de resumen globales, la presentación de los valores perdidos, la numeración de las páginas y los títulos.

**Mostrar casos.** Muestra los valores reales (o etiquetas de valor) de las variables de la columna de datos para cada caso. Esto genera un informe a modo de listado, que puede ser mucho más largo que un informe de resumen.

**Presentación preliminar.** Muestra sólo la primera página del informe. Esta opción es útil para ver una presentación preliminar del formato del informe sin tener que procesar el informe completo.

**Los datos están ordenados.** Para los informes con variables de ruptura, el archivo de datos se debe ordenar por los valores de estas variables antes de generar el informe. Si el archivo de datos ya está ordenado por estos valores, se puede ahorrar tiempo de procesamiento seleccionando esta opción. Esta opción es especialmente útil después de generar la presentación preliminar de un informe.

### **Para obtener un informe de resumen: Estadísticos en filas**

- ▶ Elija en los menús:  
   Analizar  
    Informes  
     Informe de estadísticos en filas...
- ▶ Seleccione una o más variables para las columnas de datos. En el informe se genera una columna para cada variable seleccionada.
- ▶ Para los informes ordenados y mostrados por subgrupos, seleccione una o más variables para Romper columnas por.
- ▶ Para los informes con estadísticos de resumen para los subgrupos definidos por las variables de ruptura, seleccione la variable de ruptura de la lista Romper columnas por y pulse en Resumen, en la sección Romper columnas por, para especificar las medidas de resumen.
- ▶ Para los informes con estadísticos de resumen globales, pulse en Resumen para especificar las medidas de resumen.

Figura 38-1  
 Cuadro de diálogo Informe: Estadísticos en filas



## Formato de las columnas de datos y de la ruptura de columnas del informe

Los cuadros de diálogo de formato controlan los títulos y el ancho de las columnas, la alineación del texto y la presentación de los valores de los datos o de las etiquetas de valor. El formato de las columnas de datos controla el formato de las columnas de datos situadas en la parte derecha de la página del informe. La opción Formato de ruptura controla el formato de las columnas de ruptura situadas en la parte izquierda.

Figura 38-2  
Cuadro de diálogo Informe: Formato de columna de datos

**Título de la columna.** Para la variable seleccionada, controla el título de la columna. Los títulos largos se ajustan de forma automática dentro de la columna. Utilice la tecla Intro para insertar manualmente líneas de separación donde desee ajustar los títulos.

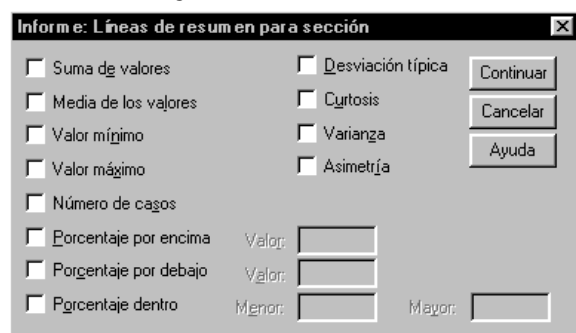
**Posición de valor en la columna.** Para la variable seleccionada, controla la alineación de los valores de los datos o de las etiquetas de valor dentro de la columna. La alineación de los valores o de las etiquetas no afecta a la alineación de los encabezados de las columnas. Puede sangrar el contenido de la columna por un número específico de caracteres o centrar el contenido.

**Contenido de la columna.** Para la variable seleccionada, controla la presentación de los valores de los datos o de las etiquetas de valor definidas. Los valores de los datos siempre se muestran para cualquier valor que no tenga etiquetas de valor definidas. No se encuentra disponible para las columnas de datos en los informes estadísticos en columnas.

## Líneas de resumen finales y Líneas de resumen del informe

Los dos cuadros de diálogo Líneas de resumen controlan la presentación de los estadísticos de resumen para los grupos de ruptura y para el informe entero. Líneas de resumen controla los estadísticos de subgrupo para cada categoría definida por las variables de ruptura. Líneas de resumen finales controla los estadísticos globales que se muestran al final del informe.

Figura 38-3  
Cuadro de diálogo Informe: Líneas de resumen

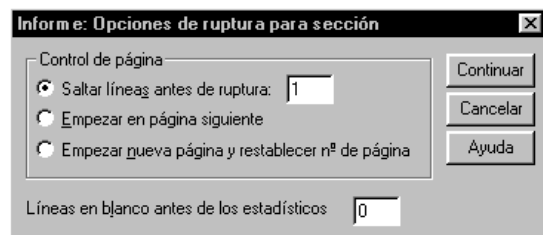


Los estadísticos de resumen disponibles son: suma, media, valor mínimo, valor máximo, número de casos, porcentaje de casos por encima y por debajo de un valor especificado, porcentaje de casos dentro de un rango de valores especificado, desviación típica, curtosis, varianza y asimetría.

### Opciones de ruptura del informe

Opciones de ruptura controla el espaciado y la paginación de la información de la categoría de ruptura.

Figura 38-4  
Cuadro de diálogo Informe: Opciones de ruptura



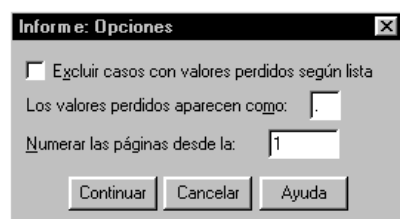
**Control de página.** Controla el espaciado y la paginación para las categorías de la variable de ruptura seleccionada. Puede especificar un número de líneas en blanco entre las categorías de ruptura o empezar cada categoría de ruptura en una página nueva.

**Líneas en blanco antes de los estadísticos.** Controla el número de líneas en blanco entre las etiquetas o los datos de la categoría de ruptura y los estadísticos de resumen. Esta opción es especialmente útil para los informes combinados que incluyan tanto el listado de los casos individuales como los estadísticos de resumen para las categorías de ruptura; en estos informes puede insertar un espacio entre el listado de los casos y los estadísticos de resumen.

### Opciones del informe

Informe: Opciones controla el tratamiento y la presentación de los valores perdidos y la numeración de las páginas del informe.

Figura 38-5  
Cuadro de diálogo Informe: Opciones



**Excluir casos con valores perdidos según lista.** Elimina (del informe) cualquier caso con valores perdidos para cualquier variable del informe.

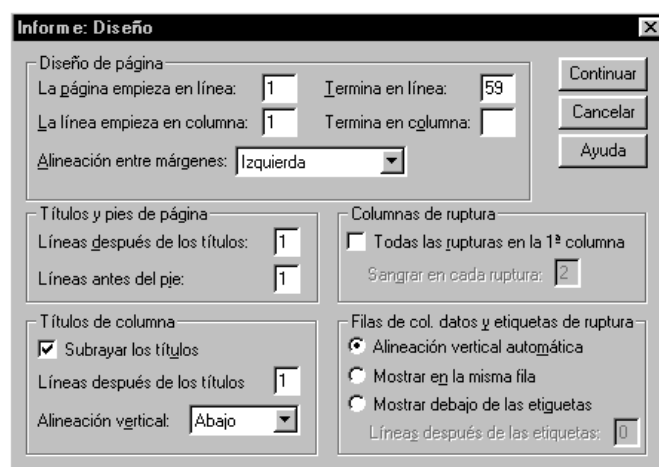
**Los valores perdidos aparecen como.** Permite especificar el símbolo que representa los valores perdidos en el archivo de datos. Este símbolo sólo puede tener un carácter y se utiliza para representar tanto los valores **perdidos por el sistema** y **valores perdidos definidos por el usuario**.

**Numerar las páginas desde la.** Permite especificar un número de página para la primera página del informe.

## Diseño del informe

Informe: Diseño controla el ancho y alto de cada página del informe, la ubicación del informe dentro de la página y la inserción de etiquetas y líneas en blanco.

Figura 38-6  
Cuadro de diálogo Informe: Diseño



**Diseño de página.** Controla los márgenes de las páginas expresados en líneas (extremos superior e inferior) y caracteres (a la izquierda y a la derecha) y la alineación del informe entre los márgenes.

**Títulos y pies de página.** Controla el número de líneas que separan los títulos y los pies de página del cuerpo del informe.

**Romper columnas por.** Controla la presentación de las columnas de ruptura. Si se especifican diversas variables de ruptura, pueden situarse en columnas diferentes o en la primera columna. Si se colocan todas en la primera columna, se generará un informe más estrecho.

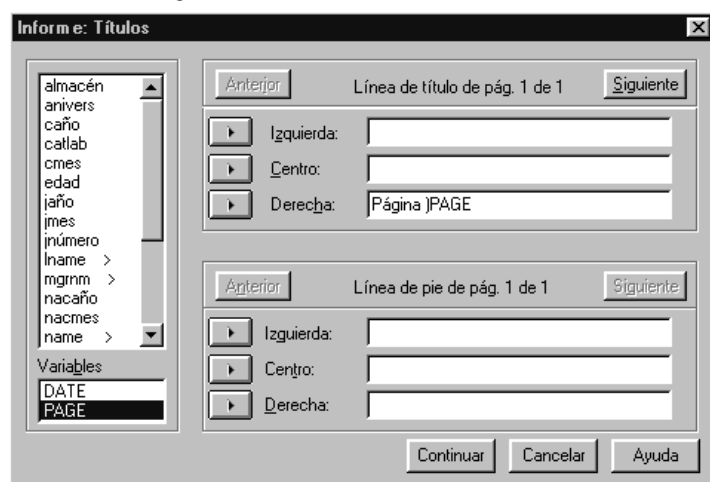
**Títulos de columna.** Controla la presentación de los títulos de columna, incluyendo el subrayado de títulos, el espacio entre los títulos y el cuerpo del informe y la alineación vertical de los títulos de columna.

**Filas de col. datos y etiquetas de ruptura.** Controla la ubicación de la información de las columnas de datos (valores de datos o estadísticos de resumen) en relación con las etiquetas de ruptura al principio de cada categoría de ruptura. La primera fila de información puede empezar en la misma línea que la etiqueta de categoría de ruptura o en un número de líneas posterior especificado. Esta sección no se encuentra disponible para los informes de estadísticos en columnas.

## Títulos del informe

Informe: Títulos controla el contenido y la ubicación de los títulos y los pies de página del informe. Puede especificar un máximo de diez líneas de títulos de página y otras tantas de pies de página, con componentes justificados a la izquierda, en el centro y a la derecha en cada línea.

Figura 38-7  
Cuadro de diálogo Informe: Títulos



Si inserta variables en los títulos o en los pies de página, la etiqueta de valor o el valor de la variable actual aparecerá en el título o en el pie de página. Para los títulos se mostrará la etiqueta de valor correspondiente al valor de la variable al principio de la página; para los pies de página, esta etiqueta se mostrará al final de la página. Si no hay etiqueta de valor, se mostrará el valor real.

**Variables especiales.** Las variables especiales *DATE* y *PAGE* permiten insertar la fecha actual o el número de página en cualquier línea de un encabezado o pie del informe. Si el archivo de datos contiene variables llamadas *DATE* o *PAGE*, no podrá utilizar estas variables en los títulos ni en los pies del informe.

## Informe de estadísticos en columnas

Informe de estadísticos en columnas genera informes de resumen en los que diversos estadísticos de resumen aparecen en columnas distintas.

**Ejemplo.** Una empresa con una cadena de tiendas registra la información de los empleados, incluyendo el salario, el cargo y la sección en la que trabaja cada uno. Se podría generar un informe que proporcione los estadísticos de salario resumidos (por ejemplo, media, mínimo y máximo) para cada sección.

**Columnas de datos.** Muestra una lista de las variables del informe para las que se desea obtener estadísticos de resumen y controla el formato de presentación y los estadísticos de resumen mostrados para cada variable.

**Romper columnas por.** Muestra una lista de las variables de segmentación opcionales que dividen el informe en grupos y controla los formatos de presentación de las columnas de ruptura. Si hay varias variables de ruptura, se creará un grupo distinto para cada una de las categorías de las variables de ruptura dentro de las categorías de la variable de ruptura anterior en la lista. Las variables de ruptura deben ser variables categóricas discretas que dividan los casos en un número limitado de categorías con sentido.

**Informe.** Controla las características globales del informe, incluyendo la presentación de los valores perdidos, la numeración de las páginas y los títulos.

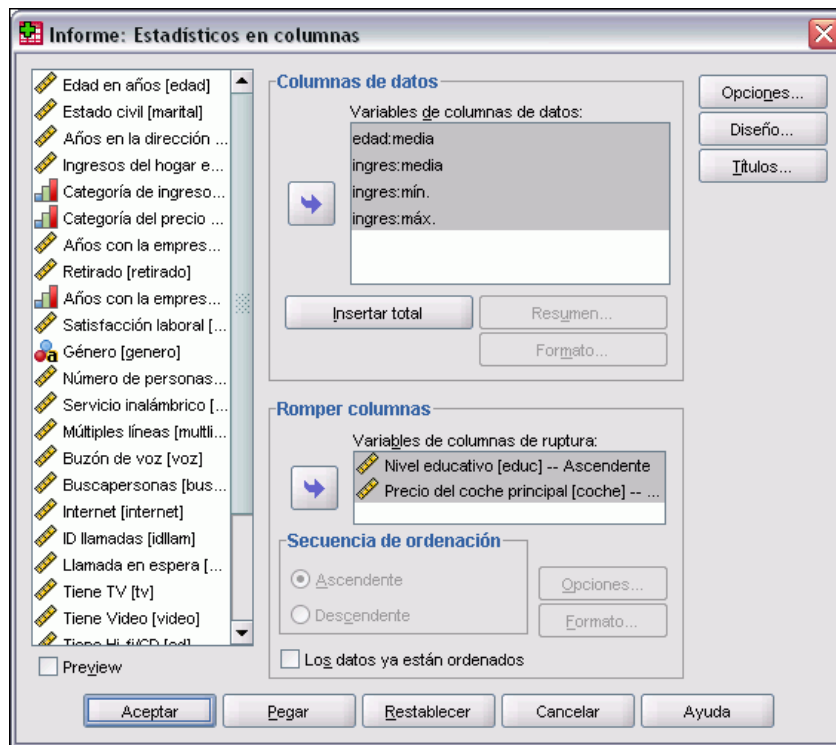
**Presentación preliminar.** Muestra sólo la primera página del informe. Esta opción es útil para ver una presentación preliminar del formato del informe sin tener que procesar el informe completo.

**Los datos están ordenados.** Para los informes con variables de ruptura, el archivo de datos se debe ordenar por los valores de estas variables antes de generar el informe. Si el archivo de datos ya está ordenado por estos valores, se puede ahorrar tiempo de procesamiento seleccionando esta opción. Esta opción es especialmente útil después de generar la presentación preliminar de un informe.

### ***Para obtener un informe de resumen: Estadísticos en columnas***

- ▶ Elija en los menús:
  - Analizar
  - Informes
  - Informe de estadísticos en columnas...
- ▶ Seleccione una o más variables para las columnas de datos. En el informe se genera una columna para cada variable seleccionada.
- ▶ Para cambiar la medida de resumen para una variable, seleccione la variable de la lista de columnas de datos y pulse en Resumen.
- ▶ Para obtener más de una medida de resumen para una variable, seleccione la variable en la lista de origen y desplácela hasta la lista Columnas de datos varias veces, una para cada medida que desee obtener.
- ▶ Para mostrar una columna con la suma, la media, la razón o cualquier otra función de las columnas existentes, pulse en Insertar total. Al hacerlo se situará una variable llamada *total* en la lista Columnas de datos.
- ▶ Para los informes ordenados y mostrados por subgrupos, seleccione una o más variables para Romper columnas por.

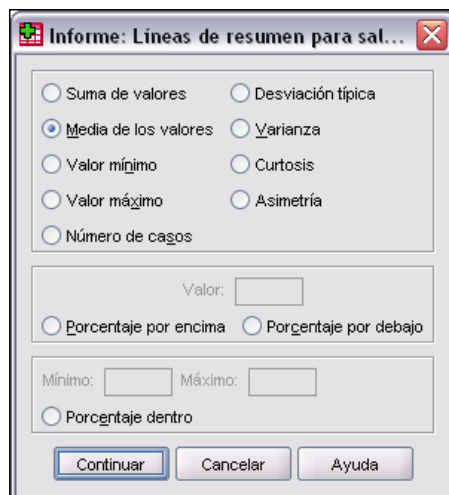
Figura 38-8  
Cuadro de diálogo Informe: Estadísticos en las columnas



### ***Función Columna de resumen total***

Líneas de resumen controla el estadístico de resumen mostrado para la variable de las columnas de datos seleccionada.

Figura 38-9  
Cuadro de diálogo Informe: Líneas de resumen





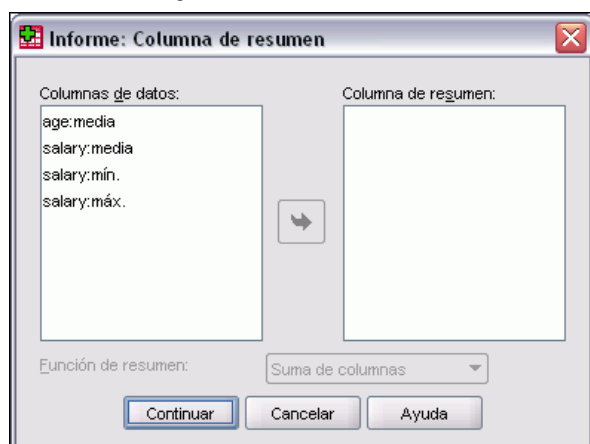
Los estadísticos de resumen disponibles son: suma, media, valor mínimo, valor máximo, número de casos, porcentaje de casos por encima y por debajo de un valor especificado, porcentaje de casos dentro de un rango de valores especificado, desviación típica, varianza, curtosis y asimetría.

### **Columna de resumen total**

Columna de resumen controla los estadísticos de resumen del total que resumen dos o más columnas de datos.

Los estadísticos de resumen del total son la suma de columnas, la media de columnas, el mínimo, el máximo, la diferencia entre los valores de dos columnas, el cociente de los valores de una columna dividido por los valores de otra y el producto de los valores de las columnas multiplicados entre sí.

Figura 38-10  
Cuadro de diálogo Informe: Columna de resumen



**Suma de columnas.** La columna *total* es la suma de las columnas de la lista Columna de resumen.

**Media de columnas.** La columna *total* es la media de las columnas de la lista Columna de resumen.

**Mínimo de columnas.** La columna *total* es el mínimo de las columnas de la lista Columna de resumen.

**Máximo de columnas.** La columna *total* es el máximo de las columnas de la lista Columna de resumen.

**1ª columna - 2ª columna.** La columna *total* es la resta de las columnas de la lista Columna de resumen. Esta lista debe contener, exactamente, dos columnas.

**1ª columna / 2ª columna.** La columna *total* es el cociente de las columnas de la lista Columna de resumen. Esta lista debe contener, exactamente, dos columnas.

**% 1ª columna / 2ª columna.** La columna *total* es el porcentaje de la primera columna dividido por la segunda columna de la lista Columna de resumen. Esta lista debe contener, exactamente, dos columnas.

**Producto de columnas.** La columna *total* es el producto de las columnas de la lista Columna de resumen.

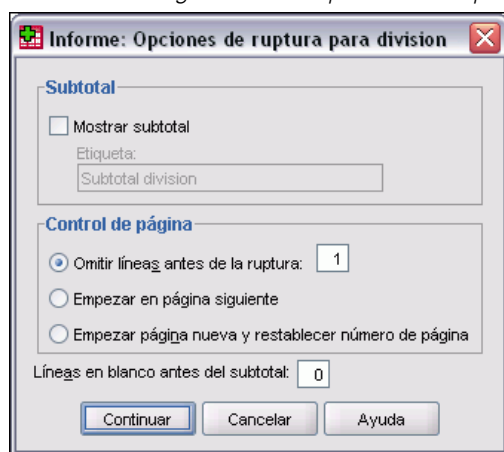
## Formato de columna del informe

Las opciones de formato de columna de ruptura y de datos para el informe de estadísticos en columnas son las mismas que las descritas para el informe de estadísticos en filas.

## Opciones de la ruptura de columnas para los estadísticos en el informe

Opciones de ruptura controla la presentación del subtotal, el espaciado y la paginación para las categorías de ruptura.

Figura 38-11  
Cuadro de diálogo Informe: Opciones de ruptura



**Subtotal.** Controla los subtotales mostrados para cada categoría de ruptura.

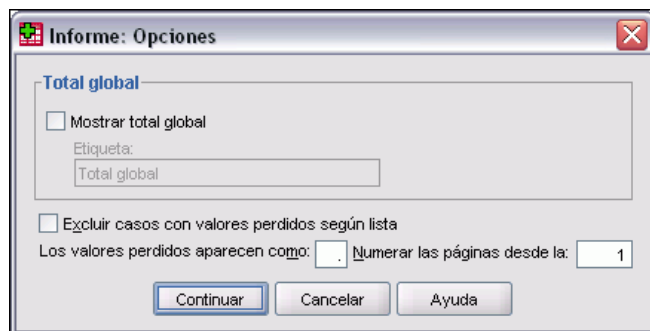
**Control de página.** Controla el espaciado y la paginación para las categorías de la variable de ruptura seleccionada. Puede especificar un número de líneas en blanco entre las categorías de ruptura o empezar cada categoría de ruptura en una página nueva.

**Líneas en blanco antes del subtotal.** Controla el número de líneas en blanco entre los datos de las categorías de ruptura y los subtotales.

## Opciones de columnas para los estadísticos en el informe

Opciones controla la presentación de los totales finales y de los valores perdidos y la paginación de los informes de estadísticos en columnas.

Figura 38-12  
Cuadro de diálogo Informe: Opciones



**Total final.** Muestra y etiqueta un total global para cada columna que aparece al final de la columna.

**Valores perdidos.** Permite excluir los valores perdidos del informe o seleccionar un único carácter para indicar estos valores.

## ***Diseño del informe para los estadísticos en columnas***

Las opciones de diseño para el informe de estadísticos en columnas son las mismas que las descritas para el informe de estadísticos en filas.

## ***Funciones adicionales del comando REPORT***

Con el lenguaje de sintaxis de comandos también podrá:

- Mostrar funciones de resumen diferentes en las columnas de una única línea de resumen.
- Insertar líneas de resumen en las columnas de datos para variables que no sean la variable de la columna de datos o para diversas combinaciones (funciones compuestas) de las funciones de resumen.
- Utilizar Mediana, Moda, Frecuencia y Porcentaje como funciones de resumen.
- Controlar de forma más precisa el formato de presentación de los estadísticos de resumen.
- Insertar líneas en blanco en diversos puntos de los informes.
- Insertar líneas en blanco después de cada  $n$ -ésimo caso de los informes en formato de listado.

Debido a la complejidad de la sintaxis de REPORT, a la hora de construir un nuevo informe con sintaxis puede resultarle útil, para aproximar el informe generado a partir de los cuadros de diálogo, copiar y pegar la sintaxis correspondiente y depurar esa sintaxis para generar exactamente el informe que le interese.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

## ***Análisis de fiabilidad***

El análisis de fiabilidad permite estudiar las propiedades de las escalas de medición y de los elementos que componen las escalas. El procedimiento Análisis de fiabilidad calcula un número de medidas de fiabilidad de escala que se utilizan normalmente y también proporciona información sobre las relaciones entre elementos individuales de la escala. Se pueden utilizar los coeficientes de correlación intraclase para calcular estimaciones de la fiabilidad inter-evaluadores.

**Ejemplo.** ¿El cuestionario mide la satisfacción del cliente de manera útil? El análisis de fiabilidad le permitirá determinar el grado en que los elementos del cuestionario se relacionan entre sí, obtener un índice global de la replicabilidad o de la consistencia interna de la escala en su conjunto e identificar elementos problemáticos que deberían ser excluidos de la escala.

**Estadísticos.** Descriptivos para cada variable y para la escala, estadísticos de resumen comparando los elementos, correlaciones y covarianzas inter-elementos, estimaciones de la fiabilidad, tabla de ANOVA, coeficientes de correlación intraclase,  $T^{\text{cuadrado}}$  de Hotelling y prueba de aditividad de Tukey.

**Modelos.** Están disponibles los siguientes modelos de fiabilidad:

- **Alfa (Cronbach).** Este modelo es un modelo de consistencia interna, que se basa en la correlación inter-elementos promedio.
- **Dos mitades.** Este modelo divide la escala en dos partes y examina la correlación entre dichas partes.
- **Guttman.** Este modelo calcula los límites inferiores de Guttman para la fiabilidad verdadera.
- **Paralelo.** Este modelo asume que todos los elementos tienen varianzas iguales y varianzas error iguales a través de las réplicas.
- **Estrictamente paralelo.** Este modelo asume los supuestos del modelo paralelo y también asume que las medias son iguales a través de los elementos.

**Datos.** Los datos pueden ser dicotómicos, ordinales o de intervalo, pero deben estar codificados numéricamente.

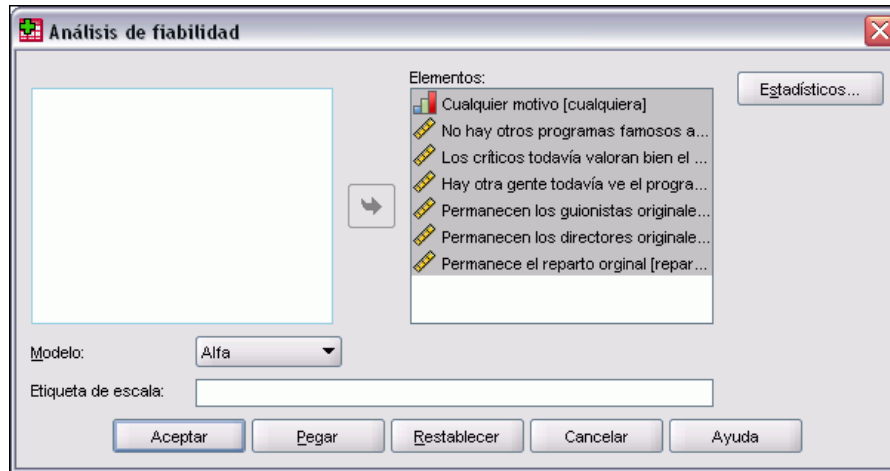
**Supuestos.** Las observaciones deben ser independientes y los errores no deben estar correlacionados entre los elementos. Cada par de elementos debe tener una distribución normal bivariada. Las escalas deben ser aditivas, de manera que cada elemento esté linealmente relacionado con la puntuación total.

**Procedimientos relacionados.** Si desea explorar la dimensionalidad de los elementos de la escala (para comprobar si es necesario más de un constructo para explicar el patrón de puntuaciones en los elementos), utilice el Análisis factorial o el Escalamiento multidimensional. Para identificar grupos homogéneos de variables, use el análisis de conglomerados jerárquico para agrupar las variables.

**Para obtener un análisis de fiabilidad**

- ▶ Elija en los menús:
  - Analizar
  - Escala
  - Análisis de fiabilidad...

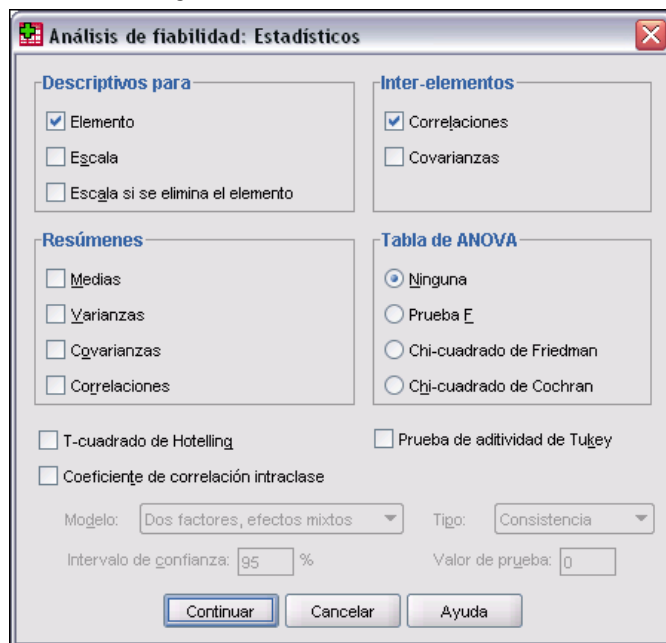
Figura 39-1  
Cuadro de diálogo Análisis de fiabilidad



- ▶ Seleccione dos o más variables como componentes potenciales de una escala aditiva.
- ▶ Elija un modelo de la lista desplegable Modelo.

## Análisis de fiabilidad: Estadísticos

Figura 39-2  
Cuadro de diálogo Análisis de fiabilidad: Estadísticos



Puede seleccionar diversos estadísticos que describen la escala y sus elementos. Los estadísticos de los que se informa por defecto incluyen el número de casos, el número de elementos y las estimaciones de la fiabilidad, según se explica a continuación:

- **Modelos Alfa.** Coeficiente Alfa; para datos dicotómicos, éste es equivalente al coeficiente 20 de Kuder-Richardson (KR20).
- **Modelos Dos mitades.** Correlación entre formas, fiabilidad de dos mitades de Guttman, fiabilidad de Spearman-Brown (longitud igual y desigual) y coeficiente alfa para cada mitad.
- **Modelos de Guttman.** Coeficientes de fiabilidad lambda 1 a lambda 6.
- **Modelos Paralelo y Estrictamente paralelo.** Prueba de bondad de ajuste del modelo; estimaciones de la varianza error, varianza común y varianza verdadera; correlación común inter-elementos estimada; fiabilidad estimada y estimación de la fiabilidad no sesgada.

**Descriptivos para.** Genera estadísticos descriptivos para las escalas o los elementos a través de los casos.

- **Elemento.** Genera estadísticos descriptivos para los elementos a través de los casos.
- **Escala.** Genera estadísticos descriptivos para las escalas.
- **Escala si se elimina el elemento.** Muestra estadísticos de resumen para comparar cada elemento con la escala compuesta por otros elementos. Los estadísticos incluyen la media de escala y la varianza si el elemento fuera a eliminarse de la escala, la correlación entre el elemento y la escala compuesta por otros elementos, y alfa de Cronbach si el elemento fuera a eliminarse de la escala.

**Resúmenes.** Proporciona estadísticos descriptivos sobre las distribuciones de los elementos a través de todos los elementos de la escala.

- **Medias.** Estadísticos de resumen para las medias de los elementos. Se muestran el máximo, el mínimo y el promedio de las medias de los elementos, el rango y la varianza de las medias de los elementos, y la razón de la mayor media sobre la menor media de los elementos.
- **Varianzas.** Estadísticos de resumen de las varianzas de los elementos (ítems). Se muestran: la menor, la mayor y el promedio de las varianzas de los ítems, el rango y la varianza de las varianzas de los ítems, y la razón de la mayor varianza sobre la menor varianza de los ítems.
- **Covarianzas.** Estadísticos de resumen para las covarianzas entre los elementos. Se muestran el máximo, el mínimo y el promedio de las covarianzas inter-elementos, el rango y la varianza de las covarianzas inter-elementos, y la razón de la mayor covarianza sobre la menor covarianza inter-elementos.
- **Correlaciones.** Estadísticos de resumen para las correlaciones entre los elementos. Se muestran el máximo, el mínimo y el promedio de las correlaciones inter-elementos, el rango y la varianza de las correlaciones inter-elementos, y la razón de la mayor correlación sobre la menor correlación inter-elementos.

**Inter-elementos.** Genera las matrices de correlaciones o covarianzas entre los elementos.

**Tabla de ANOVA.** Produce pruebas de medias iguales.

- **Prueba F.** Muestra la tabla de un análisis de varianza de medidas repetidas.
- **Chi-cuadrado de Friedman.** Muestra el chi-cuadrado de Friedman y el coeficiente de concordancia de Kendall. Esta opción es adecuada para datos que se encuentren en el formato de rangos. La prueba de chi-cuadrado sustituye a la prueba F habitual en la tabla de ANOVA.
- **Chi-cuadrado de Cochran.** Muestra la Q de Cochran. Esta opción es adecuada para datos que son dicotómicos. El estadístico Q sustituye a la F habitual en la tabla de ANOVA.

**T-cuadrado de Hotelling.** Genera un contraste multivariado sobre la hipótesis nula de que todos los elementos de la escala tienen la misma media.

**Prueba de aditividad de Tukey.** Genera un contraste sobre el supuesto de que no existe una interacción multiplicativa entre los elementos.

**Coefficiente de correlación intraclase.** Genera medidas sobre la consistencia o sobre el acuerdo de los valores entre los propios casos.

- **Modelo.** Seleccione el modelo para calcular el coeficiente de correlación intraclase. Los modelos disponibles son: Dos factores, efectos mixtos; Dos factores, efectos aleatorios y Un factor, efectos aleatorios. Seleccione Dos factores, efectos mixtos, si los efectos de personas son aleatorios y los efectos de elementos son fijos, Dos factores, efectos aleatorios, si los efectos de personas y los efectos de elementos son aleatorios; o Un factor, efectos aleatorios si los efectos de personas son aleatorios.
- **Tipo.** Seleccione el tipo de índice. Los tipos disponibles son: Los tipos disponibles son: Consistencia y Acuerdo absoluto.

- **Intervalo de confianza.** Especifica el nivel para el intervalo de confianza. El valor por defecto es 95%.
- **Valor de prueba.** Especifica el valor hipotetizado para el coeficiente, para el contraste de hipótesis. Este valor es el valor con el que se compara el valor observado. El valor por defecto es 0.

### ***Funciones adicionales del comando RELIABILITY***

Con el lenguaje de sintaxis de comandos también podrá:

- Leer y analizar una matriz de correlaciones.
- Escribir una matriz de correlaciones para su análisis posterior.
- Especificar una división distinta de la de dos mitades iguales, para el método de dos mitades.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).



## ***Escalamiento multidimensional***

El escalamiento multidimensional trata de encontrar la estructura de un conjunto de medidas de distancia entre objetos o casos. Esta tarea se logra asignando las observaciones a posiciones específicas en un espacio conceptual (normalmente de dos o tres dimensiones) de modo que las distancias entre los puntos en el espacio concuerden al máximo con las disimilaridades dadas. En muchos casos, las dimensiones de este espacio conceptual son interpretables y se pueden utilizar para comprender mejor los datos.

Si las variables se han medido objetivamente, puede utilizar el escalamiento multidimensional como técnica de reducción de datos (el procedimiento Escalamiento multidimensional permitirá calcular las distancias a partir de los datos multivariados, si es necesario). El escalamiento multidimensional puede también aplicarse a valoraciones subjetivas de disimilaridad entre objetos o conceptos. Además, el procedimiento Escalamiento multidimensional puede tratar datos de disimilaridad procedentes de múltiples fuentes, como podrían ser múltiples evaluadores o múltiples sujetos evaluados por un cuestionario.

**Ejemplo.** ¿Cómo percibe el público las diferencias entre distintos coches? Si posee datos de las valoraciones de similaridad emitidas por los sujetos sobre las diferentes marcas y modelos de coches, puede utilizar el escalamiento multidimensional para identificar las dimensiones que describan las preferencias de los consumidores. Puede encontrar, por ejemplo, que el precio y el tamaño de un vehículo definen un espacio de dos dimensiones, capaz de explicar las similitudes de las que informan los encuestados.

**Estadísticos.** Para cada modelo: Matriz de datos, Matriz de datos escalada óptimamente, S-stress (de Young), Stress (de Kruskal),  $R^2$ , Coordenadas de los estímulos, Stress promedio y  $R^2$  para cada estímulo (modelos RMDS). Para modelos de diferencias individuales (INDSCAL): ponderaciones del sujeto e índice de peculiaridad para cada sujeto. Para cada matriz en los modelos de escalamiento multidimensional replicado: stress y  $R^2$  para cada estímulo. Gráficos: coordenadas de los estímulos (de dos o tres dimensiones), diagrama de dispersión de las disparidades frente a las distancias.

**Datos.** Si los datos son de disimilaridad, todas las disimilaridades deben ser cuantitativas y deben estar medidas en la misma métrica. Si los datos son datos multivariantes, las variables pueden ser datos cuantitativos, binarios o de recuento. El escalamiento de las variables es un tema importante, ya que las diferencias en el escalamiento pueden afectar a la solución. Si las variables tienen grandes diferencias en el escalamiento (por ejemplo, una variable se mide en dólares y otra en años), debe considerar la posibilidad de tipificarlas (este proceso puede llevarse a cabo automáticamente con el propio procedimiento Escalamiento multidimensional).

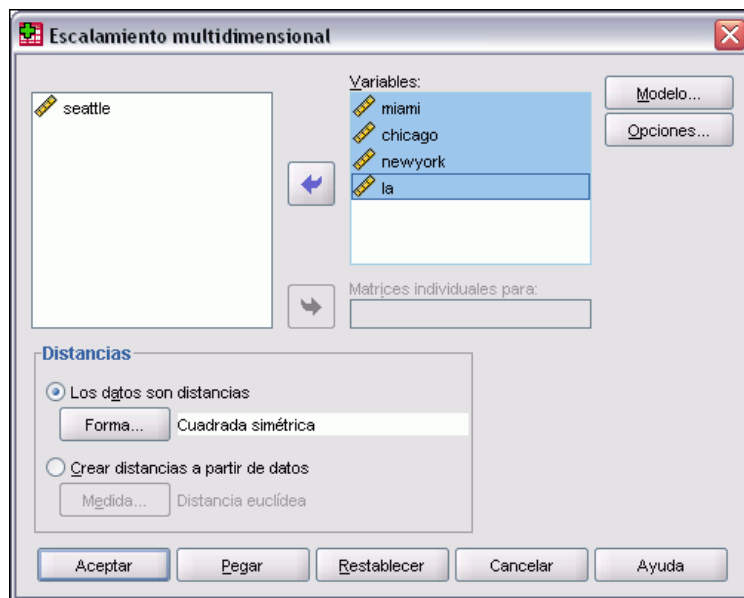
**Supuestos.** El procedimiento Escalamiento multidimensional está relativamente libre de supuestos distribucionales. Compruebe que selecciona el nivel de medida adecuado (ordinal, de intervalo, o de razón) en el cuadro de diálogo Escalamiento multidimensional: Opciones para asegurar que los resultados se calculan correctamente.

**Procedimientos relacionados.** Si su objetivo es la reducción de los datos, un método alternativo a tener en cuenta es el análisis factorial, sobre todo si las variables son cuantitativas. Si desea identificar grupos de casos similares, considere complementar el análisis de escalamiento multidimensional con un análisis de conglomerados jerárquico o de  $k$ -medias.

### **Para obtener un análisis de escalamiento multidimensional**

- ▶ Elija en los menús:  
   Analizar  
     Escala  
       Escalamiento multidimensional...

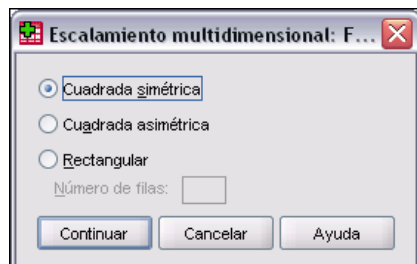
Figura 40-1  
 Cuadro de diálogo Escalamiento multidimensional



- ▶ En Distancias, seleccione Los datos son distancias o Crear distancias a partir de datos.
- ▶ Si los datos son distancias, seleccione al menos cuatro variables numéricas para el análisis. (También puede pulsar en Forma para indicar la forma de la matriz de distancias.)
- ▶ Si desea crear las distancias antes de analizarlas, seleccione al menos una variable numérica. (También puede pulsar en Medida para indicar el tipo de medida de distancia que desea.) Cada variable de agrupación puede ser numérica o de cadena, y puede crear matrices distintas para cada categoría de una variable de agrupación moviendo dicha variable a la lista Matrices individuales para.

## Escalamiento multidimensional: Forma de los datos

Figura 40-2  
Cuadro de diálogo Escalamiento multidimensional: Forma

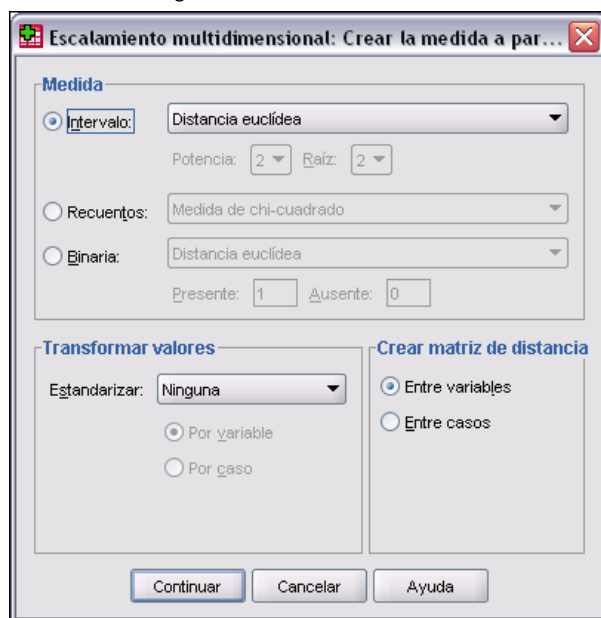


Si el conjunto de datos activo representa distancias entre uno o dos conjuntos de objetos, especifique la forma de la matriz de datos para obtener los resultados correctos.

*Nota:* No puede seleccionar Cuadrada simétrica si el cuadro de diálogo Modelo especifica la condicionalidad de filas.

## Escalamiento multidimensional: Crear la medida a partir de los datos

Figura 40-3  
Cuadro de diálogo Escalamiento multidimensional: Crear la medida a partir de los datos



El escalamiento multidimensional utiliza datos de disimilaridad para crear una solución de escalamiento. Si los datos son datos multivariantes (los valores de las variables que se han medido), debe crear los datos de disimilaridad para poder calcular una solución de escalamiento multidimensional. Puede especificar los detalles para la creación de las medidas de disimilaridad a partir de los datos.

**Medida.** Le permite especificar la medida de disimilaridad para el análisis. Seleccione una opción del grupo Medida que se corresponda con el tipo de datos y, a continuación, elija una de las medidas de la lista desplegable correspondiente a ese tipo de medida. Las opciones disponibles son:

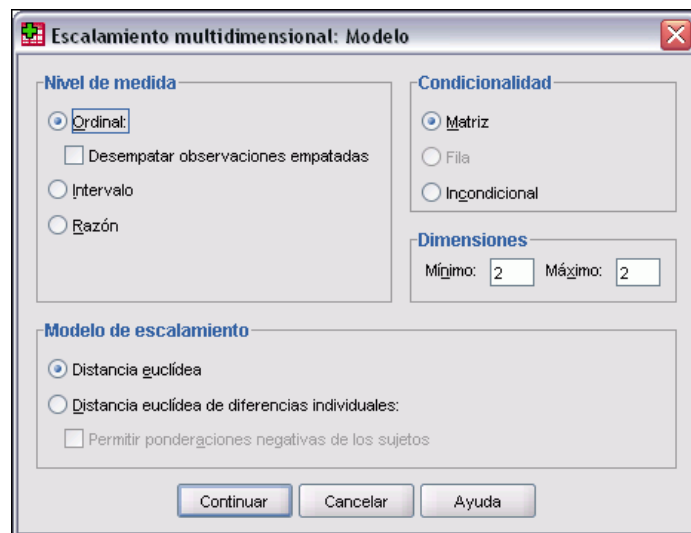
- **Intervalo.** Distancia euclídea, Distancia euclídea al cuadrado, Chebychev, Bloque, Minkowski o Personalizada.
- **Contar apariciones.** Medida de chi-cuadrado o Medida de phi-cuadrado.
- **Binaria.** Distancia euclídea, Distancia euclídea al cuadrado, Diferencia de tamaño, Diferencia de configuración, Varianza o Lance y Williams.

**Crear matriz de distancias.** Le permite elegir la unidad de análisis. Las opciones son Entre variables o Entre casos.

**Transformar valores.** En determinados casos, como cuando las variables se miden en escalas muy distintas, puede que desee tipificar los valores antes de calcular las proximidades (no es aplicable a datos binarios). Seleccione un método de estandarización en la lista desplegable Estandarizar. Si no se requiere ninguna estandarización, seleccione Ninguno.

## Escalamiento multidimensional: Modelo

Figura 40-4  
Cuadro de diálogo Escalamiento multidimensional: Modelo



La estimación correcta de un modelo de escalamiento multidimensional depende de aspectos que atañen a los datos y al modelo en sí.

**Nivel de medida.** Permite especificar el nivel de medida de los datos. Las opciones son Ordinal, Intervalo y Razón. Si las variables son ordinales, al seleccionar Desempatar observaciones empatadas se solicitará que sean consideradas como variables continuas, de forma que los empates (valores iguales para casos diferentes) se resuelvan óptimamente.

**Condicionabilidad.** Permite especificar qué comparaciones tienen sentido. Las opciones son Matriz, Fila o Incondicional.

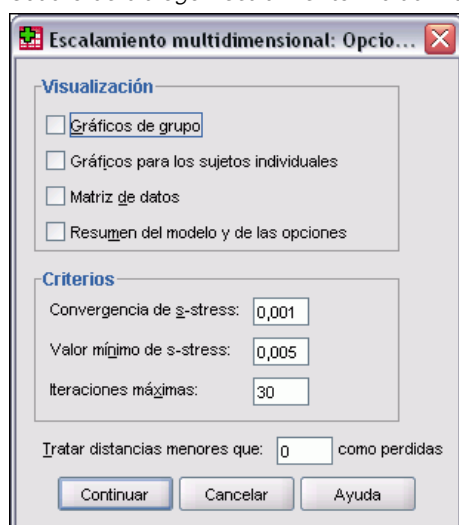
**Dimensiones.** Permite especificar la dimensionalidad de la solución o soluciones del escalamiento. Se calcula una solución para cada número del rango especificado. Especifique números enteros entre 1 y 6; se permite un mínimo de 1 sólo si selecciona Distancia euclídea como modelo de escalamiento. Para una solución única, especifique el mismo número para el mínimo y el máximo.

**Modelo de escalamiento.** Permite especificar los supuestos bajo los que se realiza el escalamiento. Las opciones disponibles son Distancia euclídea o Distancia euclídea de diferencias individuales (también conocida como INDSCAL). Para el modelo de Distancia euclídea de diferencias individuales, puede seleccionar Permitir ponderaciones negativas de los sujetos, si es adecuado para los datos.

## Escalamiento multidimensional: Opciones

Figura 40-5

Cuadro de diálogo Escalamiento multidimensional: Opciones



Puede especificar opciones para el análisis de escalamiento multidimensional.

**Mostrar.** Permite seleccionar varios tipos de resultados. Las opciones disponibles son Gráficos de grupo, Gráficos para los sujetos individuales, Matriz de datos y Resumen del modelo y de las opciones.

**Criterios.** Permite determinar cuándo debe detenerse la iteración. Para cambiar los valores por defecto, introduzca valores para la Convergencia de s-stress, el Valor mínimo de s-stress y el N° máximo de iteraciones.

**Tratar distancias menores que n como perdidas.** Las distancias menores que este valor se excluyen del análisis.

## ***Funciones adicionales del comando de ALSCAL***

Con el lenguaje de sintaxis de comandos también podrá:

- Utilizar tres tipos de modelos adicionales, conocidos como ASCAL, AINDS y GEMSCAL en la bibliografía acerca del escalamiento multidimensional.
- Realizar transformaciones polinómicas en los datos de intervalo y de razón.
- Analizar similitudes (en lugar de distancias) con los datos ordinales.
- Analizar datos nominales.
- Guardar en archivos varias matrices de ponderación y coordenadas y leerlas posteriormente para el análisis.
- Restringir el despliegamiento multidimensional.

Si desea información detallada sobre la sintaxis, consulte la referencia de sintaxis de comandos (*Command Syntax Reference*).

## ***Estadísticos de la razón***

El procedimiento Estadísticos de la razón proporciona una amplia lista de estadísticos de resumen para describir la razón entre dos variables de escala.

Se pueden ordenar los resultados por los valores de una variable de agrupación, en orden ascendente o descendente. Se puede eliminar de los resultados el informe de los estadísticos de la razón y almacenar los resultados en un archivo externo.

**Ejemplo.** ¿Existe una buena uniformidad en la razón entre el precio de tasación y el precio de venta de viviendas en cada una de las cinco regiones? En los resultados, se puede descubrir que la distribución de las razones varía considerablemente entre regiones.

**Estadísticos.** Mediana, media, media ponderada, intervalos de confianza, coeficiente de dispersión (CDD), coeficiente de variación centrado en la mediana, coeficiente de variación centrado en la media, el diferencial de precio (DRV), desviación típica, desviación absoluta promedio (DAP), rango, valores mínimos y máximos y el índice de concentración calculado dentro de un rango o porcentaje (especificados por el usuario) respecto a la razón mediana.

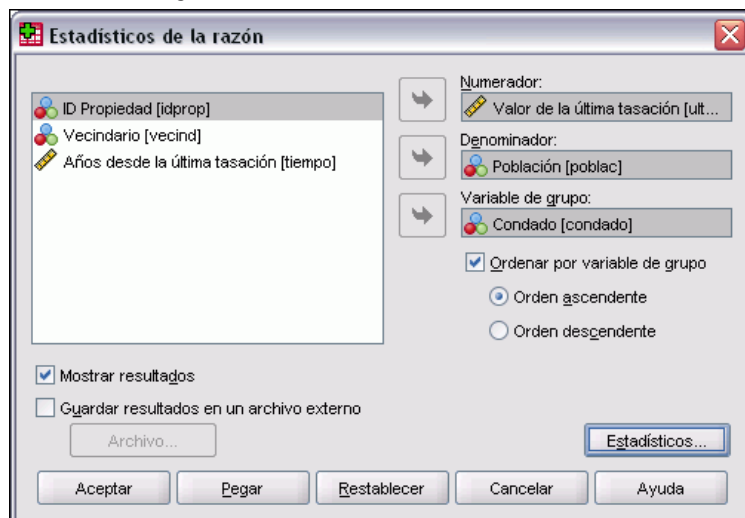
**Datos.** Utilice códigos numéricos o cadenas para codificar las variables de agrupación (medidas de nivel nominal u ordinal).

**Supuestos.** Las variables que definen el numerador y el denominador de la razón deben ser variables de escala, que toman valores positivos.

### ***Para obtener estadísticos de la razón***

- ▶ Elija en los menús:
  - Analizar
  - Estadísticos descriptivos
  - Razón...

Figura 41-1  
Cuadro de diálogo Estadísticos de la razón



- ▶ Seleccione una variable de numerador.
- ▶ Seleccione una variable de denominador.

Si lo desea:

- Seleccione una variable de agrupación y especifique el orden de los grupos en los resultados.
- Elija si desea mostrar los resultados en el Visor.
- Elija si desea guardar los resultados en un archivo externo para un uso posterior y especifique el nombre del archivo en el que se van a guardar los resultados.



## Estadísticos de la razón

Figura 41-2  
Cuadro de diálogo Estadísticos de la razón

**Tendencia central.** Las medidas de tendencia central son estadísticos que describen la distribución de las razones.

- **Mediana.** Un valor tal que el número de razones menores que este valor es igual al número de razones mayores que el mismo.
- **Media.** El resultado de sumar las razones y dividir la suma entre el número total de razones.
- **Media ponderada.** El resultado de dividir la media del numerador entre la media del denominador. La media ponderada es también la media de las razones ponderadas por el denominador.
- **Intervalos de confianza.** Muestra los intervalos de confianza para la media, la mediana y la media ponderada (si se solicita). Especifique un valor mayor o igual que 0 y menor que 100 como nivel de confianza.

**Dispersión.** Estos estadísticos miden la cantidad de variación o de dispersión entre los valores observados.

- **DAP.** La desviación absoluta promedio es el resultado de sumar las desviaciones absolutas de las razones respecto a la mediana y dividir el resultado entre el número total de razones.
- **CDD.** El coeficiente de dispersión es el resultado de expresar la desviación absoluta promedio como un porcentaje de la mediana.
- **DRP.** El diferencial relativo al precio, también conocido como el índice de regresibilidad, es el resultado de dividir la media por la media ponderada.

- **CDV centrado en la mediana.** El coeficiente de variación centrado en la mediana es el resultado de expresar la raíz de la media cuadrática de las desviaciones respecto a la mediana como un porcentaje de la mediana.
- **CDV centrado en la media.** El coeficiente de variación centrado en la media es el resultado de expresar la desviación típica como un porcentaje de la media.
- **Desviación típica.** La desviación típica es el resultado de sumar las desviaciones cuadráticas de las razones respecto a la media, dividir la suma por el número total de razones menos uno y extraer la raíz cuadrada positiva.
- **Rango.** El rango es el resultado de restar la razón mínima de la razón máxima.
- **Mínimo.** El mínimo es la razón menor.
- **Máximo.** El máximo es la razón mayor.

**Índice de concentración.** El coeficiente de concentración mide el porcentaje de razones que están dentro de un intervalo. Se puede calcular de dos maneras:

- **Razones dentro del.** En este caso, el intervalo se define de forma explícita especificando los valores superior e inferior del intervalo. Introduzca valores para las proporciones superior e inferior y pulse en Añadir para obtener un intervalo.
- **Razones en.** En este caso, el intervalo se define de forma implícita al especificar el porcentaje de la mediana. Introduzca un valor entre 0 y 100 y pulse en Añadir. El límite inferior del intervalo será igual a  $(1 - 0,01 \times \text{valor}) \times \text{mediana}$ , y el límite superior será igual a  $(1 + 0,01 \times \text{valor}) \times \text{mediana}$ .

## ***Curvas COR***

Este procedimiento es un método útil para evaluar la realización de esquemas de clasificación en los que exista una variable con dos categorías por las que se clasifiquen los sujetos.

**Ejemplo.** Un banco tiene interés en clasificar a sus clientes dependiendo de si se retrasarán o no en el pago de sus préstamos; por tanto, se desarrollan métodos especiales para tomar estas decisiones. Las curvas COR se pueden utilizar para evaluar el grado de acierto de estos métodos.

**Estadísticos.** Es un área situada bajo la curva COR con un intervalo de confianza y puntos de coordenadas de la curva COR. Gráficos: Curva COR.

**Métodos.** Se puede calcular la estimación del área situado bajo la curva COR de forma paramétrica o no paramétrica mediante un modelo exponencial binegativo.

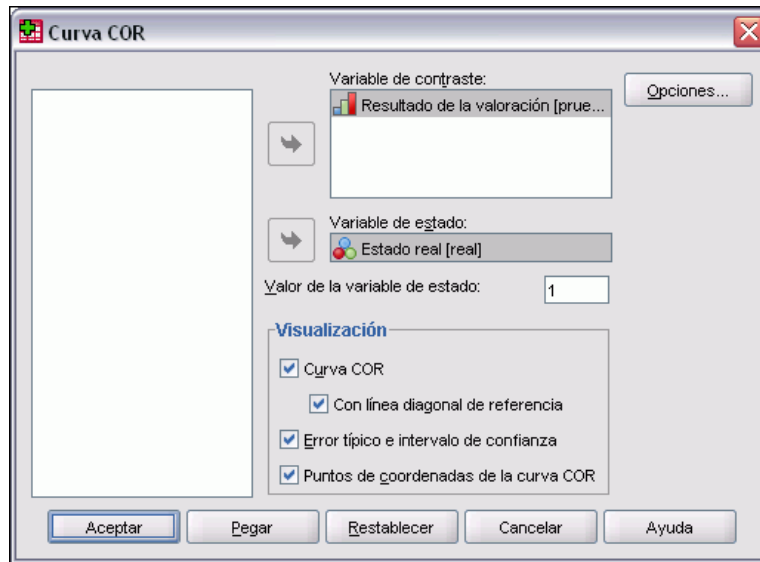
**Datos.** Las variables de contraste son cuantitativas. Las variables de contraste suelen estar constituidas por probabilidades, resultantes de un análisis discriminante o de una regresión logística, o bien compuestas por puntuaciones atribuidas en una escala arbitraria que indican el «grado de convicción» que tiene un evaluador de que el sujeto pueda pertenecer a una u otra categoría. La variable de estado puede ser de cualquier tipo e indicar la categoría real a la que pertenece un sujeto. El valor de la variable de estado indica la categoría que se debe considerar *positiva*.

**Supuestos.** Se considera que los números ascendentes de la escala del evaluador representan la creciente convicción de que el sujeto pertenece a una categoría. Por el contrario, los números descendentes representan la creciente convicción de que el sujeto pertenece a la otra categoría. El usuario deberá elegir qué dirección es *positiva*. También se considera que se conoce la categoría *real* a la que pertenece el sujeto.

### ***Para obtener una curva COR***

- ▶ Elija en los menús:  
   Analizar  
   Curva COR...

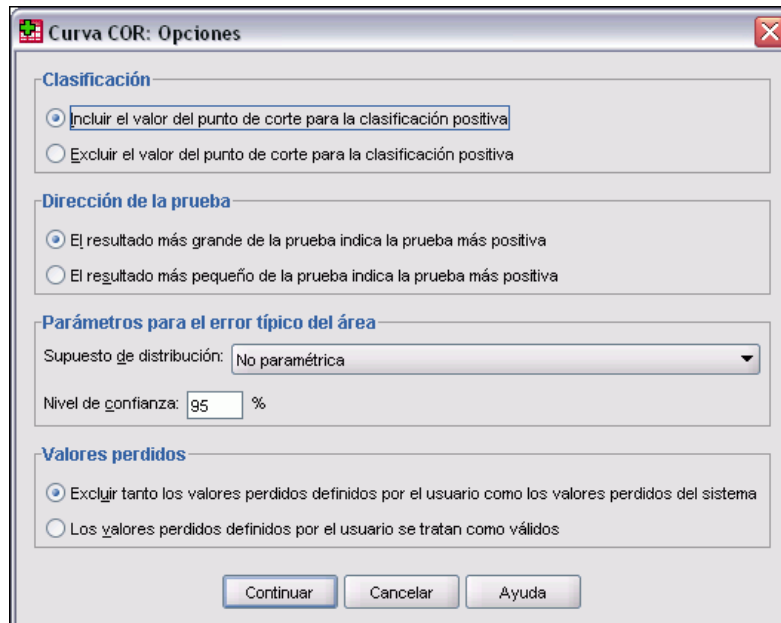
Figura 42-1  
Cuadro de diálogo Curva COR



- ▶ Seleccione una o más variables de probabilidad de contraste.
- ▶ Elija una variable de estado.
- ▶ Identifique el valor *positivo* para la variable de estado.

## Curvas COR: Opciones

Figura 42-2  
Cuadro de diálogo Curva COR: Opciones



Puede seleccionar las opciones siguientes para su análisis:

**Clasificación.** Permite especificar si se debe incluir o excluir el valor del punto de corte al realizar una clasificación *positiva*. Este ajuste no afecta a los resultados.

**Dirección de la prueba.** Permite especificar la dirección de la escala según la categoría *positiva*.

**Parámetros para el error típico del área.** Permite especificar el método de estimación del error típico del área situada bajo la curva. Los métodos disponibles son el no paramétrico y el exponencial binegativo. También se puede establecer el nivel para el intervalo de confianza. El rango disponible es entre el 50,1% y el 99,9%.

**Valores perdidos.** Permite especificar el tratamiento que reciben los valores perdidos.

# ***Conceptos básicos de la utilidad de gráficos***

Se pueden crear gráficos de alta resolución mediante los procedimientos del menú Gráficos y mediante muchos de los procedimientos del menú Analizar. Este capítulo explica los conceptos básicos de la utilidad de gráficos.

## ***Creación y modificación de gráficos***

Antes de crear un gráfico es necesario tener los datos en el Editor de datos. Es posible introducir los datos directamente en el Editor de datos; abrir un archivo de datos previamente guardado o leer una hoja de cálculo, un archivo de datos de texto delimitado por tabuladores o un archivo de base de datos. Si selecciona Tutorial en el menú Ayuda podrá ver ejemplos en pantalla de creación y modificación de gráficos. Además, el sistema de ayuda en pantalla incluye información sobre cómo crear y modificar cualquier tipo de gráfico.

## ***Generación de gráficos***

El generador de gráficos permite crear gráficos a partir de los gráficos predefinidos de la galería o a partir de los elementos individuales (por ejemplo, ejes y barras). Puede crear un gráfico arrastrando y colocando los gráficos de la galería o los elementos básicos en el **lienzo**, que es la zona grande situada a la derecha de la lista Variables del cuadro de diálogo Generador de gráficos.

A medida que genere el gráfico, el lienzo mostrará una presentación preliminar del gráfico. Aunque la presentación preliminar utiliza etiquetas de variable definidas y niveles de medida, no muestra los datos reales. En su lugar, utiliza datos generados aleatoriamente para proporcionar un esbozo aproximado de la apariencia del gráfico.

El método preferido por los nuevos usuarios es el uso de la galería. Si desea obtener información acerca del uso de la galería, consulte [Generación de un gráfico desde la galería en p. 463](#).

### ***Para iniciar el generador de gráficos***

- ▶ Elija en los menús:
  - Gráficos
  - Generador de gráficos

Se abrirá el cuadro de diálogo Generador de gráficos.

Figura 43-1  
Cuadro de diálogo *Generador de gráficos*



### **Generación de un gráfico desde la galería**

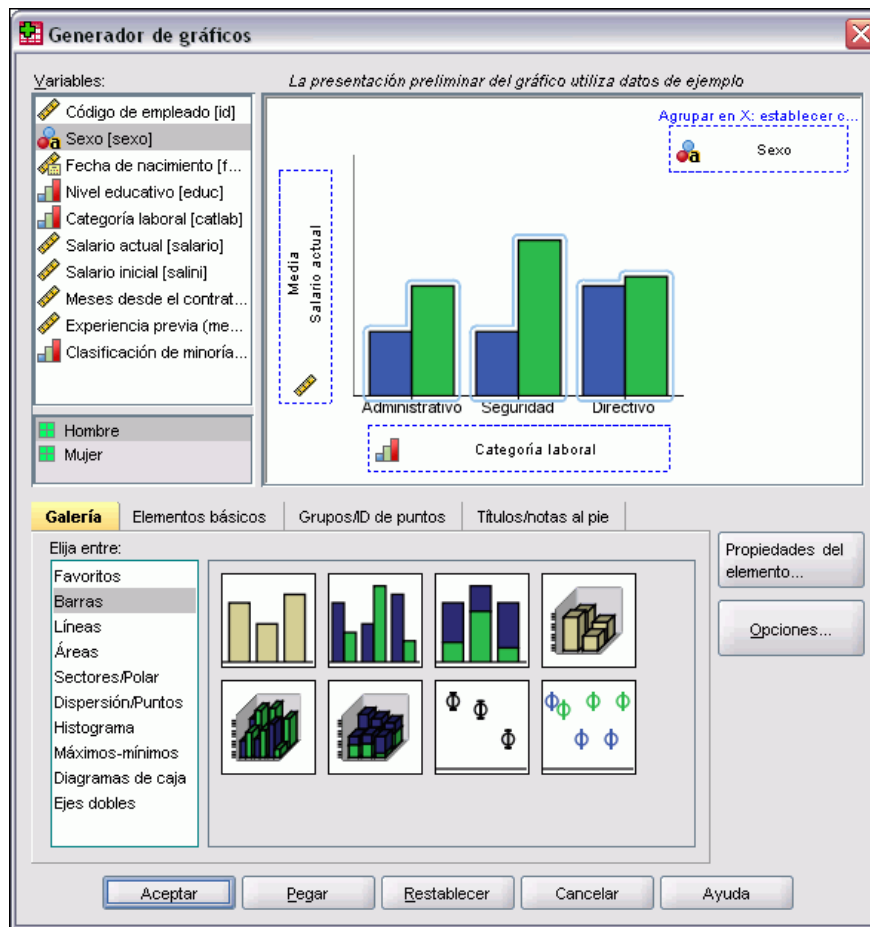
El método más sencillo para generar gráficos es utilizar la galería. A continuación, se indican los pasos generales que hay que seguir para crear un gráfico utilizando la galería.

- ▶ Si no aparece, pulse en la pestaña Galería.
- ▶ En la lista Elija entre, seleccione una categoría de gráficos. Cada categoría ofrece varios tipos.
- ▶ Arrastre la imagen del gráfico deseado al lienzo. También puede pulsar dos veces en la imagen. Si en el lienzo ya aparece un gráfico, el gráfico de la galería sustituirá al conjunto de ejes y a los elementos gráficos del gráfico.
- ▶ Arrastre variables desde la lista Variables y colóquelas en las zonas de colocación del eje y, si está disponible, en la zona de colocación de agrupamiento. Si una zona de colocación del eje ya muestra un estadístico que desea utilizar, no tendrá que arrastrar ninguna variable a la zona de colocación. Sólo deberá añadir una variable a la zona cuando el texto de la zona sea azul. Si el texto es negro, la zona ya contiene una variable o un estadístico.

*Nota:* el nivel de medida de las variables es importante. El generador de gráficos establece la configuración por defecto según el nivel de medida durante la generación de un gráfico. Además, el gráfico resultante también puede tener un aspecto distinto para los diferentes niveles de medida. Puede cambiar temporalmente el nivel de medida de una variable pulsando con el botón derecho del ratón en la variable y eligiendo una opción.

Figura 43-2

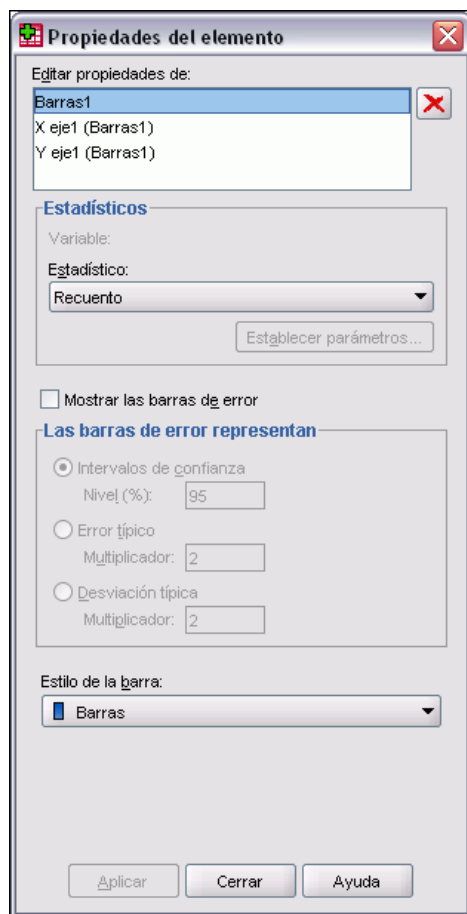
Cuadro de diálogo Generador de gráficos con las zonas de colocación rellenas



- Si necesita cambiar los estadísticos o modificar los atributos de los ejes o las leyendas (como la amplitud de la escala), pulse en Propiedades del elemento.

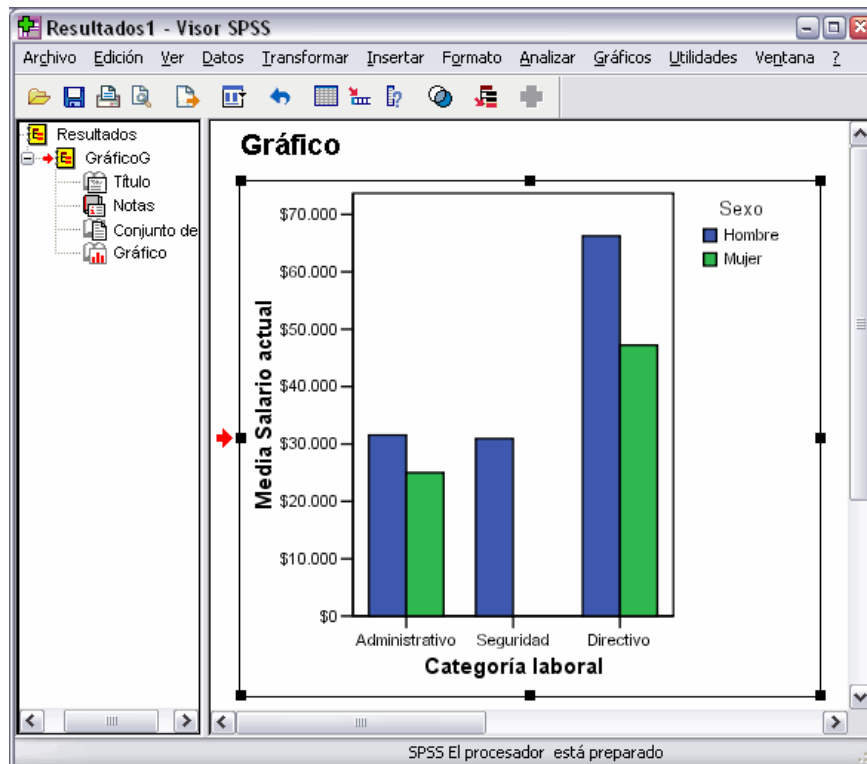


Figura 43-3  
Cuadro de diálogo Propiedades del elemento



- ▶ En la lista Editar propiedades de, seleccione el elemento que desea cambiar. (Si desea obtener información acerca de propiedades específicas, pulse en Ayuda.)
- ▶ Una vez realizados los cambios, pulse en Aplicar.
- ▶ Si necesita agregar más variables al gráfico (por ejemplo, para la agrupación o la adición de paneles), pulse en la pestaña Grupos/ID de puntos del cuadro de diálogo Generador de gráficos y seleccione una o más opciones. A continuación, arrastre las variables categóricas a las nuevas zonas de colocación que aparecen en el lienzo.
- ▶ Si desea transponer el gráfico (por ejemplo, para que las barras sean horizontales), pulse en la pestaña Elementos básicos y, a continuación, pulse en Transponer.
- ▶ Pulse en Aceptar para crear el gráfico. Aparecerá el gráfico en el Visor.

Figura 43-4  
Gráfico de barras en la ventana Visor



### Edición de gráficos

El Editor de gráficos proporciona un entorno potente y fácil de usar en el que puede personalizar sus gráficos y explorar los datos. El Editor de gráficos presenta:

- **Una interfaz de usuario sencilla e intuitiva.** Puede seleccionar y editar partes del gráfico rápidamente utilizando los menús, menús contextuales y barras de herramientas. También puede introducir texto directamente en un gráfico.
- **Un amplio conjunto de opciones de formato y de estadísticos.** Puede seleccionar entre una completa gama de opciones de estilo y opciones estadísticas.
- **Unas herramientas de exploración potentes.** Puede explorar los datos de distintas maneras, como puede ser mediante su etiquetado, reordenación y rotación. Puede cambiar los tipos de gráficos y las funciones de las variables en el gráfico. Asimismo, puede añadir curvas de distribución y líneas de ajuste, interpolación y referencia.
- **Unas plantillas flexibles para la aplicación de un aspecto y un comportamiento coherentes.** Puede crear plantillas personalizadas y utilizarlas para crear fácilmente gráficos con el aspecto y las opciones que desee. Por ejemplo, si desea contar siempre con una orientación específica para las etiquetas de los ejes, puede especificar dicha orientación en una plantilla y aplicar ésta a los demás gráficos.

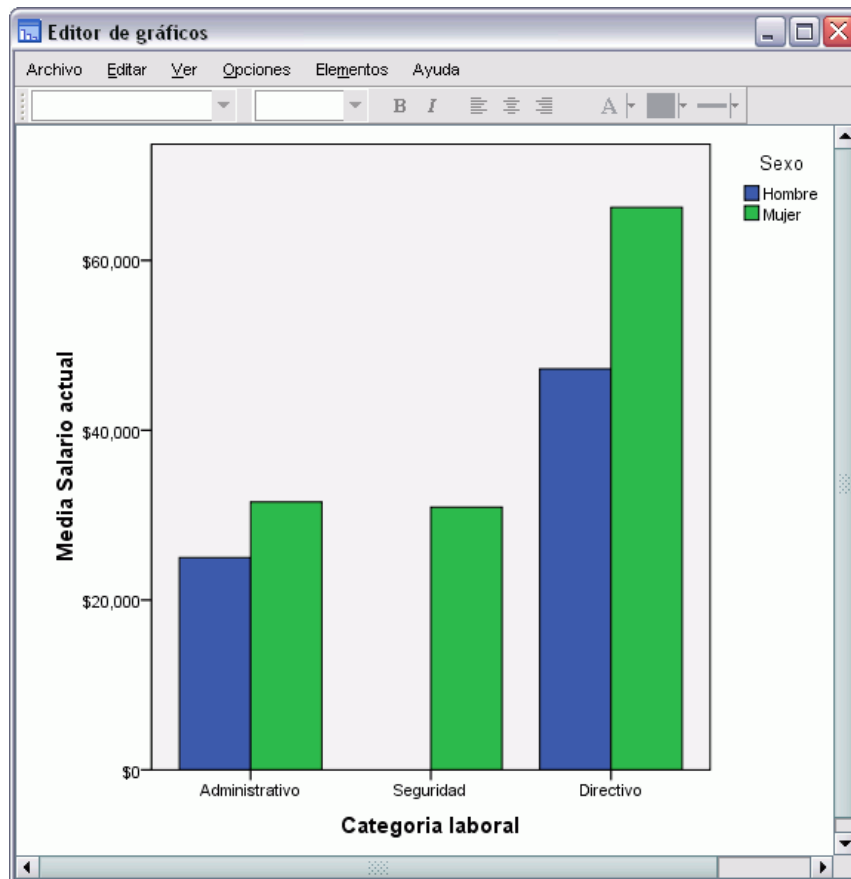
**Para ver el Editor de gráficos**

- ▶ Puede crear un gráfico en SPSS o abrir un archivo con gráficos del Visor.
- ▶ Pulse dos veces en un gráfico del Visor.

El gráfico aparecerá en el Editor de gráficos.

Figura 43-5

Gráfico mostrado en el Editor de gráficos

**Fundamentos del Editor de gráficos**

El Editor de gráficos proporciona varios métodos para la manipulación de los gráficos.

**Menús**

Muchas de las acciones que puede realizar en el Editor de gráficos se llevan a cabo con los menús, sobre todo si va a añadir un elemento al gráfico. Puede, por ejemplo, utilizar los menús para añadir una línea de ajuste a un diagrama de dispersión. Después de añadir un elemento al gráfico, utilizará con frecuencia el cuadro de diálogo Propiedades para especificar las opciones del elemento añadido.

### Cuadro de diálogo Propiedades

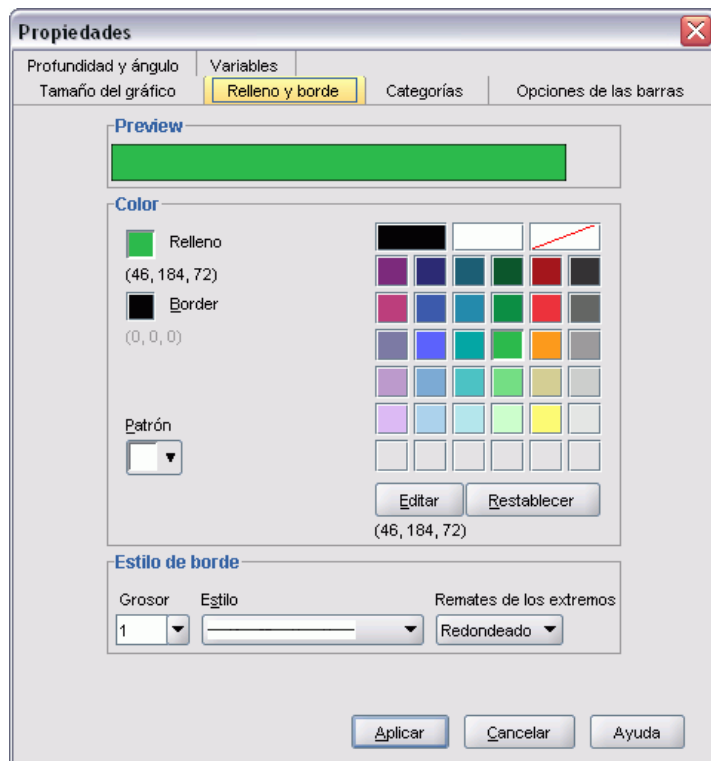
Las opciones y elementos del gráfico se encuentran en el cuadro de diálogo Propiedades.

Para ver el cuadro de diálogo Propiedades, puede:

- ▶ Pulsar dos veces en un elemento gráfico.
- o
- ▶ Seleccionar un elemento gráfico y elegir a continuación en los menús:  
Edición  
Propiedades

Además, el cuadro de diálogo Propiedades aparece automáticamente al añadir un elemento al gráfico.

Figura 43-6  
Cuadro de diálogo Propiedades, pestaña Relleno y borde



El cuadro de diálogo Propiedades incluye una serie de pestañas que le permiten definir las opciones y realizar otros cambios en el gráfico. Las pestañas que puede ver en el cuadro de diálogo Propiedades se basan en la selección actual.

Algunas pestañas incluyen una presentación preliminar que le permite hacerse una idea de cómo los cambios afectarán a la selección al aplicarlos. No obstante, el gráfico en sí no reflejará los cambios hasta que pulse en Aplicar. Puede realizar cambios en más de una pestaña antes de pulsar en Aplicar. Si desea cambiar la selección para modificar un elemento diferente del gráfico, pulse en Aplicar antes de cambiar la selección. Si no pulsa en Aplicar antes de cambiar la selección,

al pulsar en Aplicar más adelante, los cambios sólo se aplicarán en el elemento o los elementos que estén seleccionados actualmente.

Según la selección, sólo estarán disponibles algunos ajustes. La ayuda para cada una de las pestañas especifica lo que debe seleccionar para ver las pestañas. Si se han seleccionado varios elementos, sólo puede cambiar aquellos ajustes que sean comunes a todos los elementos.

### ***Barras de herramientas***

Las barras de herramientas proporcionan métodos abreviados para algunas de las funcionalidades del cuadro de diálogo Propiedades. Por ejemplo, en vez de utilizar la pestaña Texto del cuadro de diálogo Propiedades, puede utilizar la barra de herramientas Editar para cambiar la fuente y el estilo del texto.

### ***Almacenamiento de cambios***

Las modificaciones realizadas al gráfico se guardan al cerrar el Editor de gráficos y el gráfico modificado se muestra en el Visor.

## ***Opciones de definición de gráfico***

Cuando define un gráfico en el generador de gráficos, puede añadir títulos y cambiar opciones para la creación del gráfico.

### ***Adición y edición de títulos y notas al pie***

Puede añadir títulos y notas al pie a un gráfico para que a las personas que lo consultan les sea más fácil interpretarlo. El generador de gráficos también muestra automáticamente información sobre las barras de error en las notas al pie.

#### ***Para añadir títulos y notas al pie***

- ▶ Pulse en la pestaña Títulos/notas al pie.
- ▶ Seleccione uno o más títulos y notas al pie. El lienzo muestra cierto texto para indicar que se han añadido al gráfico.
- ▶ Utilice el cuadro de diálogo Propiedades del elemento para editar el texto del título o de la nota al pie.

#### ***Para eliminar un título o una nota al pie***

- ▶ Pulse en la pestaña Títulos/notas al pie.
- ▶ Anule la selección del título o de la nota al pie que desea eliminar.

**Para editar el texto del título o de la nota al pie**

Al añadir títulos y notas al pie, no es posible editar directamente el texto asociado en el gráfico. Al igual que ocurre con otros elementos del generador de gráficos, la edición se realiza mediante el cuadro de diálogo Propiedades del elemento.

- ▶ Pulse en Propiedades del elemento si no aparece el cuadro de diálogo Propiedades del elemento.
- ▶ En la lista Editar propiedades de, seleccione un título, subtítulo o nota al pie (por ejemplo, Título 1).
- ▶ En el cuadro Contenido, escriba el texto asociado con el título, subtítulo o nota al pie.
- ▶ Pulse en Aplicar.

**Para establecer las opciones generales**

El generador de gráficos ofrece opciones generales para el gráfico. Son opciones que se aplican al gráfico general en lugar de a un elemento específico del gráfico. Entre las opciones generales se incluyen el tratamiento de los valores perdidos, las plantillas, el tamaño del gráfico y el ajuste de los paneles.

- ▶ Pulse en Opciones.
- ▶ Modifique las opciones generales. A continuación se muestran detalles acerca de este tema.
- ▶ Pulse en Aplicar.

**Valores definidos como perdidos por el usuario**

**Variables de segmentación.** Si hay valores perdidos para las variables que se utilizan para definir categorías o subgrupos, seleccione Incluir para que la categoría o categorías de los valores perdidos definidos por el usuario (valores identificados como perdidos por el usuario) se incluyan en el gráfico. Estas categorías “perdidos” también actúan como variables de segmentación para calcular el estadístico. La categoría o categorías “perdidos” aparecen en el eje de categorías o en la leyenda, añadiendo, por ejemplo, una barra adicional o un sector a un gráfico de sectores. Si no hay ningún valor perdido, no aparecerán las categorías “perdidos”.

Si selecciona esta opción y quiere suprimir la presentación después de que se dibuje el gráfico, abra el gráfico en el Editor de gráficos y, a continuación, seleccione Propiedades en el menú Edición. Utilice la pestaña Categorías para mover las categorías que desea suprimir a la lista Excluidos. Sin embargo, tenga en cuenta que los estadísticos no se vuelven a calcular si oculta las categorías “perdidos”. Así, algo parecido a un porcentaje estadístico tendrá en cuenta las categorías “perdidos”.

*Nota:* este control no afecta a los valores perdidos del sistema. Siempre están excluidos del gráfico.

**Estadísticos de resumen y valores de casos.** Puede seleccionar una de las siguientes alternativas para la exclusión de casos con valores perdidos:

- **Excluir según lista para obtener una base de casos coherente para el gráfico.** Si alguna de las variables del gráfico tiene un valor perdido para un determinado caso, se excluirá el caso completo del gráfico.
- **Excluir por variable para maximizar el uso de los datos.** Si una variable seleccionada tiene algún valor perdido, los casos que tengan estos valores perdidos se excluirán al analizar dicha variable.

Para ver la diferencia existente entre la exclusión por lista y la exclusión por variables de los valores perdidos, observe las siguientes figuras, que muestran un gráfico de barras para cada una de las dos opciones.

Figura 43-7

*Exclusión por lista de los valores perdidos*

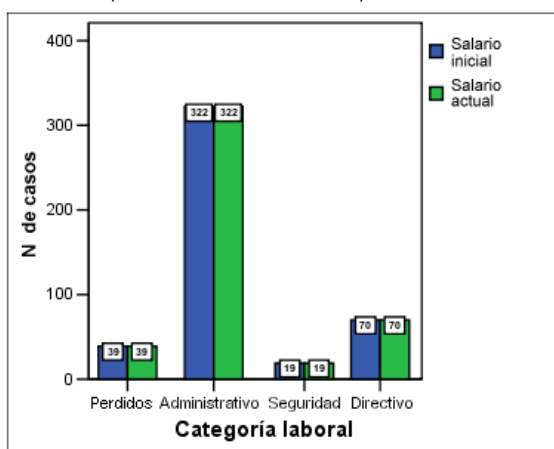
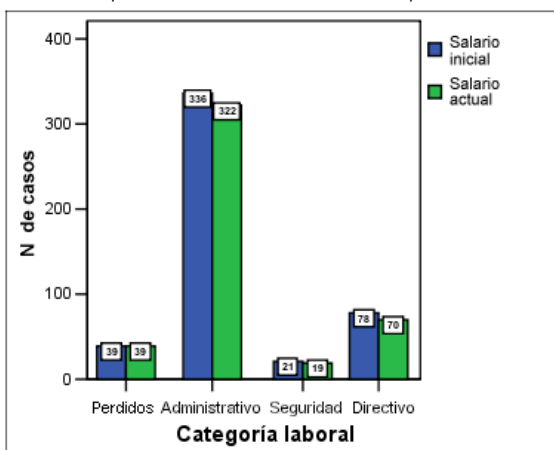


Figura 43-8

*Exclusión por variables de los valores perdidos*



Estos gráficos se crearon a partir de una versión del archivo *Employee data.sav* que se editó para que tuviese algunos valores perdidos por el sistema (en blanco) para las variables correspondientes al salario actual y a la categoría laboral. En otros casos, se introdujo el valor 0 y se definió como

perdido. En ambos gráficos, se ha seleccionado la opción Mostrar los grupos definidos por los valores perdidos, que añade la categoría *Perdidos* al resto de las categorías laborales mostradas. En cada gráfico, los valores de la función de resumen, *Número de casos*, se muestran en las etiquetas de las barras.

En ambos gráficos, 26 casos tienen un valor perdido del sistema para la categoría laboral y 13 casos tienen el valor perdido definido por el usuario (0). En el gráfico por lista, el número de casos es el mismo para ambas variables en cada agrupación de barras ya que siempre que había un valor perdido se excluyó el caso de todas las variables. En el gráfico por variables, el número de casos no perdidos de cada variable de una categoría se representa sin tener en cuenta los valores perdidos en las otras variables.

### **Plantillas**

La plantilla de gráficos permite aplicar los atributos de un gráfico a otro. Cuando abre un gráfico en el Editor de gráficos, puede guardarlo como una plantilla. A continuación, puede aplicar esa plantilla especificándolo durante la creación, o más tarde, aplicándola en el Editor de gráficos.

**Plantilla por defecto.** Ésta es la plantilla especificada por Opciones. Para acceder a estas opciones seleccione Opciones en el menú Edición del Editor de datos y, a continuación, pulse en la pestaña Gráficos. La plantilla por defecto se aplica primero, lo que significa que las otras plantillas pueden omitirla.

**Archivos de plantilla.** Pulse en Añadir para especificar una o más plantillas con el cuadro de diálogo estándar de selección de archivos. Se aplican en orden de aparición. Así, las plantillas situadas al final de la lista pueden omitir las plantillas del inicio de la lista.

### **Paneles y tamaño de los gráficos**

**Tamaño del gráfico.** Especifique un porcentaje mayor que 100 para ampliar el gráfico o menor que 100 para reducirlo. El porcentaje es relativo al tamaño del gráfico por defecto.

**Paneles.** Cuando hay numerosas columnas de panel, seleccione Ajustar paneles para que los paneles se ajusten en varias filas en vez de obligar a que encajen en una determinada fila. A menos que esta opción esté seleccionada, los paneles están ajustados para ajustarse a una fila.



## Utilidades

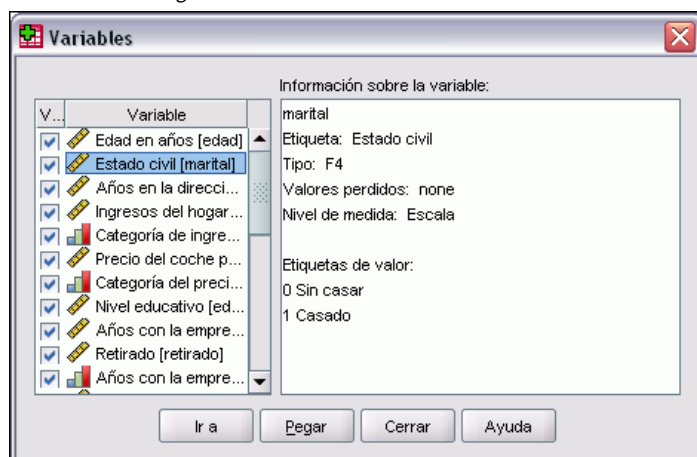
En este capítulo se describen las funciones del menú Utilidades así como la posibilidad de reordenar las listas de variables de destino.

### Información sobre la variable

El cuadro de diálogo Variables muestra información sobre la definición de la variable seleccionada actualmente, incluyendo:

- Etiqueta de variable
- Formato de datos
- Valores definidos por el usuario como perdidos
- Etiquetas de valor
- Nivel de medida

Figura 44-1  
Cuadro de diálogo Variables



**Visible.** La columna Visible de la lista de variables indica si la variable está visible actualmente en el Editor de datos y en las listas de variables del cuadro de diálogo. La visibilidad está controlada por conjuntos de variables. Si desea obtener más información, consulte [Conjuntos de variables](#) en p. 474.

**Ir a.** Se dirige a la variable seleccionada en el Editor de datos.

**Pegar.** Pega las variables seleccionadas en la posición del cursor en la ventana de sintaxis designada.

Para modificar la definición de una variable, utilice la Vista de variables en el Editor de datos.

***Para obtener información sobre la variable***

- ▶ Elija en los menús:
  - Utilidades
  - Variables...
- ▶ Seleccione la variable cuya información de definición desee mostrar.

## ***Comentarios del archivo de datos***

Puede incluir comentarios descriptivos en un archivo de datos. Para los archivos de datos con formato SPSS, estos comentarios se guardan con el archivo de datos.

***Para añadir, modificar, eliminar o visualizar los comentarios del archivo de datos***

- ▶ Elija en los menús:
  - Utilidades
  - Comentarios del archivo de datos
- ▶ Para mostrar los comentarios en el Visor, seleccione Mostrar comentarios en resultados.

Los comentarios admiten cualquier longitud, aunque están limitados a 80 bytes (por regla general, 80 caracteres en idiomas de un solo byte) por línea; las líneas se dividen automáticamente en 80 caracteres. Los comentarios se muestran en la misma fuente que los resultados de texto para reflejar de forma precisa el modo en que aparecen al mostrarlos en el Visor.

Se añade de forma automática una anotación de fecha (la fecha actual entre paréntesis) al final de la lista de comentarios siempre que se añaden o modifican los comentarios. Esto puede dar lugar a cierta ambigüedad por lo que respecta a las fechas asociadas a los comentarios si modifica un comentario existente o introduce un comentario nuevo entre los comentarios existentes.

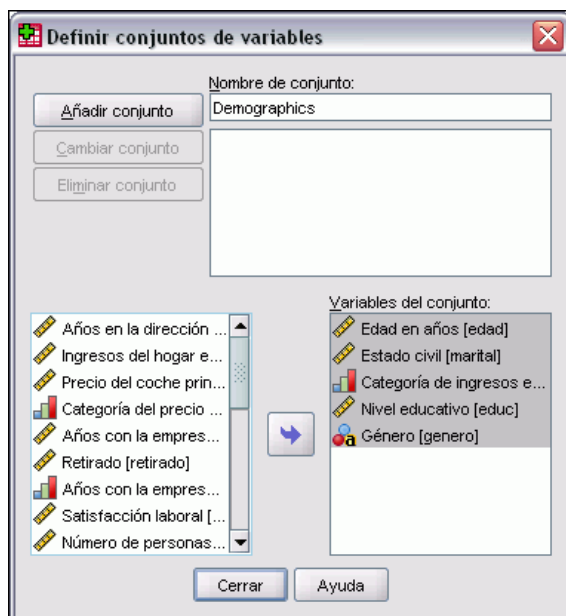
## ***Conjuntos de variables***

Puede restringir las variables que aparecen en el Editor de datos y en las listas de variables de los cuadros de diálogo definiendo y utilizando conjuntos de variables. Es especialmente útil en archivos de datos con un amplio número de variables. Los conjuntos de variables pequeños hacen que la búsqueda y la selección de variables para los análisis sea más fácil.

## ***Definir conjuntos de variables***

El cuadro de diálogo Definir conjuntos de variables crea subconjuntos de variables que aparecen en el Editor de datos y en las listas de variables de los cuadros de diálogo. Los conjuntos de variables definidos se guardan en los archivos de datos con formato SPSS.

Figura 44-2  
Cuadro de diálogo Definir conjuntos de variables



**Nombre del conjunto.** Los nombres de los conjuntos pueden tener hasta 64 bytes. Puede utilizarse cualquier carácter, incluyendo los espacios.

**Variables del conjunto.** El conjunto puede estar compuesto de cualquier combinación de variables numéricas y de cadena. El orden de las variables del conjunto no tiene ningún efecto en el orden de presentación de las variables en el Editor de datos o en las listas de variables de los cuadros de diálogo. Una variable puede pertenecer a varios conjuntos.

#### ***Para definir conjuntos de variables***

- ▶ Elija en los menús:  
Utilidades  
Definir conjuntos de variables...
- ▶ Seleccione las variables que desee incluir en el conjunto.
- ▶ Escriba el nombre del conjunto (hasta 64 bytes).
- ▶ Pulse en Añadir conjunto.

### ***Utilizar conjuntos de variables***

Utilice conjuntos de variables para restringir las variables que aparecen en el Editor de datos y en las listas de variables de los cuadros de diálogo a las variables de los conjuntos seleccionados (marcados).

Figura 44-3  
Cuadro de diálogo Utilizar conjuntos de variables



- El conjunto de variables que aparece en el Editor de datos y en las listas de variables de los cuadros de diálogo es la unión de todos los conjuntos seleccionados.
- Se puede incluir una variable en varios conjuntos seleccionados.
- El orden de las variables en los conjuntos seleccionados y el orden de los conjuntos seleccionados no tienen ningún efecto en el orden de presentación de las variables en el Editor de datos o en las listas de variables de los cuadros de diálogo.
- Aunque los conjuntos de variables definidos se guardan en los archivos de datos con formato SPSS, la lista de conjuntos seleccionados se restablece en los conjuntos por defecto preincorporados cada vez que se abre el archivo de datos.

La lista de conjuntos de variables disponibles incluye todos los conjuntos de variables definidos para el conjunto de datos activo, más dos conjuntos preincorporados:

- **ALLVARIABLES.** Este conjunto contiene todas las variables del archivo de datos, incluidas las nuevas variables creadas durante una sesión.
- **NEWVARIABLES.** Este conjunto contiene sólo las nuevas variables creadas durante la sesión.

*Nota:* Incluso si guarda el archivo de datos tras crear las nuevas variables, estas nuevas variables seguirán incluidas en el conjunto *NEWVARIABLES* hasta que cierre y vuelva a abrir el archivo de datos.

Es necesario seleccionar como mínimo un conjunto de datos. Si se selecciona *ALLVARIABLES*, ninguno de los conjuntos seleccionados tendrá ningún efecto visible, ya que este conjunto incluye todas las variables.

#### **Para seleccionar los conjuntos de variables que desea mostrar**

- Elija en los menús:  
Utilidades  
Usar conjuntos de variables...

- ▶ Seleccione los conjuntos de variables definidos que contengan las variables que desee incluir en el Editor de datos y en las listas de variables de los cuadros de diálogo.

***Para ver todas las variables***

- ▶ Elija en los menús:
  - Utilidades
  - Mostrar todas las variables

## ***Reordenación de listas de variables de destino***

Las variables aparecen en las listas de destino de un cuadro de diálogo en el orden en que son seleccionadas en la lista de origen. Si desea cambiar el orden de las variables en la lista de destino pero no quiere tener que eliminar la selección de todas las variables y volver a seleccionarlás en el nuevo orden, puede mover las variables hacia arriba y hacia abajo en la lista de destino con la tecla Ctrl (Macintosh: tecla Comando) y las teclas de dirección hacia arriba y hacia abajo. Puede mover múltiples variables simultáneamente si son contiguas (es decir, si están agrupadas unas junto a otras). No es posible mover grupos de variables no contiguas.

# *Opciones*

Controla una gran variedad de opciones, entre ellas:

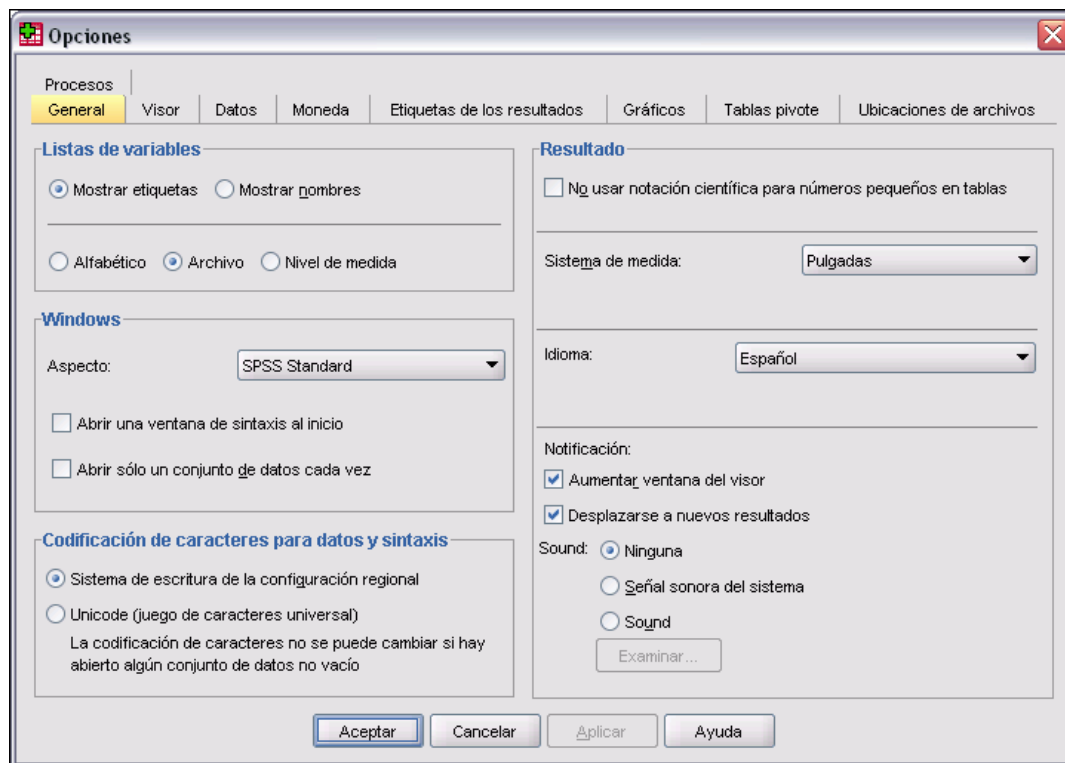
- El diario de la sesión, que guarda un registro de todos los comandos ejecutados en cada sesión
- El orden en que aparecen las variables en las listas de origen de los cuadros de diálogo
- Los elementos mostrados y ocultos en los nuevos resultados
- El Aspecto de tabla para las nuevas tablas pivote
- Formatos de moneda personalizados

## ***Para modificar la configuración de las opciones***

- ▶ Elija en los menús:
  - Edición
  - Opciones...
- ▶ Pulse en las pestañas de las selecciones que desee cambiar.
- ▶ Cambie las selecciones.
- ▶ Pulse en Aceptar o Aplicar.

## General: Opciones

Figura 45-1  
Cuadro de diálogo Opciones: pestaña General



### Listas de variables

Estos ajustes controlan la presentación de las variables en las listas de los cuadros de diálogo. Es posible mostrar los nombres o las etiquetas de las variables. Estos nombres o etiquetas pueden presentarse por orden alfabético o por orden de archivo, o agruparse por nivel de medida. El orden de presentación afecta sólo a las listas de variables de origen. Las listas de variables de destino siempre reflejan el orden en el que las variables han sido seleccionadas.

### Ventanas

**Aspecto.** Controla el aspecto básico de las ventanas y cuadros de diálogo.

**Abrir una ventana de sintaxis al inicio.** Las ventanas de sintaxis son ventanas de archivos de texto que sirven para introducir, editar y ejecutar comandos. Si utiliza con frecuencia la sintaxis de comandos, seleccione esta opción para abrir automáticamente una ventana de sintaxis al principio de cada sesión. Esta opción es especialmente útil para los usuarios avanzados que prefieren trabajar con la sintaxis de comandos en vez de con los cuadros de diálogo.

**Abrir sólo un conjunto de datos cada vez.** Cierra el origen de datos abierto actualmente cada vez que se abre uno nuevo a través de los menús y cuadros de diálogo. Por defecto, cada vez que se abre un nuevo origen de datos a través de los menús y cuadros de diálogo, se abre una nueva

ventana Editor de datos y los orígenes de datos que ya están abiertos en ventanas Editor de datos permanecen abiertos y disponibles durante la sesión hasta que se cierran de forma explícita.

Cuando se selecciona esta opción, tiene efecto de forma inmediata pero no se cierran los conjuntos de datos que ya estaban abiertos antes de cambiarla. Este ajuste no tiene efecto sobre los orígenes de datos abiertos mediante la sintaxis de comandos, que se basa en comandos `DATASET` para controlar varios conjuntos de datos. Si desea obtener más información, consulte [Trabajo con varios orígenes de datos](#) en Capítulo 6 en p. 102.

### ***Codificación de caracteres para archivos de datos y archivos de sintaxis***

Controla el comportamiento por defecto para determinar la codificación para leer y escribir archivos de datos y archivos de sintaxis. Esta opción sólo se puede seleccionar cuando no hay ningún origen de datos abierto y permanece en vigor durante las sesiones posteriores hasta que se cambie de forma explícita.

- **Sistema de escritura de la configuración regional.** Utiliza la configuración regional actual para determinar la codificación para leer y escribir archivos. A esto se le denomina **modo de página de código**.
- **Unicode (conjunto de caracteres universal).** Utiliza la codificación Unicode (UTF-8) para leer y escribir archivos. A esto se le denomina **modo Unicode**.

Existen varias implicaciones importantes relacionadas con el modo y los archivos Unicode:

- Los archivos de datos con formato SPSS y los archivos de sintaxis, guardados con una codificación Unicode, no se deben utilizar en las versiones de SPSS anteriores a la 16.0. En el caso de los archivos de sintaxis, se puede especificar la codificación al guardar el archivo. Si se trata de archivos de datos, deberá abrir el archivo en el modo de página de código y volver a guardarlo si desea leerlo con versiones anteriores.
- Cuando los archivos de datos de página de código se leen en el modo Unicode, el ancho definido de todas las variables de cadena se triplica. Para establecer de forma automática la longitud de cada variable de cadena en el valor más largo observado para dicha variable, seleccione Minimizar longitudes de cadena en función de los valores observados en el cuadro de diálogo Abrir datos.

### ***Resultados***

**No usar notación científica para números pequeños en tablas.** Suprime la presentación de la notación científica para valores decimales pequeños en el resultado. Los valores decimales muy pequeños se muestran como 0 (o 0,000).

**Sistema de medida.** El sistema de medida utilizado (puntos, pulgadas o centímetros) para especificar atributos tales como los márgenes de casillas de las tablas pivote, los anchos de casilla y el espacio entre las tablas para la impresión.

**Idioma.** Determina el idioma que se utiliza para los resultados. No se aplica a resultados en formato de sólo texto. Esta lista de idiomas disponibles depende de los archivos de idioma que estén instalados actualmente.



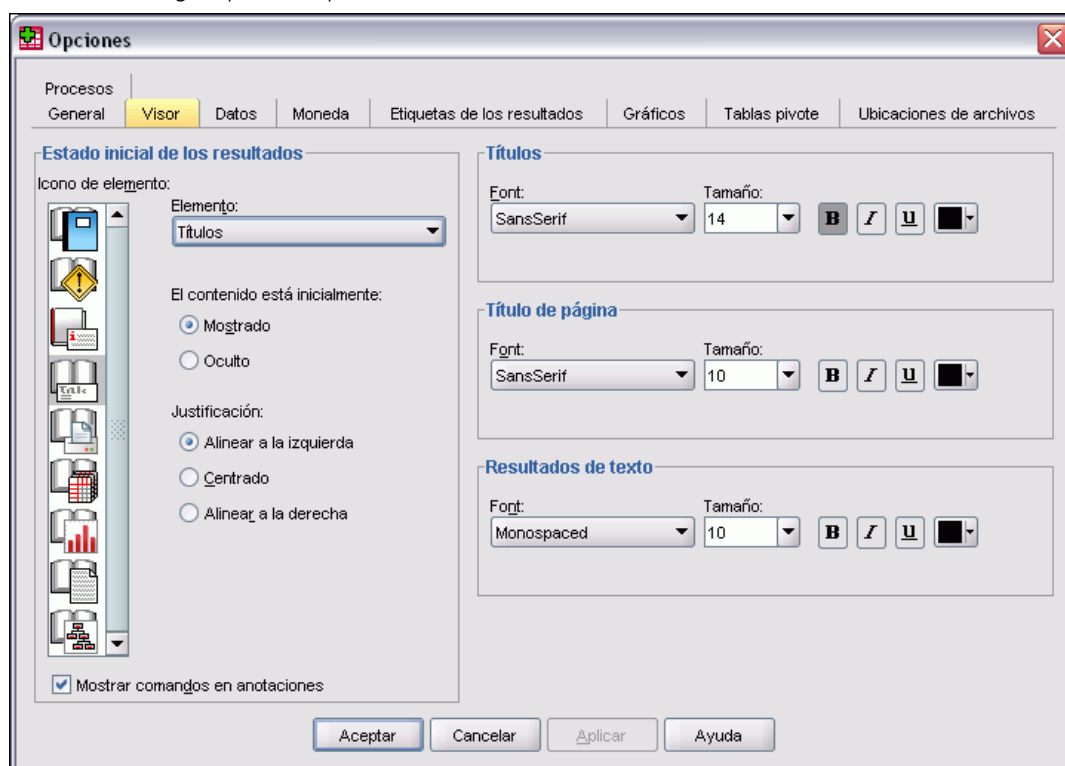
*Nota:* es posible que, en los procesos personalizados que se basan en cadenas de texto de un idioma específico, estas cadenas no se ejecuten correctamente cuando se cambie el idioma de los resultados. Si desea obtener más información, consulte [Procesos: Opciones](#) en p. 493.

**Notificación.** Determina cómo debe notificar el programa al usuario que ha finalizado la ejecución de un procedimiento y que los resultados están disponibles en el Visor.

## Opciones del Visor

Las opciones de salida de presentación de resultados en el Visor sólo afectan a los resultados obtenidos tras el cambio de la configuración. A los resultados que ya se muestran en el Visor no les afectan los cambios en la configuración.

Figura 45-2  
Cuadro de diálogo Opciones: pestaña Visor



**Estado inicial de los resultados.** Controla los elementos que se muestran y se ocultan automáticamente cada vez que se ejecuta un procedimiento, además de la alineación inicial de los elementos. Puede controlar la presentación de los siguientes elementos: registro, advertencias, notas, títulos, tablas pivote, gráficos, diagramas de árbol y resultados de texto. También se puede activar o desactivar la muestra de los comandos en el registro. Si lo desea, puede copiar la sintaxis de comandos del registro y guardarla en un archivo de sintaxis.

*Nota:* todos los elementos de resultados aparecen alineados a la izquierda en el Visor. Las selecciones de justificación sólo afectarán a la alineación de los resultados impresos. Los elementos con alineación centrada y a la derecha se identifican por un pequeño símbolo situado encima y a la izquierda del elemento.

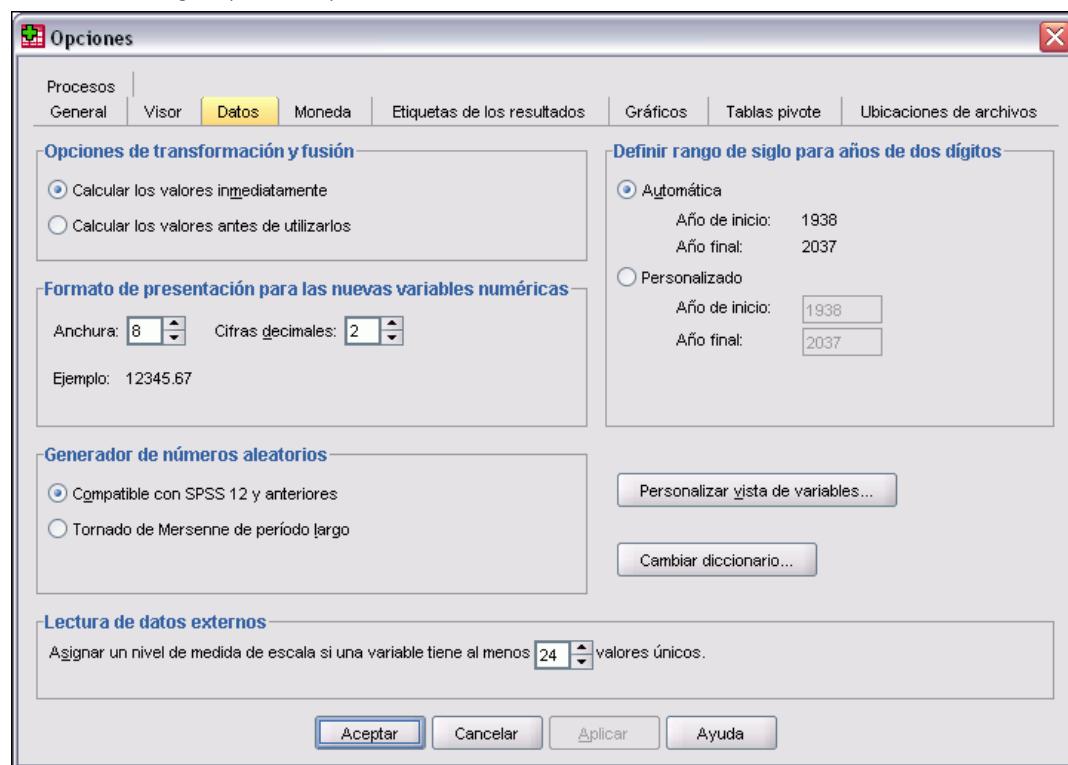
**Título.** Controla el estilo, el tamaño y el color de la fuente de los nuevos títulos de resultados.

**Título de página** Controla el estilo, tamaño y color de la fuente para los nuevos títulos de página y los títulos de página generados por la sintaxis de comandos TITLE y SUBTITLE o creados por Nuevo título de página en el menú Insertar.

**Resultados de texto** Fuente utilizada para los resultados de texto. Los resultados de texto se han diseñado para utilizarlos con fuentes de paso fijo. Si selecciona una fuente proporcional, los resultados tabulares no se alinearán adecuadamente.

## Datos: Opciones

Figura 45-3  
Cuadro de diálogo Opciones: pestaña Datos



**Opciones de transformación y fusión.** Cada vez que SPSS ejecuta un comando, lee el archivo de datos. Algunas transformaciones de datos (tales como Calcular y Recodificar) no requieren una lectura diferente de los datos; esto permite postergar su ejecución hasta que el programa lea los datos para ejecutar otro comando, como puede ser un procedimiento estadístico.

- Para archivos de datos grandes, donde puede llevar tiempo la lectura de los datos, puede seleccionar **Calcular los valores antes de usarlos** para retrasar la ejecución y ahorrar tiempo de procesamiento. Cuando se selecciona esta opción, los resultados de las transformaciones realizadas mediante los cuadros de diálogo, como **Calcular variable** no se mostrarán inmediatamente en el Editor de datos; las nuevas variables creadas por transformaciones se mostrarán sin valores de datos y los valores de datos del Editor de datos no se podrán modificar mientras haya transformaciones pendientes. Cualquier comando que lea los datos, como el procedimiento estadístico o de creación de gráficos, ejecutará las transformaciones pendientes y actualizará los datos que se muestran en el Editor de datos. Si lo desea, puede utilizar **Ejecutar transformaciones pendientes** en el menú **Transformar**.
- Con el valor por defecto **Calcular los valores inmediatamente**, cuando pega la sintaxis de comandos de los cuadros de diálogo, se copiará un comando `EXECUTE` después de cada comando de transformación. Si desea obtener más información, consulte [Varios comandos Ejecutar](#) en Capítulo 12 en p. 266.

**Formato de presentación para las nuevas variables numéricas.** Controla la presentación por defecto del ancho y el número de posiciones decimales de las nuevas variables numéricas. No existe formato de presentación por defecto para las nuevas variables de cadena. Si un valor es demasiado largo para el formato de presentación especificado, primero se redondean las posiciones decimales y después los valores se convierten a notación científica. Los formatos de presentación no afectan a los valores de datos internos. Por ejemplo, el valor 123456,78 se puede redondear a 123457 para la presentación, pero se utilizará el valor original sin redondear en cualquier cálculo.

**Definir rango de siglo para años de dos dígitos.** Define el rango de años para las variables con formato de fecha introducidas o mostradas con un año de dos dígitos (por ejemplo, 10/28/86, 29-OCT-87). La opción de rango automático se basa en el año actual; es decir, comienza 69 años antes del actual y finaliza 30 años después (sumando el año en curso hace un total de 100 años). En el rango personalizado, el año final se establece de forma automática en función del valor introducido en el año inicial.

**Generador de números aleatorios.** Hay dos generadores de números aleatorios disponibles:

- **Compatible con la versión 12.** El generador de números aleatorios utilizado en la versión 12 y versiones anteriores. Utilice este generador de números aleatorios si necesita reproducir los resultados aleatorizados generados por versiones previas basadas en una semilla de aleatorización especificada.
- **Tornado de Mersenne.** Un generador de números aleatorios nuevo que es más fiable en los procesos de simulación. Utilice este generador de números aleatorios si no es necesario reproducir resultados aleatorizados correspondientes a la versión 12 o anteriores.

**Lectura de datos externos** Para los datos leídos de formatos de archivo externos y los archivos de datos creados en versiones anteriores de SPSS (anteriores a la versión 8.0), es posible especificar el número mínimo de valores de datos que deben considerarse en una variable numérica para clasificarla como variable de escala o nominal. Las variables con un número de valores únicos inferior al especificado se clasificarán como nominales.

**Personalizar Vista de variables.** Controla la presentación por defecto y el orden de los atributos en la Vista de variables. Si desea obtener más información, consulte [Cambio de la Vista de variables por defecto](#) en p. 484.

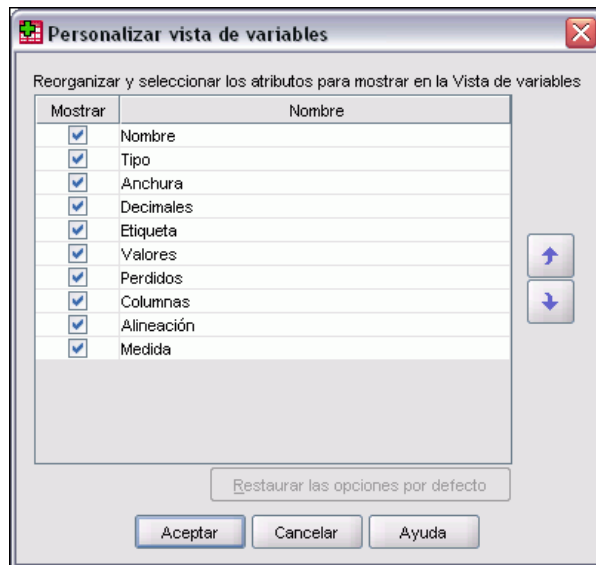
**Cambiar diccionario.** Controla la versión de idioma del diccionario que se utiliza para realizar la revisión ortográfica de los elementos de la Vista de variables. Si desea obtener más información, consulte [Revisión ortográfica de las etiquetas de los valores y las variables](#) en Capítulo 5 en p. 94.

## ***Cambio de la Vista de variables por defecto***

Puede utilizar la opción Personalizar Vista de variables para controlar qué atributos se muestran por defecto en la Vista de variables (por ejemplo, nombre, tipo, etiqueta) y el orden en el que aparecen.

Pulse en Personalizar Vista de variables.

Figura 45-4  
*Personalizar Vista de variables (por defecto).*



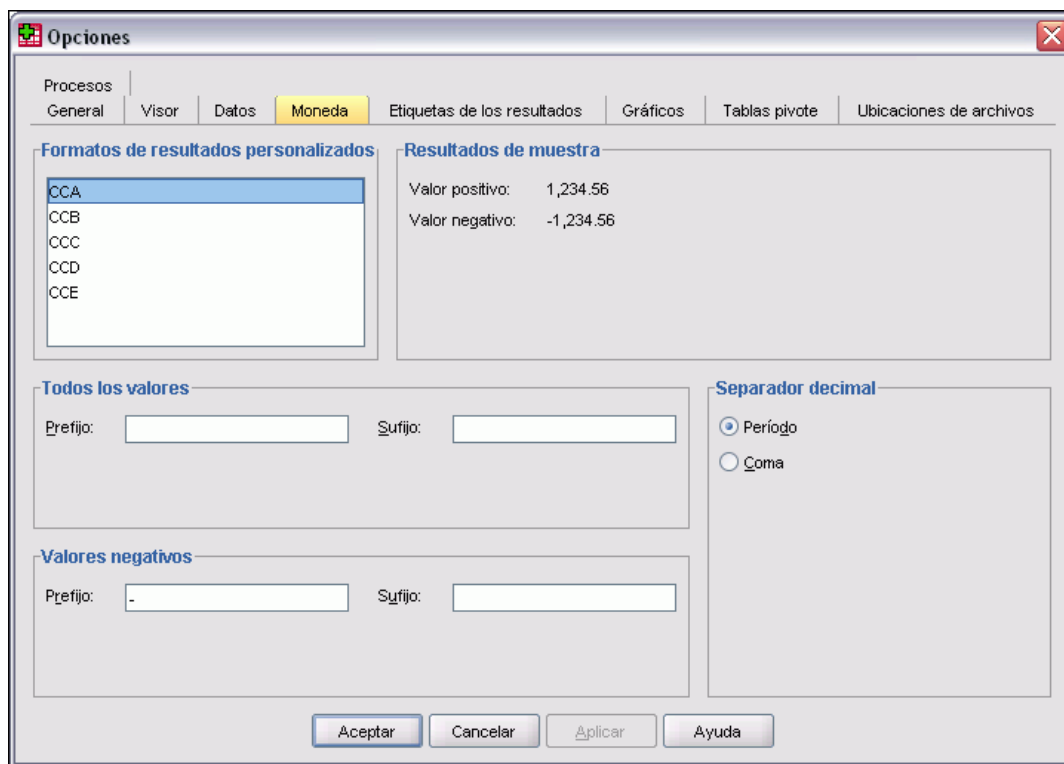
- ▶ Seleccione (marque) los atributos de variable que desea mostrar.
- ▶ Utilice los botones de dirección hacia arriba y hacia abajo para cambiar el orden de la presentación de los atributos.

## ***Moneda: Opciones***

Puede crear hasta cinco formatos de presentación de moneda personalizados que pueden incluir caracteres de prefijo y sufijo especiales además de un tratamiento especial para los valores negativos.

Los nombres de los cinco formatos de moneda personalizados son MPA, MPB, MPC, CCD y MPE. No se pueden cambiar los nombres de los formatos ni añadir otros nuevos. Para modificar un formato de moneda personalizado, seleccione el nombre del formato de la lista de origen y realice los cambios que desee.

Figura 45-5  
Cuadro de diálogo Opciones: pestaña Moneda



Los prefijos, los sufijos y los separadores decimales definidos para los formatos monetarios personalizados sólo afectan a la presentación en la pantalla. No es posible introducir valores en el Editor de datos utilizando caracteres de moneda personalizados.

### ***Para crear formatos de moneda personalizados***

- ▶ Pulse en la pestaña Moneda.
- ▶ Seleccione uno de los formatos de moneda de la lista (MPA, MPB, MPC, MPD y MPE).
- ▶ Introduzca el prefijo, el sufijo y los valores indicadores de decimales.
- ▶ Pulse en Aceptar o Aplicar.

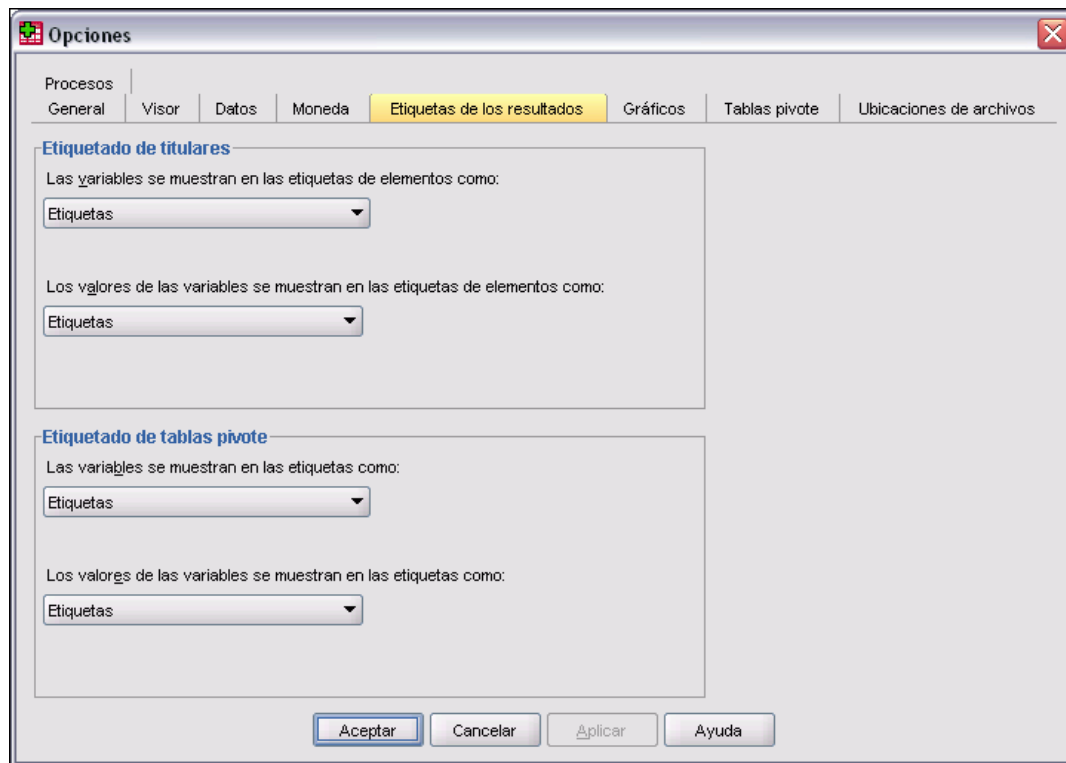
### ***Etiqueta de los resultados: Opciones***

Las opciones de Etiquetas de los resultados controlan la presentación de la información sobre variables y valores de datos en las tablas pivote y en los titulares. Puede visualizar nombres de variable, etiquetas de variable definidas y valores de datos reales, etiquetas de valor definidas o una combinación de estas opciones.

Etiquetas descriptivas de variable y de valor (Vista de variables del Editor de datos, columnas de *Etiqueta* y *Valores*) a menudo facilitan la interpretación de los resultados. Sin embargo, las etiquetas largas pueden crear dificultades en algunas tablas.

Figura 45-6

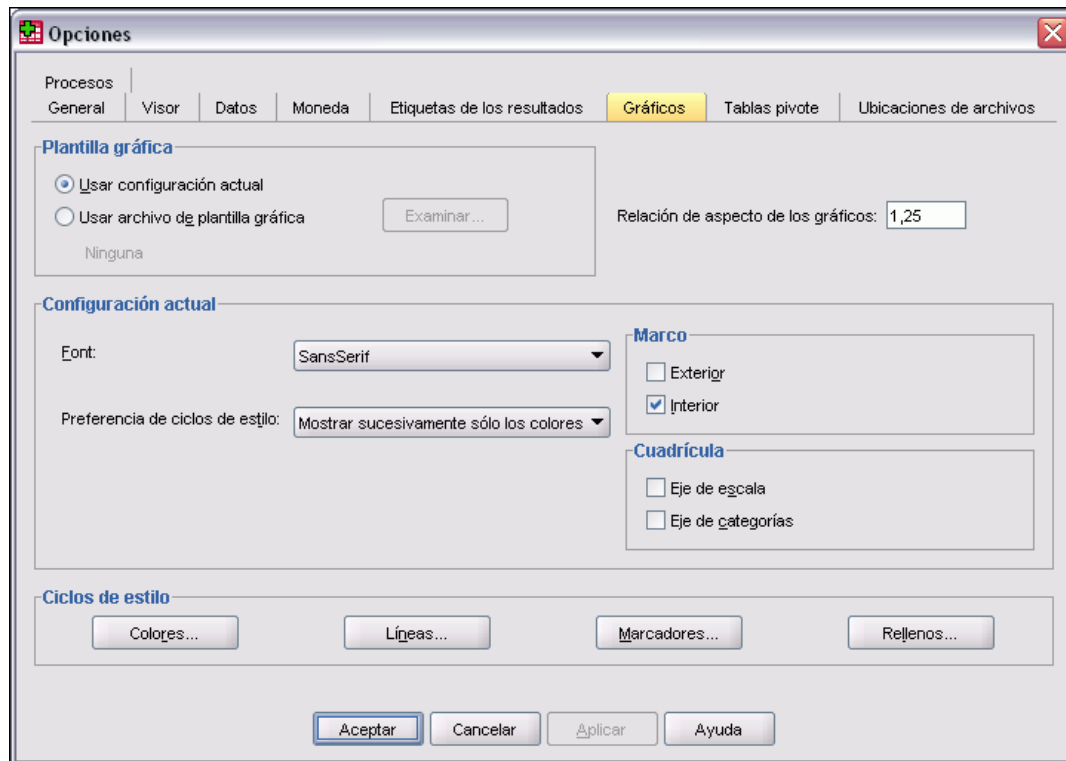
Cuadro de diálogo opciones: pestaña *Etiquetas de los resultados*



Las opciones de etiquetas de los resultados afectan sólo a los nuevos resultados que se producen después de cambiar las selecciones. A los resultados que ya se muestran en el Visor no les afectan los cambios en la configuración. Esta configuración sólo afecta a los resultados de las tablas pivote. Los resultados de texto no se verán afectados por estas selecciones.

## Opciones de gráfico

Figura 45-7  
Cuadro de diálogo Opciones: pestaña Gráficos



**Plantilla gráfica.** Los gráficos nuevos pueden utilizar tanto las opciones seleccionadas aquí, como las opciones de un archivo de plantilla gráfica. Pulse en Examinar para seleccionar un archivo de plantilla gráfica. Para crear un archivo de plantilla gráfica, cree un gráfico con los atributos que desee y guárdelo como una plantilla (seleccione Guardar plantilla gráfica en el menú Archivo).

**Relación de aspecto de los gráficos.** La relación ancho-alto del marco exterior de los nuevos gráficos. Puede especificar una relación ancho-alto entre los valores 0,1 y 10,0. Los valores inferiores a 1 generan gráficos que son más altos que anchos. Los valores mayores que 1 producen gráficos que son más anchos que altos. Un valor de 1 produce un gráfico cuadrado. Una vez creado un gráfico, no es posible cambiar su relación de aspecto.

**Especificaciones actuales.** Entre los ajustes disponibles se incluyen:

- **Fuente.** La fuente utilizada para todo el texto en los nuevos gráficos.
- **Preferencia de ciclos de estilo.** La asignación inicial de colores y tramas para nuevos gráficos. Mostrar sucesivamente sólo los colores utiliza únicamente colores para diferenciar los elementos gráficos y no utiliza tramas. Mostrar sucesivamente sólo las tramas sólo utiliza estilos de línea, símbolos de marcador o tramas de relleno para diferenciar los elementos gráficos y no utiliza color.

- **Marco.** Controla la presentación de los marcos interno y externo en los nuevos gráficos.
- **Líneas de cuadrícula.** Controla la presentación de las líneas de cuadrícula de los ejes de categorías y escala en los nuevos gráficos.

**Ciclos de estilo.** Personaliza los colores, estilos de línea, símbolos de marcador y tramas de relleno para los gráficos nuevos. Puede cambiar el orden de los colores o tramas utilizados al crear un gráfico nuevo.

### **Colores de los elementos de datos**

Especifique el orden en que se deben utilizar los colores para los elementos de datos (por ejemplo, barras y marcadores) en el nuevo gráfico. Los colores se utilizan siempre que se selecciona una opción que incluye *color* en el grupo Preferencia de ciclos de estilo del cuadro de diálogo principal Opciones de gráfico.

Si crea, por ejemplo, un gráfico de barras agrupadas con dos grupos y selecciona Mostrar sucesivamente los colores y, a continuación, las tramas en el cuadro de diálogo principal Opciones de gráfico, los dos primeros colores de la lista Gráficos agrupados se utilizan como colores de barra en el nuevo gráfico.

#### **Para cambiar el orden en que se utilizan los colores**

- ▶ Seleccione Gráficos simples y, a continuación, seleccione un color utilizado para los gráficos sin categorías.
- ▶ Seleccione Gráficos agrupados para cambiar el ciclo de color de los gráficos con categorías. Para cambiar el color de una categoría, selecciónela y elija un color en la paleta para dicha categoría.

Si lo desea, puede:

- Insertar una nueva categoría por encima de la categoría seleccionada.
- Desplazar una categoría seleccionada.
- Eliminar una categoría seleccionada.
- Restablecer la secuencia por defecto.
- Editar un color mediante la selección de su casilla y pulsando en Editar.

### **Líneas de los elementos de datos**

Especifique el orden en que se deben utilizar los estilos para los elementos de datos de línea en el nuevo gráfico. Los estilos de línea se utilizan siempre que en el gráfico hay elementos de datos de línea y se selecciona una opción que incluye *tramas* en el grupo Preferencia de ciclos de estilo del cuadro de diálogo principal Opciones de gráfico.

Si crea, por ejemplo, un gráfico de líneas con dos grupos y selecciona Mostrar sucesivamente sólo las tramas en el cuadro de diálogo principal Opciones de gráfico, los dos primeros estilos de la lista Gráficos agrupados se utilizan como tramas de línea en el nuevo gráfico.



**Para cambiar el orden en que se utilizan los estilos de línea**

- ▶ Seleccione Gráficos simples y, a continuación, seleccione un estilo de línea utilizado para los gráficos de líneas sin categorías.
- ▶ Seleccione Gráficos agrupados para cambiar el ciclo de trama de los gráficos de líneas con categorías. Para cambiar el estilo de línea de una categoría, selecciónela y elija un estilo de línea en la paleta para dicha categoría.

Si lo desea, puede:

- Insertar una nueva categoría por encima de la categoría seleccionada.
- Desplazar una categoría seleccionada.
- Eliminar una categoría seleccionada.
- Restablecer la secuencia por defecto.

**Marcadores de los elementos de datos**

Especifique el orden en que se deben utilizar los símbolos para los elementos de datos de marcadores en el nuevo gráfico. Los estilos de marcadores se utilizan siempre que en el gráfico hay elementos de datos de marcadores y se selecciona una opción que incluye *tramas* en el grupo Preferencia de ciclos de estilo del cuadro de diálogo principal Opciones de gráfico.

Si crea, por ejemplo, un gráfico de diagramas de dispersión con dos grupos y selecciona Mostrar sucesivamente sólo las tramas en el cuadro de diálogo principal Opciones de gráfico, los dos primeros símbolos de la lista Gráficos agrupados se utilizan como marcadores en el nuevo gráfico.

**Para cambiar el orden en que se utilizan los estilos de marcadores**

- ▶ Seleccione Gráficos simples y, a continuación, seleccione un símbolo de marcador utilizado para los gráficos sin categorías.
- ▶ Seleccione Gráficos agrupados para cambiar el ciclo de trama de los gráficos con categorías. Para cambiar el símbolo de marcador de una categoría, selecciónela y elija un símbolo en la paleta para dicha categoría.

Si lo desea, puede:

- Insertar una nueva categoría por encima de la categoría seleccionada.
- Desplazar una categoría seleccionada.
- Eliminar una categoría seleccionada.
- Restablecer la secuencia por defecto.

**Rellenos de los elementos de datos**

Especifique el orden en que se deben utilizar los estilos de relleno para los elementos de datos de la barra y área en el nuevo gráfico. Los estilos de relleno se utilizan siempre que en el gráfico hay elementos de datos de barra o área y se selecciona una opción que incluye *tramas* en el grupo Preferencia de ciclos de estilo del cuadro de diálogo principal Opciones de gráfico.

Si crea, por ejemplo, un gráfico de barras agrupadas con dos grupos y selecciona **Mostrar** sucesivamente sólo las tramas en el cuadro de diálogo principal **Opciones de gráfico**, los dos primeros estilos de la lista **Gráficos agrupados** se utilizan como tramas de relleno de barra en el nuevo gráfico.

***Para cambiar el orden en que se utilizan los estilos de relleno***

- ▶ Seleccione **Gráficos simples** y, a continuación, seleccione una trama de relleno utilizada para los gráficos sin categorías.
- ▶ Seleccione **Gráficos agrupados** para cambiar el ciclo de trama de los gráficos con categorías. Para cambiar la trama de relleno de una categoría, selecciónela y elija una trama de relleno en la paleta para dicha categoría.

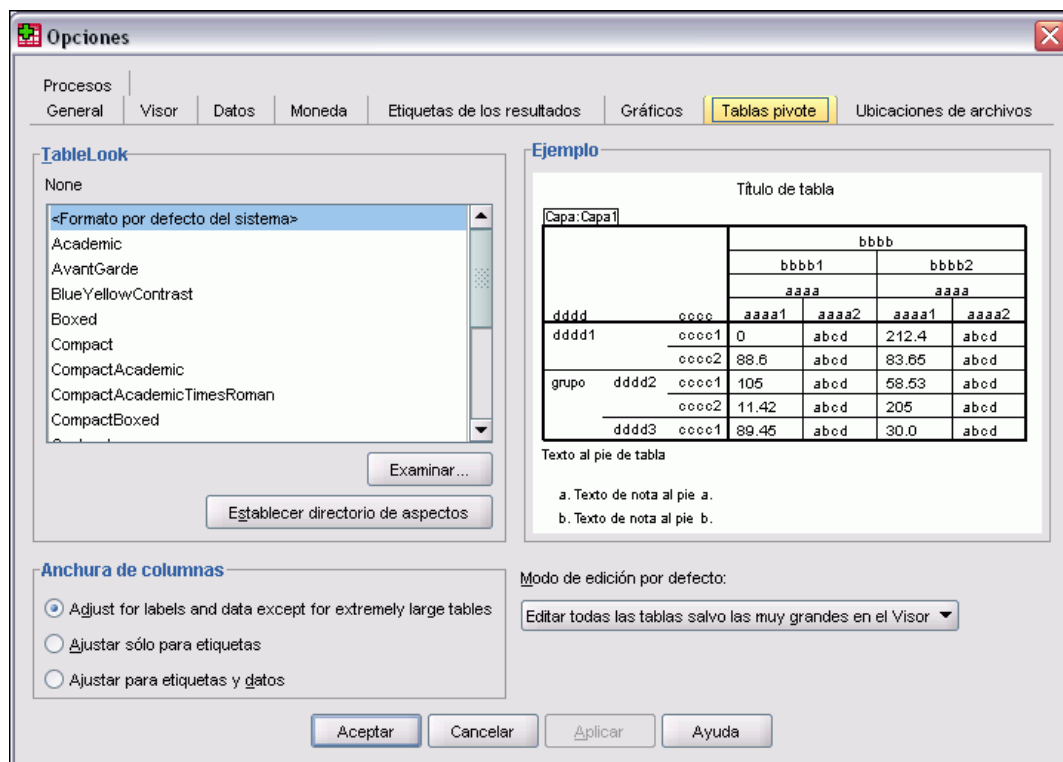
Si lo desea, puede:

- Insertar una nueva categoría por encima de la categoría seleccionada.
- Desplazar una categoría seleccionada.
- Eliminar una categoría seleccionada.
- Restablecer la secuencia por defecto.

## ***Tabla pivote: Opciones***

Las opciones de **Tabla pivote** establecen el aspecto de tabla por defecto, utilizado para los resultados de nuevas tablas pivote. Aspectos de tabla puede controlar una variedad de atributos de tabla pivote, incluyendo la presentación y el ancho de las líneas de cuadrícula; el estilo, tamaño y color de la fuente; así como los colores de fondo.

Figura 45-8  
Cuadro de diálogo Opciones: pestaña Tablas pivote



**Aspecto de tabla.** Seleccione un aspecto de tabla en la lista de archivos y pulse en Aceptar o Aplicar. Puede utilizar uno de los aspectos de tabla que se incluyen en SPSS, o bien crear uno propio en el editor de tablas pivote (menú Formato, Aspectos de tabla).

- **Examinar.** Permite seleccionar un aspecto de tabla de otro directorio.
- **Establecer directorio de aspectos.** Le permite cambiar el directorio de aspectos por defecto.

*Nota:* los aspectos de tabla creados en versiones anteriores de SPSS no se pueden utilizar en la versión 16.0 o posterior.

**Ajustar los anchos de columna para** Controla el ajuste automático de los anchos de columna en las tablas pivote.

- **Sólo las etiquetas.** Ajusta el ancho de columna al ancho de la etiqueta de la columna. Así se generan tablas más compactas, pero no se mostrarán los valores de los datos más anchos que la etiqueta (los asteriscos indican valores demasiado anchos para ser mostrados).
- **Etiquetas y Datos.** Ajusta el ancho de columna al más ancho de entre la etiqueta de columna y el mayor de los valores de los datos. Así se generan tablas más anchas, pero se asegura que se mostrarán todos los valores.

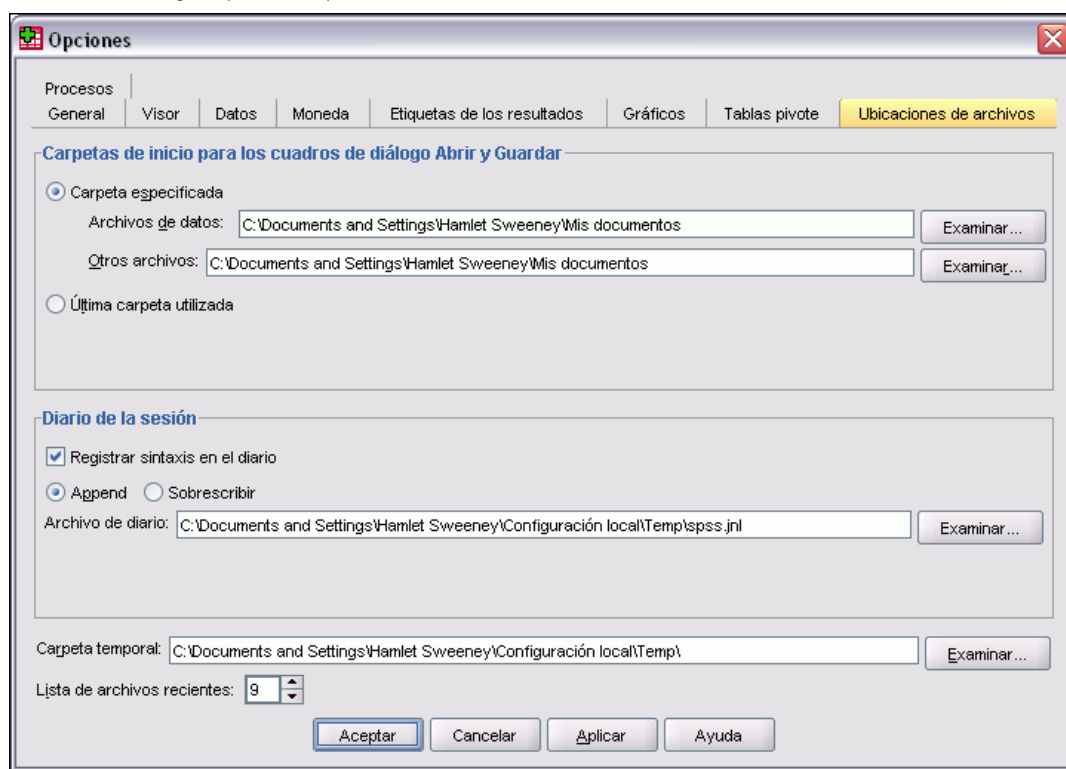
**Modo de edición por defecto.** Controla la activación de las tablas pivote en la ventana del Visor o en una ventana independiente. Por defecto, cuando se pulsa dos veces en una tabla pivote, se activan todas las tablas en la ventana del Visor, excepto las que son muy grandes. Puede elegir entre activar las tablas pivote en una ventana independiente o seleccionar una opción de tamaño

que abra las tablas pivote más pequeñas en la ventana del Visor y las más grandes en una ventana independiente.

## Opciones de ubicaciones de archivos

Las opciones de la pestaña Ubicaciones de archivos controlan la ubicación por defecto que la aplicación utilizará para abrir y guardar archivos al inicio de cada sesión, la ubicación del archivo de diario, la ubicación de la carpeta temporal y el número de archivos que aparecerán en la lista de archivos utilizados recientemente.

Figura 45-9  
Cuadro de diálogo Opciones: pestaña Ubicaciones de archivos



### **Carpeta de inicio para los cuadros de diálogo Abrir y Guardar**

- **Carpeta especificada.** La carpeta especificada se emplea como ubicación por defecto al comienzo de cada sesión. Puede especificar ubicaciones por defecto distintas para los archivos de datos y el resto de archivos (que no sean de datos).
- **Última carpeta utilizada.** La última carpeta utilizada para abrir o guardar archivos en la sesión anterior es la que se utiliza por defecto al comienzo de la siguiente sesión. Se aplica tanto a los archivos de datos como al resto de archivos (que no sean de datos).

Estos ajustes sólo se aplican a los cuadros de diálogo para abrir o guardar archivos y la “última carpeta utilizada” está determinada por el último cuadro de diálogo que se utilizó para abrir o guardar un archivo. Los archivos abiertos o guardados a través de la sintaxis de comandos no tienen efecto sobre estos ajustes, ni se ven afectados por ellos.

### ***Diario de la sesión***

Puede utilizar el diario de la sesión para registrar automáticamente los comandos ejecutados en cada sesión. Incluye comandos introducidos y ejecutados en ventanas de sintaxis y comandos generados por elecciones de cuadros de diálogo. Puede editar este archivo y volver a utilizar los comandos en otras sesiones. Puede activar o desactivar el registro de sesión, añadir o sustituir el archivo diario y seleccionar el nombre y la ubicación del mismo. Si lo desea, puede copiar la sintaxis de comandos del archivo de diario y guardarla en un archivo de sintaxis.

### ***Carpeta temporal***

Especifica la ubicación de los archivos temporales creados durante una sesión. En el análisis en modo distribuido (disponible con la versión de servidor), esto no afecta a la ubicación de los archivos de datos temporales. En el modo distribuido, la ubicación de los archivos de datos temporales está controlada por la variable de entorno *SPSSTMPDIR*, la cual se puede definir sólo en el ordenador que ejecuta la versión del servidor del software. Si necesita cambiar la ubicación del directorio temporal, póngase en contacto con el administrador del sistema.

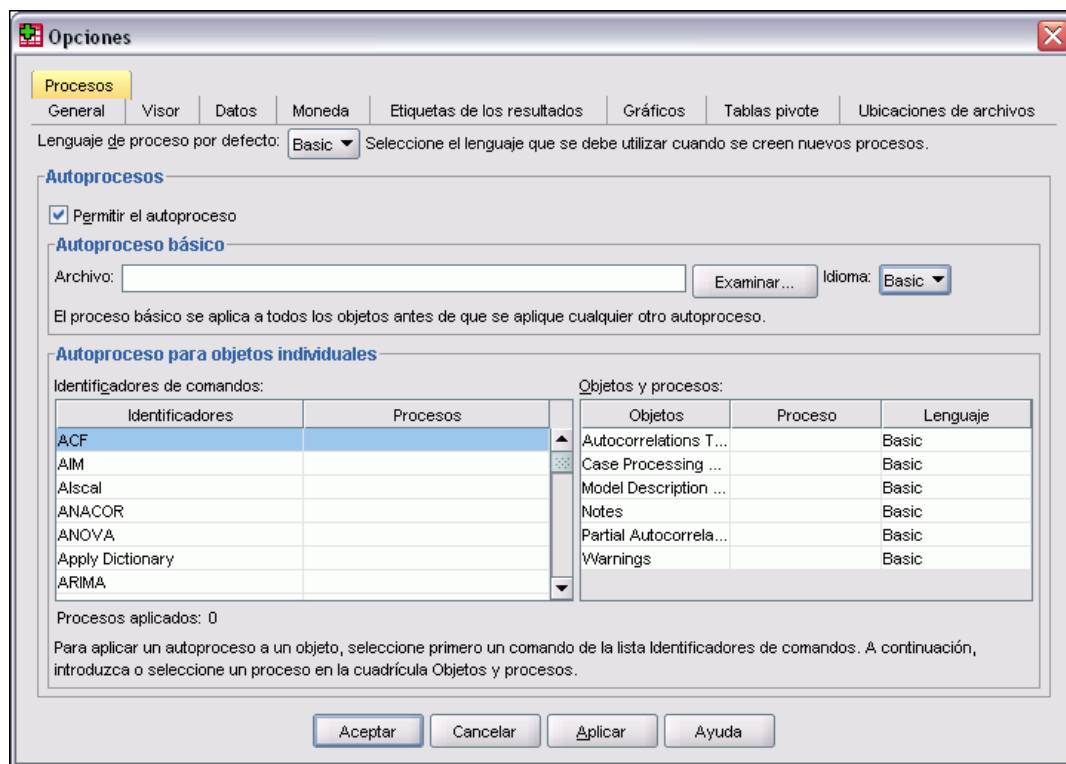
### ***Lista de archivos utilizados recientemente***

Controla el número de archivos utilizados recientemente que aparecen en el menú Archivo.

## ***Procesos: Opciones***

Utilice la pestaña Procesos para especificar el lenguaje de procesos por defecto y los autoprocesos que desea utilizar. Puede utilizar procesos para automatizar muchas funciones, incluyendo la personalización de tablas pivote.

Figura 45-10  
Cuadro de diálogo Opciones: pestaña Procesos



*Nota:* los usuarios del antiguo Sax Basic deberán convertir manualmente los autoprocesos personalizados. Los autoprocesos instalados con versiones anteriores a la 16.0 están disponibles como un conjunto independiente de archivos de procesos ubicado en el subdirectorio *Samples* del directorio donde se ha instalado SPSS . Por defecto, ningún elemento de resultados está asociado con los autoprocesos. Deberá asociar manualmente todos los autoprocesos a los elementos de resultados deseados, tal como se describe a continuación.

**Lenguaje de procesos por defecto.** Esta opción determina el editor de procesos que se ejecuta cuando se crean nuevos procesos. También especifica el lenguaje por defecto cuyo ejecutable se utilizará para ejecutar autoprocesos. Los lenguajes de procesos disponibles dependen de la plataforma. Para Windows, los lenguajes de procesos disponibles son Basic, instalado con el sistema Base, y el lenguaje de programación Python. Para el resto de plataformas, el procesamiento está disponible con el lenguaje de programación Python.

*Nota:* Para activar la ejecución de procesos en lenguaje de programación Python, deberá instalar por separado Python y el complemento de integración SPSS-Python. El lenguaje de programación Python no se instala junto con el sistema Base, pero está disponible en el CD de instalación para Windows. El complemento de integración SPSS-Python está disponible en el CD de instalación para todas las plataformas..

**Permitir el autoproceso.** Esta casilla de verificación permite activar o desactivar el autoprocesamiento. Por defecto, el autoprocesamiento está activado.

**Autoproceso básico.** Proceso opcional que se aplica a todos los objetos nuevos del Visor antes de la aplicación de cualquier otro autoproceso. Especifique el archivo de procesos que se utilizará como autoproceso básico y el lenguaje cuyo ejecutable se utilizará para ejecutar el proceso.

#### ***Para aplicar autoprocesos a los elementos de resultados***

- ▶ En la cuadrícula Identificadores de comandos, elija un comando que genere elementos de resultados a los que se aplicarán autoprocesos.

La columna *Objetos*, de la cuadrícula Objetos y procesos, muestra una lista de los objetos asociados con el comando seleccionado. La columna *Proceso* muestra los procesos existentes para el comando seleccionado.

- ▶ Especifique un proceso para cualquiera de los elementos que aparecen en la columna *Objetos*. Pulse en la correspondiente casilla *Proceso*. Escriba la ruta al proceso o haga clic en el botón con puntos suspensivos (...) para buscar el proceso.
- ▶ Especifique el lenguaje cuyo ejecutable se utilizará para ejecutar el proceso. *Nota:* el lenguaje seleccionado no se verá afectado al modificar el lenguaje del proceso por defecto.
- ▶ Pulse en Aplicar o Aceptar.

#### ***Para eliminar asociaciones de procesos***

- ▶ En la cuadrícula Objetos y procesos, pulse en la casilla de la columna Proceso correspondiente al proceso cuya asociación desea eliminar.
- ▶ Borre la ruta del proceso y, a continuación, pulse en otra casilla de la cuadrícula Objetos y procesos.
- ▶ Pulse en Aplicar o Aceptar.

#### ***Para establecer las opciones de procesos***

Elija en los menús:

Edición  
Opciones

- ▶ Pulse en la pestaña Procesos.
- ▶ Seleccione los ajustes deseados.
- ▶ Pulse en Aplicar o Aceptar.

# ***Personalización de menús y barras de herramientas***

## ***Editor de menús***

Puede personalizar los menús utilizando el Editor de menús. Con el Editor de menús es posible:

- Añadir elementos de menú que ejecuten procesos personalizados.
- Añadir elementos de menú que ejecuten archivos de sintaxis de comandos.
- Añadir elementos de menú que ejecuten otras aplicaciones y envíen los datos automáticamente a otras aplicaciones.

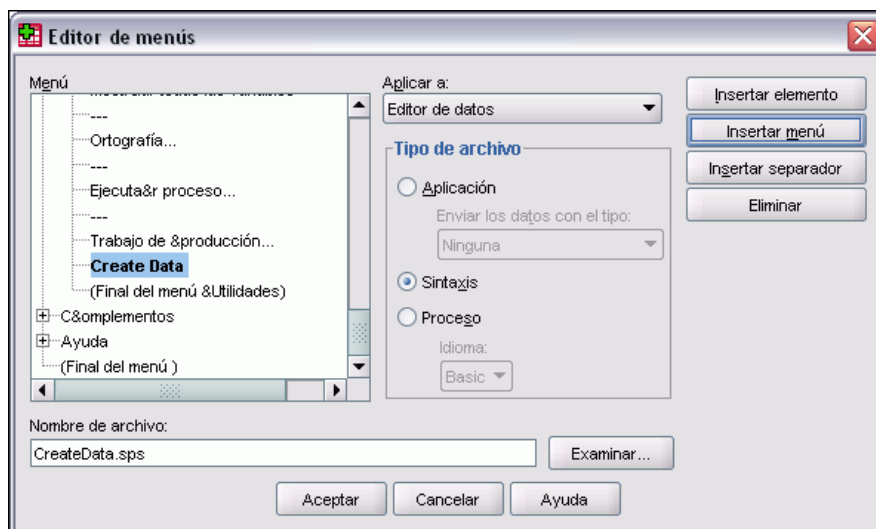
Puede enviar datos a otras aplicaciones en los siguientes formatos: SPSS, Excel, Lotus 1-2-3, delimitado por tabuladores y dBASE IV.

### ***Para añadir elementos a los menús***

- ▶ Elija en los menús:
  - Ver
  - Editor de menús...
- ▶ En el cuadro de diálogo Editor de menús, pulse dos veces (o pulse en el icono del signo más) en el menú donde desee añadir un nuevo elemento.
- ▶ Seleccione el elemento de menú sobre el que desea que aparezca el nuevo elemento.
- ▶ Pulse en Insertar elemento para insertar un nuevo elemento de menú.
- ▶ Seleccione el tipo de archivo para el nuevo elemento (archivo de proceso, archivo de sintaxis de comandos o aplicación externa).
- ▶ Pulse en Examinar para seleccionar un archivo que sea anexionado al elemento de menú.



Figura 46-1  
Cuadro de diálogo Editor de menús



También se pueden añadir menús completamente nuevos y separadores entre los elementos de menú. Si lo desea, puede enviar automáticamente el contenido del Editor de datos a otra aplicación cuando seleccione esa aplicación en el menú.

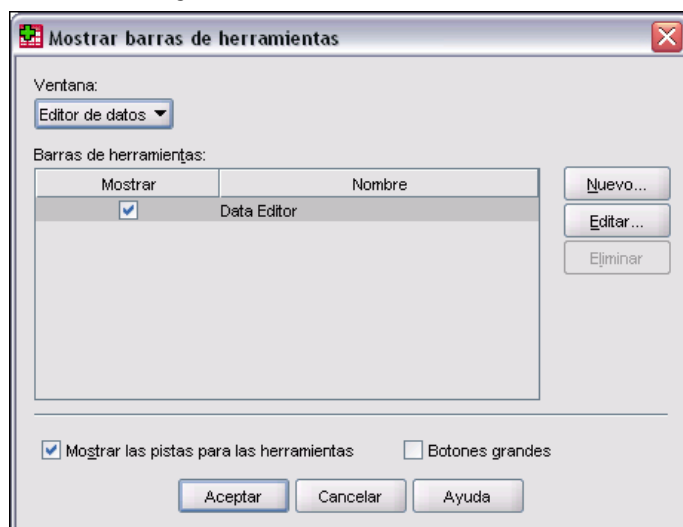
## ***Personalización de las barras de herramientas***

Puede personalizar las barras de herramientas y crear nuevas barras de herramientas. En las barras de herramientas puede incluirse cualquier herramienta disponible, incluso la de cualquier acción de menú. Además pueden contener herramientas personalizadas que ejecutan otras aplicaciones, que ejecutan archivos de sintaxis de comandos o archivos de procesos.

## ***Mostrar barras de herramientas***

Utilice Mostrar barras de herramientas para mostrar u ocultar, personalizar y crear nuevas barras de herramientas. En las barras de herramientas puede incluirse cualquier herramienta disponible, incluso la de cualquier acción de menú. Además pueden contener herramientas personalizadas que ejecutan otras aplicaciones, que ejecutan archivos de sintaxis de comandos o archivos de procesos.

Figura 46-2  
Cuadro de diálogo *Mostrar barras de herramientas*



## ***Para personalizar las barras de herramientas***

- ▶ Elija en los menús:  
Ver  
Barras de herramientas  
Personalizar
- ▶ Seleccione la barra de herramientas que desea personalizar y pulse en Edición o pulse en Nueva para crear una nueva barra de herramientas.
- ▶ Para las barras de herramientas nuevas, introduzca un nombre para la barra de herramientas, seleccione las ventanas en las que desea que aparezca y pulse en Edición.
- ▶ Seleccione un elemento en la lista Categorías para que se visualicen las herramientas disponibles en esa categoría.
- ▶ Arrastre y suelte las herramientas que desee en la barra de herramientas que aparece en el cuadro de diálogo.
- ▶ Para eliminar una herramienta de la barra de herramientas, arrástrela a cualquier punto fuera de la barra de herramientas que aparece en el cuadro de diálogo.

Para crear una herramienta personalizada que abra un archivo, ejecute un archivo de sintaxis de comandos o ejecute un proceso:

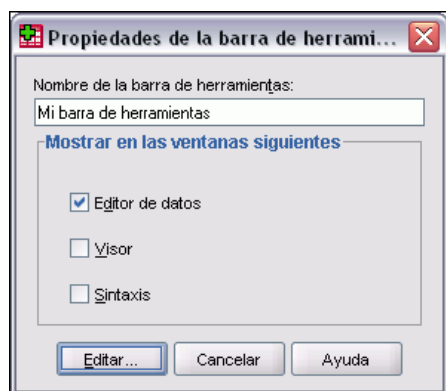
- ▶ Pulse en Nueva herramienta en el cuadro de diálogo Barra de herramientas de edición.
- ▶ Introduzca una etiqueta descriptiva para la herramienta.
- ▶ Seleccione la acción que desee realizar con la herramienta (abrir un archivo, ejecutar un archivo de sintaxis de comandos o ejecutar un proceso).
- ▶ Pulse en Examinar para seleccionar un archivo o una aplicación para asociarlos a la herramienta.

Las nuevas herramientas se muestran en la categoría Personales , que además contiene los elementos de menú definidos por el usuario.

## **Propiedades de la barra de herramientas**

Utilice Propiedades de la barra de herramientas para seleccionar los tipos de ventana en los que desee que aparezca la barra de herramientas seleccionada. Este cuadro de diálogo también se utiliza para crear nombres para nuevas barras de herramientas.

Figura 46-3  
Cuadro de diálogo Propiedades de la barra de herramientas



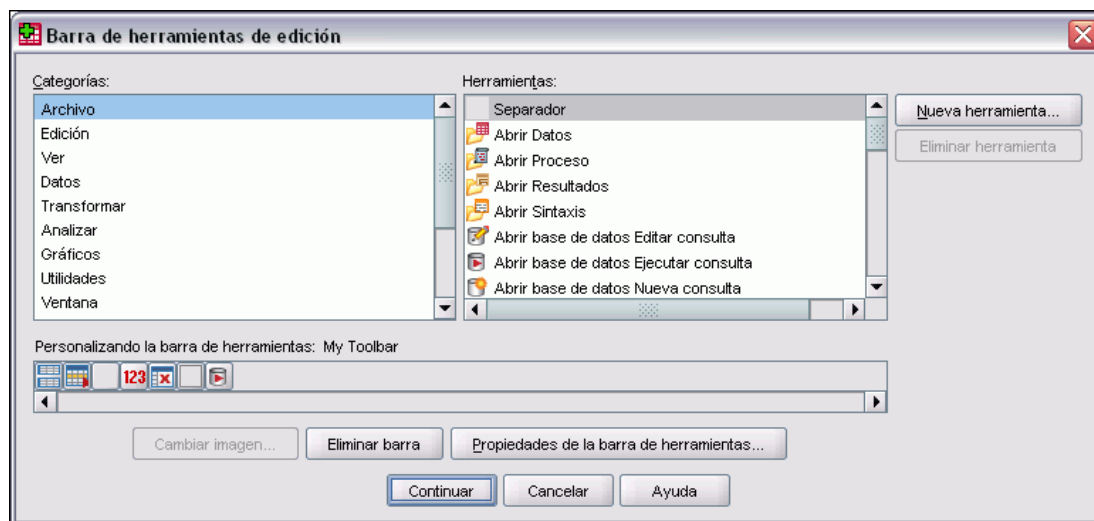
### **Para establecer las propiedades de la barra de herramientas**

- ▶ Elija en los menús:  
Ver  
Barras de herramientas  
Personalizar
- ▶ Para las barras de herramientas existentes, pulse en Edición y, a continuación, en Propiedades en el cuadro de diálogo Barra de herramientas de edición.
- ▶ En el caso de barras de herramientas nuevas, pulse en Nueva herramienta.
- ▶ Seleccione los tipos de ventana en los que desee que aparezca la barra de herramientas. Para las barras de herramientas nuevas, introduzca además un nombre.

## **Barra de herramientas de edición**

Utilice el cuadro de diálogo Barra de herramientas de edición para personalizar las barras de herramientas existentes y para crear nuevas barras. En las barras de herramientas puede incluirse cualquier herramienta disponible, incluso la de cualquier acción de menú. Además pueden contener herramientas personalizadas que ejecutan otras aplicaciones, que ejecutan archivos de sintaxis de comandos o archivos de procesos.

Figura 46-4  
Cuadro de diálogo Personalizar barra de herramientas



### **Para cambiar las imágenes de la barra de herramientas**

- ▶ Seleccione la herramienta cuya imagen desea cambiar en la barra de herramientas.
- ▶ Pulse en Cambiar imagen.
- ▶ Seleccione el archivo de imagen que desea utilizar para la herramienta. Se admiten los siguientes formatos de imagen: BMP, PNG, GIF, JPG.
  - Las imágenes deben ser cuadradas. Las imágenes no cuadradas se recortan hasta formar un cuadrado.
  - El tamaño de las imágenes se ajusta automáticamente. Para una visualización óptima, utilice imágenes de 16x16 píxeles para imágenes pequeñas de la barra de herramientas, o de 32x32 píxeles para imágenes grandes.

### **Crear nueva herramienta**

Utilice el cuadro de diálogo Crear nueva herramienta para crear herramientas personalizadas para ejecutar otras aplicaciones, ejecutar archivos de sintaxis de comandos y ejecutar archivos de procesos.

Figura 46-5  
Cuadro de diálogo *Crear nueva herramienta*

**Crear nueva herramienta**

**Descripción**

Etiqueta:

**Acción**

Aplicación

Enviar los datos con el tipo:

Sintaxis

Proceso

Idioma:

Nombre de archivo:

## ***Trabajos de producción***

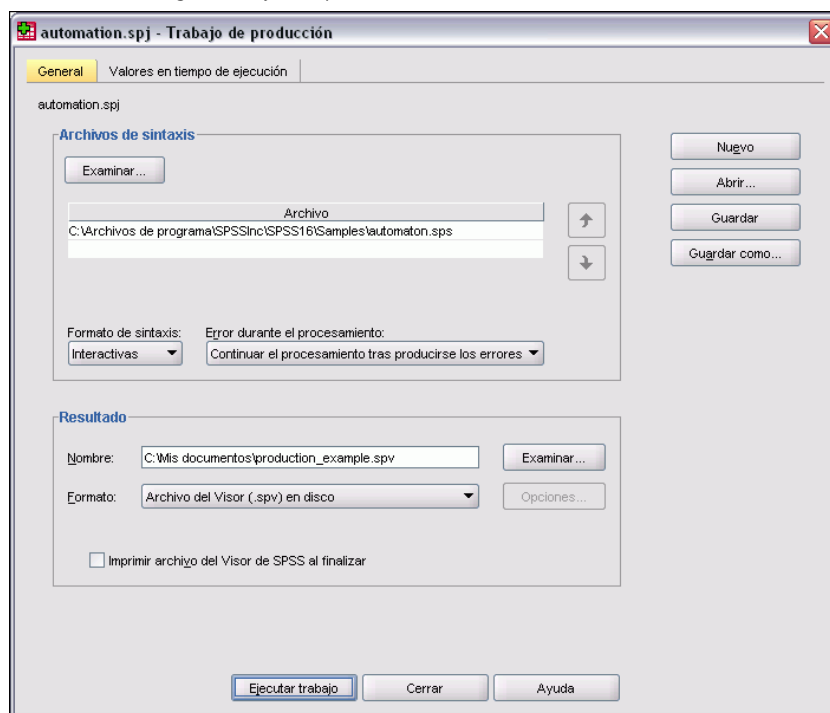
Los trabajos de producción proporcionan la capacidad de ejecutar SPSS automáticamente. El programa se ejecuta sin intervención del usuario y finaliza tras ejecutar el último comando, lo que le permite realizar otras tareas mientras se ejecuta o programar el nuevo trabajo de producción para que se ejecute automáticamente en las horas programadas. Los trabajos de producción son útiles si se ejecuta a menudo el mismo conjunto de análisis que consumen mucho tiempo, tales como los informes semanales.

Los trabajos de producción utilizan archivos de sintaxis de comandos para indicarle a SPSS lo que debe hacer. Un archivo de sintaxis de comandos es un archivo de texto con sintaxis de comandos. Puede utilizar cualquier editor de textos para crearlo. También puede generar la sintaxis de comandos pegando las selecciones del cuadro de diálogo en una ventana de sintaxis o editando el archivo de diario. Si desea obtener más información, consulte [Trabajar con sintaxis de comandos](#) en Capítulo 12 en p. 261.

Para crear un trabajo de producción:

- ▶ En los menús de cualquier ventana, seleccione:
  - Utilidades
  - Trabajo de producción

Figura 47-1  
Cuadro de diálogo Trabajo de producción



- Pulse en Nuevo para crear un nuevo trabajo de producción o en Abrir para abrir un trabajo de producción existente.

*Nota:* los archivos de trabajo de la unidad de producción (.spj) creados con versiones anteriores a la 16.0 no funcionarán en la versión 16.0 o versiones posteriores. Hay disponible una utilidad de conversión para convertir los archivos de trabajo de la unidad de producción de Windows y Macintosh a trabajos de producción (.spj). Si desea obtener más información, consulte [Conversión de los archivos de la unidad de producción](#) en p. 510.

- Seleccione uno o más archivos de sintaxis de comandos.

*Nota:* en el análisis en modo distribuido, las ubicaciones de los archivos de sintaxis deben introducirse manualmente y las rutas deben ser relativas al servidor remoto. Si desea obtener más información, consulte [Especificaciones de rutas absolutas frente a rutas relativas](#) en Capítulo 4 en p. 77.

- Seleccione el nombre del archivo de resultados, su ubicación y formato.
- Pulse en Guardar o en Guardar como para guardar el trabajo de producción.

**Formato de sintaxis.** Controla la forma de las reglas de sintaxis utilizadas para el trabajo.

- **Interactivo.** Cada comando debe finalizar con un punto. Los puntos pueden aparecer en cualquier parte del comando y los comandos pueden continuar en varias líneas, aunque un punto como último carácter no en blanco en una línea se interpreta como el final del comando. Las líneas de continuación y los comandos nuevos pueden empezar en cualquier parte de

una línea nueva. Se trata de reglas “interactivas” que surten efecto al seleccionar y ejecutar comandos en una ventana de sintaxis.

- **Lote.** Cada comando debe comenzar al principio de una línea nueva (sin espacios en blanco antes del inicio del comando) y las líneas de continuación se deben sangrar como mínimo un espacio. Si desea sangrar comandos nuevos, puede utilizar un signo más o un punto como primer carácter al inicio de la línea y, a continuación, sangrar el comando. El punto del final del comando es opcional. Este ajuste es compatible con las reglas de sintaxis para los archivos de comandos incluidos con el comando `INCLUDE`.

*Nota:* no utilice la opción Lote si sus archivos de sintaxis contienen la sintaxis de comandos de `GGRAPH` que incluya instrucciones `GPL`. Las instrucciones `GPL` sólo se ejecutarán según reglas interactivas.

**Procesamiento de errores.** Controla el tratamiento de errores en el trabajo.

- **Continuar procesando después de los errores.** Los errores del trabajo no detienen automáticamente el procesamiento de los comandos. Los comandos de los archivos de trabajo de producción se tratan como parte de la secuencia de comandos normal y el procesamiento de comandos continúa del modo normal.
- **Detener procesamiento inmediatamente.** El procesamiento de comandos se detiene si se detecta un primer error en el archivo de trabajo de producción. Esto es compatible con el comportamiento de los archivos de comandos incluidos con el comando `INCLUDE`.

**Resultados.** Controla el nombre, la ubicación y el formato de los resultados del trabajo de producción. Se encuentran disponibles las siguientes opciones de formato:

- **Archivo del Visor al disco (.spv).** Los resultados se guardan en el formato del Visor de SPSS , en la ubicación especificada del archivo.
- **Archivo del Visor (.spv) a PES Repository.** Es necesaria la opción SPSS Adaptor para Enterprise Services.
- **Informes Web (.spv) a PES Repository.** Es necesaria la opción SPSS Adaptor para Enterprise Services.
- **Word/RTF (\*.doc).** Las tablas pivote se exportan como tablas de Word, con todos los atributos de formato intactos (por ejemplo, bordes de casilla, estilos de fuente y colores de fondo). Los resultados de texto se exportan en formato RTF. En los sistemas operativos Windows, los gráficos se incluyen en el documento en formato EMF (metarchivo mejorado). En los demás sistemas operativos, los gráficos se incluyen en formato PNG. Los resultados de texto siempre se muestran con un tipo de letra de paso fijo (monoespaciadas) y se exportan con los mismos atributos de fuente. Es necesario utilizar una fuente de paso fijo para que los resultados de texto separado por espacios queden correctamente alineados; por defecto, la mayoría de los navegadores de Web utilizan una fuente de paso fijo para el texto con formato previo.

*Nota:* es posible que Microsoft Word no muestre correctamente las tablas muy anchas.

- **Excel (\*.xls).** Las filas, columnas y casillas de tablas pivote se exportan como filas, columnas y casillas de Excel, con todos los atributos de formato intactos, por ejemplo, bordes de casilla, estilos de fuente y colores de fondo, etc. Los resultados de texto se exportan con todos los atributos de fuente intactos. Cada línea del resultado de texto constituye una fila del archivo de Excel y se incluye todo su contenido en una sola casilla. No se incluyen los gráficos.



- **HTML (\*.htm).** Las tablas pivote se exportan como tablas HTML. Los resultados de texto se exportan como formato previo de HTML. Los gráficos se incrustan mediante referencias y deben exportarse en un formato adecuado para incluirlos en los documentos HTML (por ejemplo, PNG y JPEG).
- **Formato de documento portátil (\*.pdf).** Todos los resultados se exportan como aparecen en la presentación preliminar, con todos los atributos de formato intactos.
- **Archivo de PowerPoint (\*.ppt).** Las tablas pivote se exportan como tablas de Word y se incrustan en diapositivas diferentes en el archivo de PowerPoint (una diapositiva por cada tabla pivote). Se conservan todos los atributos de formato de la tabla pivote, por ejemplo, bordes de casillas, estilos de fuente y colores de fondo. Los gráficos se exportan en formato TIFF. No se incluyen los resultados de texto.

*Nota:* la exportación a PowerPoint sólo está disponible en los sistemas operativos Windows y no está disponible con la versión para estudiantes.

- **Texto (\*.txt).** Entre los formatos de resultados de texto se incluyen texto sin formato, UTF-8, and UTF-16. Las tablas pivote se pueden exportar en formato separado por tabuladores o por espacios. Todos los resultados de texto se exportan en formato separado por espacios. Para los gráficos, se inserta en el archivo de texto una línea por cada gráfico, indicando el nombre de archivo del gráfico exportado.

**Imprimir archivo del Visor de SPSS al finalizar.** Envía el archivo de resultados finales del Visor a la impresora al finalizar el trabajo de producción.

**Ejecutar trabajo.** Ejecuta el trabajo de producción en otra sesión. Resulta muy útil para comprobar los nuevos trabajos de producción antes de implementarlos.

### ***Trabajos de producción con comandos OUTPUT***

Los trabajos de producción tienen en cuenta comandos OUTPUT de SPSS, como OUTPUT SAVE, OUTPUT ACTIVATE y OUTPUT NEW. Los comandos OUTPUT SAVE, que se ejecutan durante un trabajo de producción, escribirán el contenido de los documentos de resultados especificados en las ubicaciones especificadas. Esto se añade al archivo de resultados creado por el trabajo de producción. Al utilizar OUTPUT NEW para crear un nuevo documento de resultados, se recomienda que lo guarde expresamente con el comando OUTPUT SAVE.

El archivo de resultados del trabajo de producción está compuesto por el contenido del documento de resultados activo durante la finalización del trabajo. Para aquellos trabajos que contengan comandos OUTPUT, el archivo de resultados no puede contener todos los resultados creados en la sesión. Por ejemplo, supongamos que el trabajo de producción está compuesto por un número de procedimientos de SPSS seguidos de un nuevo comando OUTPUT NEW, seguidos de más procedimientos de SPSS, pero no de ningún comando OUTPUT. El comando OUTPUT NEW define un nuevo documento de salida activo. Al finalizar el trabajo de producción, contendrá únicamente los resultados de los procedimientos que se hayan ejecutado después del comando OUTPUT NEW.

## ***Opciones HTML***

**Opciones de tabla.** No hay disponibles opciones de tabla para el formato HTML. Las tablas pivote se convierten en tablas HTML.

**Opciones de imagen.** Los tipos de imagen disponibles son: EPS, JPEG, TIFF, PNG y BMP. En los sistemas operativos Windows, también está disponible el formato EMF (metarchivo mejorado). También puede ajustar la escala del tamaño de imagen del 1% al 200%.

## ***Opciones de PowerPoint***

**Opciones de tabla.** Puede utilizar las entradas de titulares del Visor como los títulos de las diapositivas. Cada diapositiva contiene un único elemento de resultado. El título se genera a partir de la entrada del titular para el elemento en el panel de titulares del Visor.

**Opciones de imagen.** Puede ajustar la escala del tamaño de imagen del 1% al 200%. (Las imágenes se exportan a PowerPoint en formato TIFF.)

*Nota:* el formato PowerPoint sólo está disponible en sistemas operativos Windows y necesita PowerPoint 97 o una versión posterior.

## ***Opciones de PDF***

**Incrustar marcadores.** Esta opción incluye en el documento PDF los marcadores correspondientes a las entradas de titulares del Visor. Al igual que el panel de titulares del Visor, los marcadores pueden facilitar mucho la navegación por los documentos que tienen un gran número de objetos de resultados.

**Incrustar fuentes.** Al incrustar fuentes se garantiza que el documento PDF presentará el mismo aspecto en cualquier ordenador. De lo contrario, si algunas fuentes utilizadas en el documento no están disponibles en el ordenador donde se visualiza o imprime el documento PDF, la sustitución de fuentes puede resultar en una calidad menor.

## ***Opciones del texto***

**Opciones de tabla.** Las tablas pivote se pueden exportar en formato separado por tabuladores o por espacios. Para el formato separado por espacios, también puede controlar:

- **Ancho de columna.** Autoajuste no muestra nunca el contenido de las columnas en varias líneas. Cada columna tiene el ancho correspondiente al valor o etiqueta más ancho que haya en dicha columna. Personalizado establece un ancho de columna máximo que se aplica a todas las columnas de la tabla. Los valores que superen dicho ancho se mostrarán en otra línea en la misma columna.
- **Carácter de borde de fila/columna.** Controla los caracteres utilizados para crear los bordes de fila y columna. Para suprimir la visualización de los bordes de fila y columna, especifique espacios en blanco para estos valores.

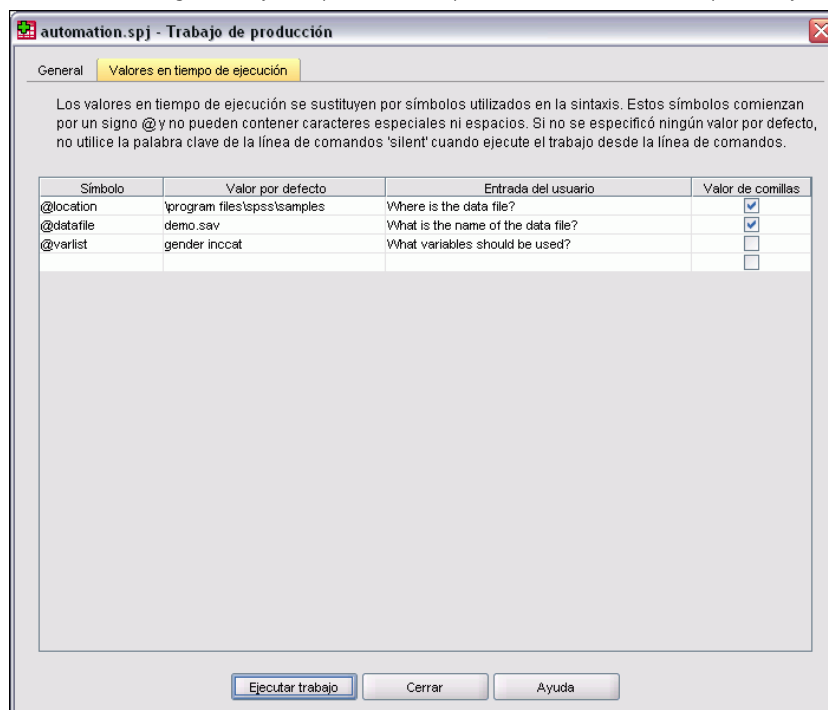
**Opciones de imagen.** Los tipos de imagen disponibles son: EPS, JPEG, TIFF, PNG y BMP. En los sistemas operativos Windows, también está disponible el formato EMF (metarchivo mejorado). También puede ajustar la escala del tamaño de imagen del 1% al 200%.

## Valores de tiempo de ejecución

Los valores de tiempo de ejecución definidos en un archivo de trabajo de producción y utilizados en un archivo de sintaxis de comandos simplifican tareas tales como la ejecución del mismo análisis para diversos archivos de datos o del mismo conjunto de comandos para distintos conjuntos de variables. Por ejemplo, podría definir los valores de tiempo de ejecución `@datafile` para que se le solicite el nombre de un archivo de datos cada vez que ejecute un trabajo de producción que utilice la cadena `@datafile` en vez de un nombre de archivo en el archivo de sintaxis de comandos.

Figura 47-2

Cuadro de diálogo Trabajo de producción, pestaña Valores de tiempo de ejecución



**Símbolo.** La cadena del archivo de sintaxis de comandos que desencadena que el trabajo de producción solicite el valor al usuario. El nombre del símbolo debe comenzar con un signo @ y debe cumplir las normas de denominación de variables. Si desea obtener más información, consulte [Nombres de variable](#) en Capítulo 5 en p. 82.

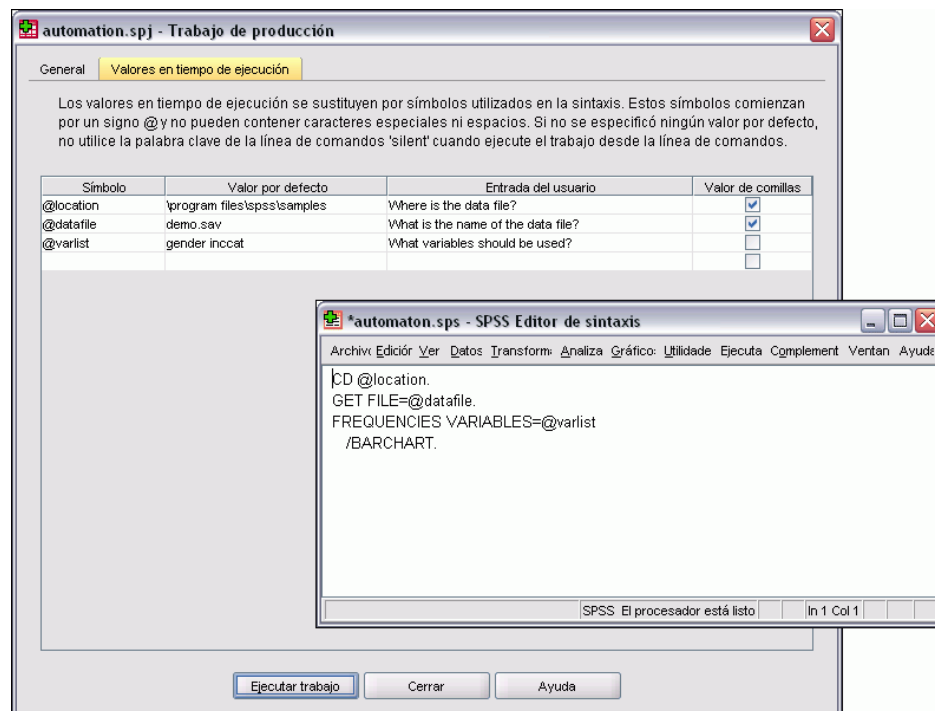
**Valor por defecto.** Valor que utiliza por defecto el trabajo de producción si no se introduce un valor nuevo. Este valor se muestra cuando el trabajo de producción solicita información al usuario. Puede sustituir o modificar el valor en el momento de la ejecución. Si no introduce un valor por defecto, no utilice la palabra clave `silent` al ejecutar el trabajo de producción con modificadores de la línea de comandos, a no ser que también utilice el modificador `-symbol` para especificar los valores de tiempo de ejecución. Si desea obtener más información, consulte [Ejecución de trabajos de producción desde una línea de comandos](#) en p. 509.

**Entrada del usuario.** Etiqueta descriptiva que se muestra cuando el trabajo de producción solicita al usuario que introduzca información. Por ejemplo, podría utilizar la frase “¿Qué archivo de datos desea utilizar?” para identificar un campo que requiera el nombre de un archivo de datos.

**Valor entre comillas.** Encierra entre comillas el valor por defecto o el valor introducido por el usuario. Por ejemplo, las especificaciones de archivo deben ir entre comillas.

Figura 47-3

*Símbolos del tiempo de ejecución en un archivo de sintaxis de comandos*

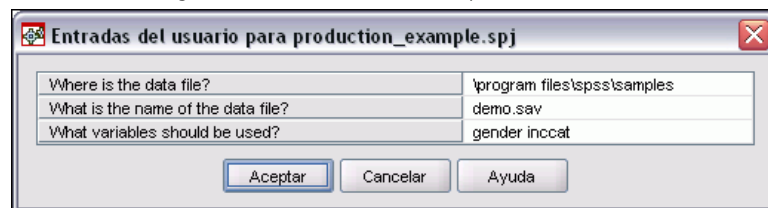


## Entradas del usuario

El trabajo de producción solicita valores al usuario cuando se ejecuta un trabajo de producción que contiene símbolos de tiempo de ejecución que han sido definidos. Se pueden sustituir o modificar los valores mostrados por defecto. Posteriormente, estos valores se sustituyen por los símbolos de tiempo de ejecución en todos los archivos de sintaxis de comandos asociados al trabajo de producción.

Figura 47-4

*Cuadro de diálogo Entradas del usuario de producción*



## Ejecución de trabajos de producción desde una línea de comandos

Los modificadores de la línea de comandos permiten al usuario programar los trabajos de producción automáticamente para que se ejecuten cada cierto tiempo con utilidades de programación disponibles en su sistema operativo. El formato básico del argumento de la línea de comandos es:

```
spss filename.spj -production
```

En función de cómo se haya invocado el trabajo de producción, es posible que sea necesario incluir rutas de acceso de los directorios para el archivo ejecutable *spss* (ubicado en el directorio donde se ha instalado la aplicación) y/o el archivo del trabajo de producción.

Puede ejecutar trabajos de producción desde una línea de comandos con los siguientes modificadores:

**-production [prompt|silent]**. Inicia la aplicación en modo de producción. Las palabras clave `prompt` y `silent` especifican si se desea mostrar el cuadro de diálogo que solicita los valores en tiempo de ejecución, en caso de que se hayan especificado en el trabajo. La palabra clave `prompt` es el valor por defecto y muestra el cuadro de diálogo. La palabra clave `silent` suprime dicho cuadro de diálogo. Si utiliza la palabra clave `silent`, puede definir los símbolos de tiempo de ejecución con el modificador `-symbol`. En caso contrario, se utilizará el valor por defecto. Los modificadores `-switchserver` y `-singleseat` se ignoran cuando se utiliza el modificador `-production`.

**-symbol <valores>**. Lista de parejas símbolo-valor utilizadas en el trabajo de producción. Cada nombre de símbolo comienza por `@`. Los valores que contienen espacios deben escribirse entre comillas. Las reglas que permiten incluir comillas o apóstrofes en los literales de cadena pueden cambiar de un sistema operativo a otro, pero por lo general es posible escribir una cadena que contenga comillas simples o apóstrofes entre comillas dobles (por ejemplo, 'un valor entre comillas').

Para ejecutar trabajos de producción en un servidor remoto en un análisis en modo distribuido, se debe introducir la información de acceso al servidor:

**-server <inet:nombrehost:puerto>**. Nombre o dirección IP y número de puerto del servidor. Sólo Windows.

**-user <nombre>**. Nombre válido de usuario. Si es necesario un nombre de dominio, escriba antes del nombre del usuario el nombre del dominio y una barra invertida (`\`). Sólo Windows.

**-password <contraseña>**. Contraseña del usuario. Sólo Windows.

### Ejemplo

```
spss \production_jobs\prodjob1.spj -production silent -symbol @datafile /data/July_data.sav
```

- Este ejemplo supone que está ejecutando la línea de comandos desde el directorio de instalación, así que no es necesaria una ruta para el archivo ejecutable *spss*.
- Este ejemplo también supone que el trabajo de producción especifica que el valor de `@datafile` debe estar entre comillas (casilla de verificación Valor entre comillas de la pestaña Valores de tiempo de ejecución), así que las comillas no son necesarias al especificar el archivo de datos en la línea de comandos. De lo contrario, sería necesario especificar, por ejemplo,

"/data/July\_data.sav" para incluir las comillas con la especificación del archivo de datos, ya que las especificaciones de archivos deben figurar entre comillas en la sintaxis de comandos.

- La ruta de acceso al directorio de la ubicación del trabajo de producción utiliza la barra diagonal inversa de Windows. En Macintosh y Linux, se utilizan las barras diagonales normales. Las barras diagonales normales en la especificación del archivo de datos entre comillas funcionarán en todos los sistemas operativos ya que la cadena entre comillas se inserta en el archivo de sintaxis de comandos y estas barras se aceptan en comandos que incluyen especificaciones de archivo (por ejemplo, GET FILE, GET DATA, SAVE) en todos los sistemas operativos.
- La palabra clave `silent` suprime las entradas del usuario en el trabajo de producción y el modificador `-symbol` inserta el nombre y la ubicación del archivo de datos entre comillas donde aparece el símbolo del tiempo de ejecución `@datafile` en los archivos de sintaxis de comandos incluidos en el trabajo de producción.

## ***Conversión de los archivos de la unidad de producción***

los archivos de trabajo de la unidad de producción (.spp) creados con versiones anteriores a la 16.0 no funcionarán en la versión 16.0 o versiones posteriores. Para los archivos de trabajo de la unidad de producción de Windows y Macintosh creados con versiones anteriores, se puede utilizar *prodconvert*, ubicado en el directorio de instalación, y de esta manera convertir esos archivos en nuevos archivos de trabajo de producción (.spj). La especificación básica (y única) es:

```
prodconvert filename.spp
```

Se creará un nuevo archivo con el mismo nombre pero con la extensión *.spj* en el mismo directorio del archivo original.

### ***Limitaciones***

- Los formatos de gráficos WMF y EMF no son compatibles. Se utilizará el formato PNG en su lugar.
- Las opciones de exportación Documento de resultados (sin gráficos), Sólo gráficos y Nada no son compatibles. Los objetos de resultados compatibles con el formato seleccionado están incluidos.
- La opción Mostrar SPSS y mantener abierto tras la finalización del trabajo no es compatible. La sesión de producción siempre se cierra tras la finalización del trabajo.
- Se ignorará la configuración del servidor remoto. Para especificar la configuración del servidor remoto para el análisis distribuido, se debe ejecutar el trabajo de producción desde una línea de comandos, utilizando los modificadores de la línea de comando para especificar la configuración del servidor. Si desea obtener más información, consulte [Ejecución de trabajos de producción desde una línea de comandos](#) en p. 509.
- Se ignorará la configuración Publicar en Web.

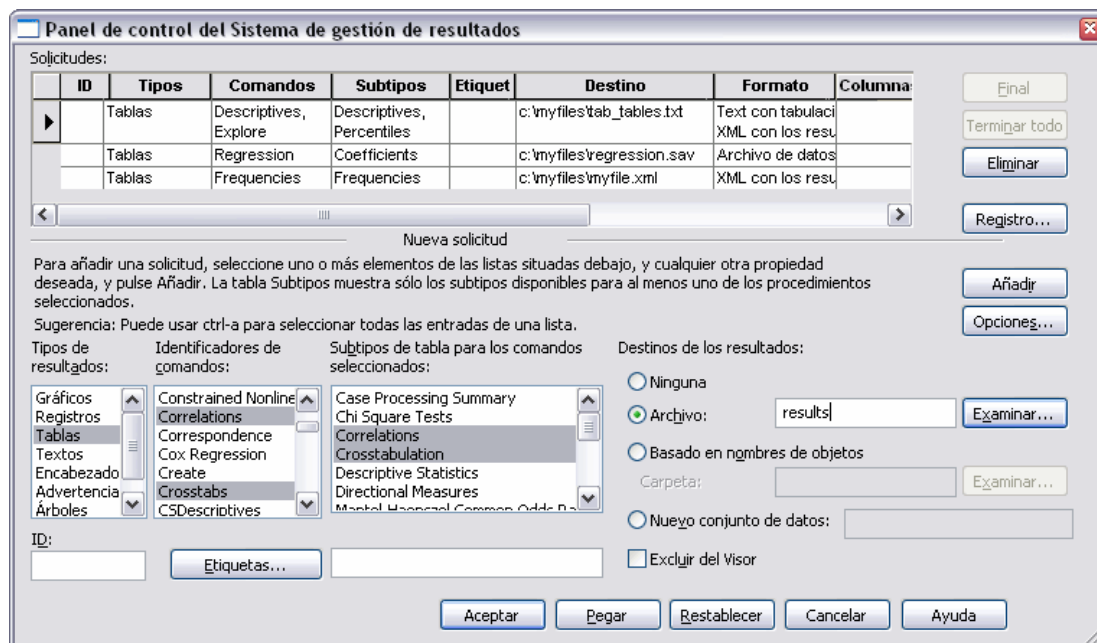
# Sistema de gestión de resultados

El Sistema de gestión de resultados (SGR) ofrece la posibilidad de escribir de forma automática las categorías de resultados seleccionadas en diferentes archivos de resultados y en distintos formatos. Los formatos incluyen: formato de archivo de datos de SPSS (.sav), formato de archivo del Visor de SPSS (.spv), XML, HTML y texto. Si desea obtener más información, consulte [Opciones de SGR](#) en p. 517.

## Para utilizar el panel de control del Sistema de gestión de resultados

- ▶ Elija en los menús:  
Utilidades  
Panel de control de SGR...

Figura 48-1  
Panel de control del Sistema de gestión de resultados



Puede utilizar el panel de control para iniciar y detener el envío de los resultados a los distintos destinos.

- Cada solicitud de SGR permanece activa hasta que finaliza de forma explícita o hasta el final de la sesión.

- Un archivo de destino especificado en una solicitud de SGR no estará disponible para otros procedimientos y otras aplicaciones hasta que finalice la solicitud de SGR.
- Mientras que una solicitud de SGR está activa, los archivos de destino especificados se almacenan en la memoria (RAM) y, por consiguiente, las solicitudes de SGR activas que escriben una gran cantidad de resultados en archivos externos pueden consumir una cantidad considerable de memoria.
- Las distintas solicitudes de SGR son independientes. Un mismo resultado se puede enviar a distintas ubicaciones en formatos diversos que dependen de las especificaciones de las distintas solicitudes de SGR.
- El orden de los objetos de resultados en un destino concreto equivale al orden en que se han creado, lo cual se determina mediante el orden y el funcionamiento de los procedimientos que generan los resultados.

#### **Para añadir solicitudes de SGR nuevas**

- ▶ Seleccione los tipos de resultados (tablas, gráficos, etc.) que desee incluir. Si desea obtener más información, consulte [Tipos de objetos de resultados](#) en p. 514.
- ▶ Seleccione los comandos que desee incluir. Si desea incluir todos los resultados, seleccione todos los elementos de la lista. Si desea obtener más información, consulte [Identificadores de comandos y subtipos de tabla](#) en p. 515.
- ▶ En el caso de los comandos que generan resultados de tablas pivote, seleccione los tipos de tabla concretos que desee incluir.

La lista muestra sólo las tablas disponibles en los comandos seleccionados; cualquier tipo de tabla disponible en uno o más comandos seleccionados se muestra en la lista. Si no se selecciona ningún comando, se muestran todos los tipos de tabla. Si desea obtener más información, consulte [Identificadores de comandos y subtipos de tabla](#) en p. 515.

- ▶ Para seleccionar tablas basadas en etiquetas de texto en lugar de subtipos, pulse en Etiquetas. Si desea obtener más información, consulte [Etiquetas](#) en p. 516.
- ▶ Pulse en Opciones para especificar el formato del resultado (por ejemplo, archivo de datos de SPSS, XML o HTML). Por defecto, se utiliza el formato XML con los resultados. Si desea obtener más información, consulte [Opciones de SGR](#) en p. 517.
- ▶ Especifique un destino de resultados:
  - **Archivo.** Todos los resultados seleccionados se envían a un único archivo.
  - **Basado en nombres de objetos.** Los resultados se envían a varios archivos de destino según los nombres de objetos. Se crea un archivo independiente para cada objeto de resultados, con un nombre de archivo basado o en los nombres de subtipo de tabla o en las etiquetas de tabla. Escriba el nombre de la carpeta de destino.
  - **Nuevo conjunto de datos.** En el caso de los resultados con formato de archivo de datos de SPSS, puede enviar los resultados a un conjunto de datos. El conjunto de datos está disponible para su uso posterior durante la misma sesión, pero no se guardará a menos que se haya guardado explícitamente antes de que finalice la sesión. Esta opción sólo está disponible para los resultados con formato de archivo de datos de SPSS. El nombre de un conjunto de datos



debe cumplir las normas de denominación de variables. Si desea obtener más información, consulte [Nombres de variable](#) en Capítulo 5 en p. 82.

- ▶ Si lo desea:
  - Excluya los resultados seleccionados del Visor. Si selecciona Excluir del Visor, los tipos de resultados de la solicitud de SGR no se mostrarán en la ventana del Visor. Si varias solicitudes SGR activas incluyen los mismos tipos de resultados, la presentación de dichos tipos de resultados en el Visor se determina mediante la solicitud de SGR más reciente que contiene los tipos de resultados. Si desea obtener más información, consulte [Exclusión de presentación de resultados del Visor](#) en p. 522.
  - Asigne una cadena de ID a la solicitud. A todas las solicitudes se les asigna automáticamente un valor de ID; puede anular la cadena de ID por defecto del sistema con un ID descriptivo, que puede resultar útil si dispone de varias solicitudes activas que desea identificar fácilmente. Los valores de ID que asigne no pueden empezar por un signo de dólar (\$).

Los siguientes consejos son útiles para seleccionar varios elementos de una lista:

- Pulse Ctrl+A para seleccionar todos los elementos de una lista.
- Mantenga pulsada la tecla Mayús para seleccionar varios elementos contiguos.
- Mantenga pulsada la tecla Ctrl para seleccionar varios elementos no contiguos.

### **Para finalizar y eliminar solicitudes de SGR**

Las solicitudes de SGR activas y nuevas se muestran en la lista Solicitudes con la solicitud más reciente en la parte superior. Puede cambiar los anchos de las columnas de información si pulsa y arrastra los bordes, y puede desplazar la lista horizontalmente para ver más información sobre una solicitud en concreto.

Un asterisco (\*) después de la palabra Activa en la columna *Estado* indica que se ha creado una solicitud de SGR con una sintaxis de comandos que incluye funciones no disponibles en el panel de control.

Para finalizar una solicitud de SGR activa concreta:

- ▶ Pulse en cualquier casilla de la fila para dicha solicitud en la lista Solicitudes.
- ▶ Pulse en Terminar.

Para finalizar todas las solicitudes de SGR activas:

- ▶ Pulse en Terminar todo.

Para eliminar una solicitud nueva (una solicitud que se ha añadido pero que aún no está activa):

- ▶ Pulse en cualquier casilla de la fila para dicha solicitud en la lista Solicitudes.
- ▶ Pulse en Eliminar.

*Nota:* las solicitudes de SGR activas no finalizan hasta que pulsa en Aceptar.

## ***Tipos de objetos de resultados***

Hay siete tipos de objetos de resultados distintos:

**Gráficos.** Los objetos de gráfico sólo se incluyen con los formatos de destino XML y HTML. Para el formato HTML, los archivos de imagen del gráfico se guardan en un subdirectorio (carpeta) independiente.

**Registros.** Objetos de texto de registro. Los objetos de registro contienen determinados tipos de mensajes de error y advertencia. Dependiendo de la configuración de las opciones (menú Edición, Opciones, pestaña Visor), los objetos de registro pueden contener también la sintaxis de comandos que se ha ejecutado durante la sesión. Los objetos de registro tienen la etiqueta *Registro* en el panel de titulares del Visor.

**Tablas.** Objetos de resultados que son tablas pivote en el Visor (incluye tablas de notas). Las tablas son los únicos objetos de resultados que se pueden dirigir al formato del archivo de datos de SPSS (.sav).

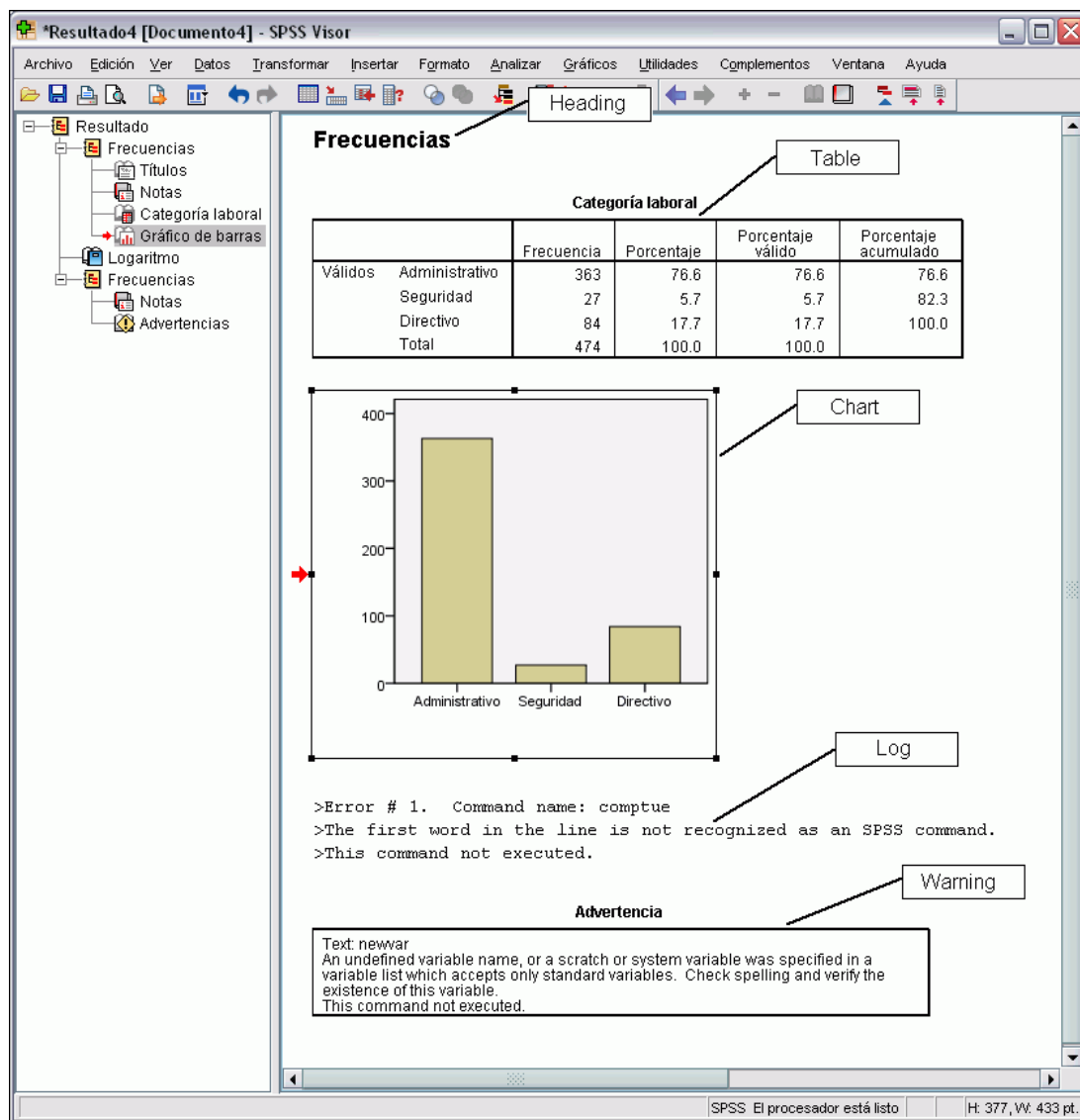
**Textos.** Objetos de texto que no son registros ni encabezados (incluyen los objetos con la etiqueta *Resultados de texto* en el panel de titulares del Visor).

**Encabezados.** Objetos de texto con la etiqueta *Título* en el panel de titulares del Visor. No se incluyen objetos de texto de encabezado con el formato XML de los resultados.

**Advertencias.** Los objetos de advertencias contienen determinados tipos de mensajes de error y advertencia.

**Árboles.** Diagramas de modelo de árbol generados por la opción *Árbol de clasificación*. Los objetos de árbol sólo se incluyen con los formatos de destino XML y HTML.

Figura 48-2  
Tipos de objetos de resultados



## Identificadores de comandos y subtipos de tabla

### Identificadores de comandos

Los identificadores de comandos están disponibles para todos los procedimientos estadísticos y de gráficos y para cualquier otro comando que genere bloques de resultados con su propio encabezado identificable en el panel de titulares del Visor. Estos identificadores son por lo general (aunque no siempre) iguales o similares a los nombres de los procedimientos de los menús y los títulos de los cuadros de diálogo, los cuales son por lo general (aunque no siempre) similares a los nombres

de los comandos subyacentes. Por ejemplo, el identificador de comandos para el procedimiento Frecuencias es “Frecuencias” y el nombre del comando subyacente es también el mismo.

No obstante, hay algunos casos en que el nombre del procedimiento y el identificador de comandos o el nombre del comando no son en absoluto similares. Por ejemplo, todos los procedimientos del submenú Pruebas no paramétricas (del menú Analizar) utilizan el mismo comando subyacente y el identificador de comandos es el mismo que el nombre de comando subyacente: Pruebas no paramétricas.

### ***Subtipos de tabla***

Los subtipos de tabla son los diferentes tipos de tablas pivote que se pueden generar. Algunos subtipos sólo están generados por un comando; otros subtipos se pueden generar mediante varios comandos (aunque las tablas pueden no presentar un aspecto similar). Aunque los nombres de subtipos de tabla suelen ser bastante descriptivos, puede haber muchos entre los que elegir (sobre todo si ha seleccionado un número considerable de comandos); además, dos subtipos pueden tener nombres muy similares.

### ***Para buscar identificadores de comandos y subtipos de tabla***

En caso de duda, puede buscar los nombres de los identificadores de comandos y los subtipos de tabla en la ventana del Visor:

- ▶ Ejecute el procedimiento para generar algunos resultados en el Visor.
- ▶ Pulse con el botón derecho del ratón en el elemento del panel de titulares del Visor.
- ▶ Seleccione Copiar identificador de comandos de SGR o Copiar subtipo de tablas de SGR.
- ▶ Pegue el nombre del identificador de comandos o del subtipo de tabla copiado en cualquier editor de texto (por ejemplo, una ventana Editor de sintaxis).

## ***Etiquetas***

En lugar de los nombres de subtipos de tabla, puede seleccionar tablas basadas en el texto mostrado en el panel de titulares del Visor. También puede seleccionar otros tipos de objeto basados en las etiquetas. Las etiquetas resultan útiles para diferenciar varias tablas del mismo tipo en las que el texto del titular refleja algún atributo del objeto de resultados concreto como las etiquetas o los nombres de las variables. Hay, no obstante, ciertos factores que pueden afectar al texto de la etiqueta:

- Si el procesamiento de segmentación del archivo está activado, es posible que se añada a la etiqueta una identificación con el grupo de segmentación del archivo.
- Las etiquetas que incluyen información sobre las variables o los valores se ven afectadas por la configuración actual de las opciones de las etiquetas de resultados (menú Edición, Opciones, pestaña Etiquetas de los resultados).
- Las etiquetas se ven afectadas por el ajuste actual del idioma de los resultados (menú Edición, Opciones, pestaña General).

**Para especificar las etiquetas que se van a utilizar para identificar los objetos de resultados**

- ▶ En el panel de control del Sistema de gestión de resultados, seleccione uno o más tipos de resultados y, a continuación, seleccione uno o más comandos.
- ▶ Pulse en Etiquetas.

Figura 48-3  
Cuadro de diálogo Etiquetas de SGR



- ▶ Introduzca la etiqueta *exactamente* como aparece en el panel de titulares de la ventana del Visor. También puede pulsar con el botón derecho del ratón en el elemento del titular, elegir Copiar etiqueta de SGR y pegar la etiqueta copiada en el campo de texto Etiqueta.
- ▶ Pulse en Añadir.
- ▶ Repita el proceso con cada etiqueta de tabla que desee incluir.
- ▶ Pulse en Continuar.

**Comodines**

Puede utilizar un asterisco (\*) como último carácter de la cadena de etiqueta y como carácter comodín. Se seleccionarán todas las etiquetas que empiecen por la cadena especificada (excepto el asterisco). Este proceso sólo funciona si el asterisco es el último carácter, ya que los asteriscos pueden aparecer como caracteres válidos en una etiqueta.

## Opciones de SGR

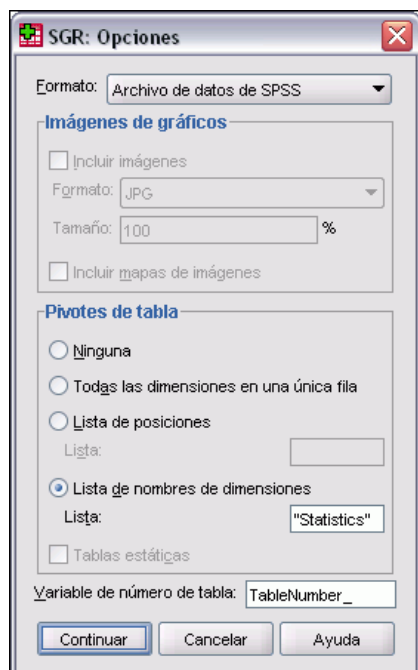
Puede utilizar el cuadro de diálogo Opciones de SGR para:

- Especificar el formato del resultado.
- Especificar el formato de la imagen (para formatos de resultados HTML y XML de resultados).
- Especificar qué elementos de la dimensión de tabla deben tener la dimensión de las filas.
- Incluya una variable que identifique el número de tabla secuencial que sea el origen en cada caso (para formato de archivo de datos de SPSS).

### Para especificar las opciones de SGR

- Pulse en Opciones en el panel de control del Sistema de gestión de resultados.

Figura 48-4  
Cuadro de diálogo SGR: Opciones



### Formato

**XML con los resultados.** XML que se adapta al esquema de *resultados de SPSS*. Los gráficos se incluyen como XML según el esquema vizml ([xml.spss.com/spss/visualization](http://xml.spss.com/spss/visualization)). Los diagramas de árbol se incluyen como XML según el esquema pmml ([www.dmg.org](http://www.dmg.org)). En lugar de incluir gráficos y diagramas de árbol en formato XML, puede guardarlos como archivos independientes en el formato gráfico seleccionado:

**HTML.** Objetos de resultados que son tablas pivote en el Visor y se convierten en tablas HTML simples. No se admite ningún atributo de aspecto de tabla (características de fuente, estilos de borde, colores, etc.). Los objetos de resultados de texto tienen la etiqueta <PRE> en HTML. Si opta por incluir gráficos, se exportarán como archivos independientes en el formato gráfico seleccionado y se incrustarán mediante referencias.

**Archivo de datos de SPSS.** Este formato es un formato de archivo binario. Todos los tipos de objetos de resultados distintos de las tablas se excluyen. Cada columna de una tabla se convierte en una variable en el archivo de datos. Para utilizar un archivo de datos creado con SGR en la misma sesión, deberá finalizar la solicitud de SGR activa si quiere abrir el archivo de datos. Si desea obtener más información, consulte [Envío de resultados a archivos de datos de SPSS](#) en p. 523.

**SPV.** Formato de archivo del Visor de SPSS. Es el mismo formato empleado para guardar el contenido de la ventana Visor.

**Texto.** Texto separado por espacios. Los resultados se escriben como texto con los resultados tabulares alineados con espacios para las fuentes de paso fijo. Se excluyen todos los gráficos.

**Texto con tabulaciones.** Texto delimitado por tabulaciones. Para los resultados que se muestran como tablas pivote en el Visor, las tabulaciones delimitan los elementos de columnas de tabla. Las líneas de bloque de texto se escriben sin cambios; no se realiza ningún intento de dividirlos con tabulaciones en las posiciones útiles. Se excluyen todos los gráficos.

### **Imágenes de gráficos**

Para los formatos HTML y XML de resultados, puede incluir gráficos y diagramas de modelo de árbol como archivos de imagen. Se crea un archivo de imagen diferente para cada gráfico y/o árbol.

- Para el formato de documento HTML, se incluyen las etiquetas estándar `<IMG SRC='nombrearchivo'>` en el documento HTML para cada imagen de archivo.
- Para el formato de documento XML de los resultados, el archivo XML contiene un elemento `chart` con un atributo `ImageFile` con el formato general `<chart imageFile="rutaarchivo/nombrearchivo"/>` para cada archivo de imagen.
- Los archivos de imagen se guardan en un subdirectorio (carpeta) independiente. El nombre del subdirectorio es el nombre del archivo de destino sin ninguna extensión y con `_files` añadido al final. Por ejemplo, si el archivo de destino es `datosjulio.htm`, el subdirectorio de imágenes se llamará `datosjulio_files`.

**Formato.** Los formatos de imagen disponibles son PNG, JPG, EMF, BMP y VML.

- El formato EMF (metarchivo mejorado) sólo está disponible en los sistemas operativos Windows.
- El formato de imagen VML sólo está disponible para el formato de documento HTML.
- El formato de imagen VML no crea archivos de imagen adicionales. El código VML que procesa la imagen está incrustado en el HTML.
- El formato de imagen VML no incluye diagramas de árbol.

**Tamaño.** Puede ajustar la escala del tamaño de imagen del 10% al 200%.

**Incluir mapas de imagen.** Para el formato de documento HTML, esta opción crea pistas de mapa de imagen que muestran información sobre algunos elementos del gráfico, como el valor del punto seleccionado en un gráfico de líneas o de la barra seleccionada en un gráfico de barras.

### **Pivotes de tabla**

Para los resultados de las tablas pivote, puede especificar los elementos de dimensión que deben aparecer en las columnas. El resto de los elementos de dimensión aparecen en las filas. Para el formato de archivo de datos de SPSS, las columnas de tabla se convierten en variables y las filas en casos.

- Si especifica varios elementos de dimensión para las columnas, estos se anidan en las columnas en el orden en que aparecen. Para el formato de archivo de datos de SPSS, los nombres de variable se generan mediante elementos de columna anidados. Si desea obtener más información, consulte [Nombres de variable en los archivos de datos generados por SGR](#) en p. 530.

- Si la tabla no contiene ninguno de los elementos de dimensión que aparecen, todos los elementos de dimensión para dicha tabla aparecerán en las filas.
- Los pivotes de tabla especificados aquí no surtirán ningún efecto en las tablas que se muestran en el Visor.

Cada dimensión de una tabla (fila, columna, capa) puede contener cero o más elementos. Por ejemplo, una tabla de contingencia sencilla de dos dimensiones contiene un único elemento de dimensión de fila y un único elemento de dimensión de columna, cada uno de los cuales contiene una de las variables utilizadas en la tabla. Puede utilizar argumentos de posición o “nombres” de elementos de dimensión para especificar los elementos de dimensión que desea colocar en la dimensión de columna.

**Todas las dimensiones en filas.** Crea una fila única para cada tabla. Para los archivos de datos con formato SPSS, significa que cada tabla es un caso único y que todos los elementos de tabla son variables.

**Lista de posiciones.** El formato general de un argumento de posición es una letra que indica la posición por defecto del elemento (C para columna, R para fila o L para capa) seguida de un número entero positivo que indica la posición por defecto en la dimensión. Por ejemplo, R1 indica el elemento de dimensión de fila más exterior.

- Para especificar varios elementos de diversas dimensiones, separe cada dimensión con un espacio: por ejemplo, R1 C2.
- La letra de dimensión seguida de ALL indica todos los elementos de dicha dimensión en el orden por defecto. Por ejemplo, CALL equivale al comportamiento por defecto (utilizando todos los elementos de columna en el orden por defecto para generar columnas).
- CALL RALL LALL (o RALL CALL LALL, etc.) coloca los elementos de dimensión en las columnas. Para el formato de archivo de datos de SPSS, se genera de este modo una fila o caso por tabla en el archivo de datos.

Figura 48-5  
Argumentos de posición de filas y columnas

		C1				
R1	R2	R3	Experiencia previa (meses)	Meses desde el contrato	Salario inicial	Fecha de nacimiento
Modelo						
1	Correlaciones	Experiencia previa (meses)	1,000	,067	-,087	,805
		Meses desde el contrato	,067	1,000	,012	,085
		Salario inicial	-,087	,012	1,000	-,075
		Fecha de nacimiento	,805	,085	-,075	1,000
	Covarianzas	Experiencia previa (meses)	31,307	12,940	-,022	7,096E-06
		Meses desde el contrato	12,940	1205,248	,019	4,635E-06
		Salario inicial	-,022	,019	,002	-5,236E-09
		Fecha de nacimiento	7,096E-06	4,635E-06	-5,236E-09	2,485E-12

**Lista de nombres de dimensiones.** En lugar de argumentos de posición, puede utilizar “nombres” de elementos de dimensión, que son las etiquetas de texto que aparecen en la tabla. Por ejemplo, una tabla de contingencia sencilla de dos dimensiones contiene un único elemento de dimensión de fila y un único elemento de dimensión de columna, cada uno de los cuales incluye etiquetas basadas



en las variables de dichas dimensiones además de un único elemento de dimensión de capa con la etiqueta *Estadísticos* (si el idioma de los resultados es el inglés).

- Los nombres de elementos de dimensión pueden variar según el idioma de los resultados y la configuración que influye en la presentación de los nombres de variable o las etiquetas de las tablas.
- Cada nombre de elemento de dimensión debe aparecer entre comillas simples o dobles. Para especificar varios nombres de elementos de dimensión, incluya un espacio entre cada nombre entre comillas.

Las etiquetas asociadas con los elementos de dimensión pueden no ser siempre evidentes.

***Para ver todos los elementos de dimensión y sus etiquetas para una tabla pivote***

- ▶ Active (pulse dos veces en) la tabla en el Visor.
- ▶ Elija en los menús:
  - Ver
  - Mostrar todo
- y/o
- ▶ Si no se muestran los paneles de pivotado, elija en los menús:
  - Pivotar
  - Paneles de pivotado

Las etiquetas del elemento aparecerán en los paneles de pivotado.

Figura 48-6  
Nombres de elementos de dimensión mostrados en la tabla y los paneles de pivotado

**Tabla sin las etiquetas de elementos de dimensión visibles**

Sexo		Salario inicial	Salario actual
Mujer	Media	\$13,091.97	\$26,031.92
	N	216	216
Hombre	Media	\$20,301.40	\$41,441.78
	N	258	258
Total	Media	\$17,016.09	\$34,419.57
	N	474	474

**Tabla con las etiquetas de elementos de dimensión visibles**

		Variables	
Sexo	Estadísticos	Salario inicial	Salario actual
Mujer	Media	\$13,091.97	\$26,031.92
	N		
Hombre	Media		
	N		
Total	Media		
	N		

## Registro

Puede registrar la actividad de SGR en un registro de XML o formato de texto.

- El registro realiza un seguimiento de todas las solicitudes de SGR nuevas para la sesión, pero no incluye las solicitudes de SGR activas antes de solicitar un registro.
- El archivo de registro actual finaliza si especifica un nuevo archivo de registro o si anula selección (desactiva) Registrar actividad de SGR.

Para especificar el registro de SGR:

- ▶ Pulse en Registro en el panel de control del Sistema de gestión de resultados.

## Exclusión de presentación de resultados del Visor

La casilla de verificación Excluir del Visor afecta a todos los resultados que se seleccionan en la solicitud de SGR al suprimir la presentación de esos resultados en la ventana del Visor. Este proceso suele resultar útil para los trabajos de producción que generan una gran cantidad de resultados y cuando no se necesitan los resultados como un documento del Visor (archivo .spv).

Además, puede utilizar esta funcionalidad para suprimir la presentación de determinados objetos de resultados que simplemente no desea ver sin enviar ningún otro resultado a un archivo y formato externos.

Para suprimir la presentación de determinados objetos de resultados sin dirigir otros resultados a un archivo externo:

- ▶ Cree una solicitud de SGR que identifique los resultados no deseados.
- ▶ Seleccione Excluir del Visor.
- ▶ Para el destino de los resultados, seleccione Archivo, pero deje el campo Archivo en blanco.
- ▶ Pulse en Añadir.

Los resultados seleccionados se excluyen del Visor, mientras que el resto de los resultados se muestran en el Visor del modo normal.

## ***Envío de resultados a archivos de datos de SPSS***

Un archivo de datos de SPSS consta de variables en las columnas y de casos en las filas, lo que representa básicamente el formato al que las tablas pivote pasan cuando se convierten en archivos de datos:

- Las columnas de la tabla son variables en el archivo de datos. Los nombres de variable válidos se generan a partir de etiquetas de columna.
- Las etiquetas de fila de la tabla se convierten en variables con nombres de variable genéricos (*Var1*, *Var2*, *Var3*, etc.) en el archivo de datos. Los valores de estas variables son etiquetas de fila en la tabla.
- En el archivo de datos se incluyen automáticamente tres variables de identificador de tabla: *Comando\_*, *Subtipo\_* y *Etiqueta\_*. Las tres son variables de cadena. Las dos primeras variables se corresponden con los identificadores de comandos y subtipos. Si desea obtener más información, consulte [Identificadores de comandos y subtipos de tabla](#) en p. 515. *Etiqueta\_* contiene el texto de título de tabla.
- Las filas de la tabla se convierten en casos en el archivo de datos.

### ***Ejemplo: Tabla sencilla de dos dimensiones***

En el caso más simple (una tabla sencilla de dos dimensiones), las columnas de tabla se convierten en variables y las filas se convierten en casos en el archivo de datos.

Figura 48-7  
Tabla sencilla de dos dimensiones

Informe			
Sexo		Salario actual	Salario inicial
Hombre	Media	\$41.441,78	\$20.301,40
	Mediana	\$32.850,00	\$15.750,00
	Mínimo	\$19.650	\$9.000
	Máximo	\$135.000	\$79.980
Mujer			
Total			

Comman d_	Subtype_	Label_	Var1	Var2	Salarioactual	Salarioini
1	Means Report	Informe	Hombre	Media	\$41,441.78	\$20,301.40
2	Means Report	Informe	Hombre	Mediana	\$32,850.00	\$15,750.00
3	Means Report	Informe	Hombre	Mínimo	\$19,650	\$9,000
4	Means Report	Informe	Hombre	Máximo	\$135,000	\$79,980
5	Means Report	Informe	Mujer	Media	\$26,031.92	\$13,091.97
6	Means Report	Informe	Mujer	Mediana	\$24,300.00	\$12,375.00
7	Means Report	Informe	Mujer	Mínimo	\$15,750	\$9,000
8	Means Report	Informe	Mujer	Máximo	\$58,125	\$30,000
9	Means Report	Informe	Total	Media	\$34,419.57	\$17,016.09

- Las primeras tres variables identifican la tabla de origen por comando, subtipo y etiqueta.
- Los dos elementos que definen las filas de la tabla (valores de la variable *Sexo* y medidas estadísticas) reciben los nombres de variable genéricos *Var1* y *Var2*. Estas dos variables son variables de cadena.
- Las etiquetas de columna de la tabla se utilizan para crear nombres de variable válidos. En este caso, dichos nombres de variable se basan en las etiquetas de variable de las tres variables de escala resumidas en la tabla. Si las variables no tienen etiquetas de variable definidas o bien ha optado por mostrar los nombres de variable en lugar de las etiquetas de variable como las etiquetas de columna en la tabla, los nombres de variable del nuevo archivo de datos deben ser los mismos que los del archivo de datos de origen.

### Ejemplo: Tablas con capas

Además de filas y columnas, una tabla puede contener una tercera dimensión: la dimensión de capas.

Figura 48-8  
Tabla con capas

		Sexo	
		Hombre	Mujer
Categoría laboral	Administrativo	47	40
	Seguridad		
	Directivo		

		Sexo	
		Hombre	Mujer
Categoría laboral	Administrativo	110	166
	Seguridad	14	0
	Directivo	70	10

Var1	Var2	Var3	Var4	Hombre	Mujer
1	Categoría laboral	Administrativo	Clasificación de minorías No	110	166
2	Categoría laboral	Administrativo	Clasificación de minorías Sí	47	40
3	Categoría laboral	Seguridad	Clasificación de minorías No	14	0
4	Categoría laboral	Seguridad	Clasificación de minorías Sí	13	0
5	Categoría laboral	Directivo	Clasificación de minorías No	70	10
6	Categoría laboral	Directivo	Clasificación de minorías Sí	4	0

- En la tabla, la variable con la etiqueta *Clasificación étnica* define las capas. En el archivo de datos, se crean dos variables adicionales: una que identifica el elemento de capa y otra que identifica las categorías del elemento de capa.
- Al igual que en el caso de las variables creadas a partir de elementos de fila, las variables creadas a partir de elementos de capa son variables de cadena con nombres de variable genéricos (el prefijo *Var* seguido de un número secuencial).

### Archivos de datos creados a partir de varias tablas

Si se envían varias tablas al mismo archivo de datos, cada una de ellas se añadirá al archivo de datos de un modo similar a la fusión de archivos de datos mediante la adición de casos de un archivo de datos a otro (menú Datos, Fundir archivos, Añadir casos).

- Cada tabla siguiente añade siempre casos al archivo de datos.
- Si las etiquetas de columna de las tablas difieren, cada tabla puede añadir variables al archivo de datos, con valores perdidos para los casos de otras tablas que no tienen una columna con un etiquetado idéntico.

#### Ejemplo: Varias tablas con las mismas etiquetas de columna

Si hay varias tablas con las mismas etiquetas de columna se suelen generar los archivos de datos más útiles de inmediato (archivos de datos que no requieren ninguna manipulación adicional). Por ejemplo, dos o más tablas de frecuencia del procedimiento Frecuencias tienen etiquetas de columna idénticas.

Figura 48-9  
 Dos tablas con etiquetas de columna idénticas

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos Hombre	258	54.4	54.4	54.4
Mujer	216	45.6	45.6	100.0

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos Administrativo	363	76.6	76.6	76.6
Seguridad	27	5.7	5.7	82.3
Directivo	84	17.7	17.7	100.0

Command_	Subtype_	Label_	Var1	Var2	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
1	Frecuencias	Sexo	Válidos	Hombre	258	54.4	54.4	54.4
2	Frecuencias	Sexo	Válidos	Mujer	216	45.6	45.6	100.0
3	Frecuencias	Sexo	Válidos	Total	474	100.0	100.0	.
4	Frecuencias	Categoría laboral	Válidos	Administrativo	363	76.6	76.6	76.6
5	Frecuencias	Categoría laboral	Válidos	Seguridad	27	5.7	5.7	82.3
6	Frecuencias	Categoría laboral	Válidos	Directivo	84	17.7	17.7	100.0
7	Frecuencias	Categoría laboral	Válidos	Total	474	100.0	100.0	.

- La segunda tabla contribuye con casos adicionales (filas) al archivo de datos, pero no aporta ninguna variable nueva, dado que las etiquetas de columna son exactamente iguales; por consiguiente, no hay ningún archivo de revisión de datos perdidos de gran tamaño.
- Aunque los valores para *Comando\_* y *Subtipo\_* son los mismos, el valor *Etiqueta\_* identifica la tabla de origen para cada grupo de casos dado que las dos tablas de frecuencia tienen distintos títulos.

**Ejemplo: Varias tablas con distintas etiquetas de columna**

Se crea una variable nueva en el archivo de datos para cada etiqueta de columna única en las tablas que se envían al archivo de datos. Este proceso ofrece como resultado bloques de valores perdidos si las tablas contienen distintas etiquetas de columna.

Figura 48-10  
 Dos tablas con distintas etiquetas de columna

The figure illustrates two tables with different column labels. The first table, titled 'Informe', shows salary data for Female and Male, categorized by Media and Mediana. The second table, also titled 'Informe', shows educational level and months since hire for Female and Male, categorized by Media and Mediana. The SPSS interface shows a data view with 12 rows and 10 variables, including Subtype, Label, Var1, Var2, BeginningSalary, CurrentSalary, EducationalLevelyears, and MonthssinceHire.

Informe			
Gender		Beginning Salary	Current Salary
Female	Media	\$13.091,97	\$26.031,92
	Mediana	\$12.375,00	\$24.300,00
Male	Media	\$20.301,40	\$41.441,78

Informe			
Gender		Educational Level (years)	Months since Hire
Female	Media	12,37	80,38
	Mediana	12,00	81,00
Male	Media	14,43	81,72

Subtype	Label	Var1	Var2	BeginningSalary	CurrentSalary	EducationalLevelyears	MonthssinceHire
1	Report	Informe	Female	Media	\$13,091.97	\$26,031.92	.
2	Report	Informe	Female	Mediana	\$12,375.00	\$24,300.00	.
3	Report	Informe	Male	Media	\$20,301.40	\$41,441.78	.
4	Report	Informe	Male	Mediana	\$15,750.00	\$32,850.00	.
5	Report	Informe	Total	Media	\$17,016.09	\$34,419.57	.
6	Report	Informe	Total	Mediana	\$15,000.00	\$28,875.00	.
7	Report	Informe	Female	Media	.	.	12,37
8	Report	Informe	Female	Mediana	.	.	12,00
9	Report	Informe	Male	Media	.	.	14,43
10	Report	Informe	Male	Mediana	.	.	15,00
11	Report	Informe	Total	Media	.	.	13,49
12	Report	Informe	Total	Mediana	.	.	12,00

- La primera tabla tiene columnas con las etiquetas *Salario inicial* y *Salario actual*, no presentes en la segunda tabla, lo que resulta en una serie de valores perdidos para dichas variables en los casos de la segunda tabla.
- Por el contrario, la segunda tabla tiene columnas con las etiquetas *Nivel de formación académica* y *Meses desde el contrato*, no presentes en la primera tabla, lo que resulta en una serie de valores perdidos para dichas variables en los casos de la primera tabla.
- Las variables discordantes como las de este ejemplo se pueden generar incluso con tablas del mismo subtipo. En este ejemplo, ambas tablas son del mismo subtipo.

**Ejemplo: Archivos de datos no creados a partir de varias tablas**

Si alguna tabla no tiene el mismo número de elementos de fila que otras tablas, no se creará ningún archivo de datos. El número de filas no tiene que ser el mismo; el número de *elementos* de fila que se convierten en variables en el archivo de datos debe ser el mismo. Por ejemplo, una tabla de contingencia de dos variables y una tabla de contingencia de tres variables contienen distintos números de elementos de fila dado que la variable “capa” está anidada en la variable de fila de la presentación de la tabla de contingencia de tres variables por defecto.

Figura 48-11

Tablas con distintos números de elementos de fila

**Tabla de contingencia Categoría laboral \* Sexo**

		Sexo		Total
		Mujer	Hombre	
Categoría laboral	Administrativo	206	157	363
	Seguridad	0	27	27
	Directivo	10	74	84
Total		216	258	474

**Tabla de contingencia Categoría laboral \* Sexo \* Clasificación de minorías**

Clasificación de minorías			Sexo		Total
			Mujer	Hombre	
No	Categoría	Administrativo	166	110	276
	laboral	Seguridad	0	14	14
		Directivo	10	70	80
	Total		176	194	370
Sí	Categoría	Administrativo	40	47	87
	laboral	Seguridad	0	13	13
		Directivo	0	4	4
	Total		40	64	104

**Control de elementos de columna para las variables de control del archivo de datos**

En el cuadro de diálogo Opciones del panel de control del Sistema de gestión de resultados puede especificar qué elementos de dimensión deben estar en las columnas y, por tanto, se utilizan para crear variables en el archivo de datos generado. Este proceso es equivalente al pivotado de la tabla en el Visor.

Por ejemplo, el procedimiento Frecuencias genera una tabla de estadísticos descriptivos con estadísticos en las filas, mientras que el procedimiento Descriptivos genera una tabla de estadísticos descriptivos con estadísticos en las columnas. Para incluir ambos tipos de tabla en el mismo archivo de datos de forma significativa, debe cambiar la dimensión de columna de uno de los tipos de tabla.

Dado que ambos tipos de tabla utilizan el nombre de elemento “Estadísticos” para la dimensión de estadísticos, podemos colocar los estadísticos de la tabla de estadísticos Frecuencias en las columnas si especifica “Estadísticos” (entre comillas) en la lista de nombres de dimensiones del cuadro de diálogo Opciones de SGR.



Figura 48-12  
Cuadro de diálogo SGR: Opciones

**SGR: Opciones**

Formato: Archivo de datos de SPSS

**Imágenes de gráficos**

Incluir imágenes

Formato: JPG

Tamaño: 100 %

Incluir mapas de imágenes

**Pivotes de tabla**

Ninguna

Todas las dimensiones en una única fila

Lista de posiciones

Lista:

Lista de nombres de dimensiones

Lista: "Statistics"

Tablas estáticas

Variable de número de tabla: TableNumber\_

Continuar Cancelar Ayuda

Figura 48-13

Combinación de distintos tipos de tablas en un archivo de datos mediante el pivotado de elementos de dimensión

Default Frequencies Statistics Table			
		Salario inicial	Salario actual
N	Válidos	474	474
	Perdidos	0	0
Media		\$17,016.09	\$34,419.57
Mínimo			
Máximo			

Pivoted Frequencies Statistics Table					
	N		Media	Mínimo	Máximo
	Válidos	Perdidos			
Salario inicial	474	0	\$17,016.09	\$9,000	\$79,980
Salario actual					\$135,000

Default Descriptive Statistics Table				
	N	Mínimo	Máximo	Media
Salario inicial	474	\$9,000	\$79,980	\$17,016.09
Salario actual	474	\$15,750	\$135,000	\$34,419.57

The screenshot shows the SPSS Editor de datos window with a pivot table. The table has columns for 'Var1', 'Válidos', 'Perdidos', 'Media', 'Mínimo', 'Máximo', and 'N'. The rows are numbered 1 through 7. Row 1: Salario inicial, Válidos: 474, Perdidos: 0, Media: \$17,016.09, Mínimo: \$9,000, Máximo: \$79,980. Row 2: Salario actual, Válidos: 474, Perdidos: 0, Media: \$34,419.57, Mínimo: \$15,750, Máximo: \$135,000. Row 3: Salario inicial, Válidos: ., Perdidos: ., Media: \$17,016.09, Mínimo: \$9,000, Máximo: \$79,980, N: 474. Row 4: Salario actual, Válidos: ., Perdidos: ., Media: \$34,419.57, Mínimo: \$15,750, Máximo: \$135,000, N: 474. Row 5: N válido (según lista), Válidos: ., Perdidos: ., Media: ., Mínimo: ., Máximo: ., N: 474. Rows 6 and 7 are empty.

Algunas de las variables incluyen valores perdidos, dado que las estructuras de tabla no son exactamente iguales con los estadísticos de las columnas.

### ***Nombres de variable en los archivos de datos generados por SGR***

SGR genera nombres de variable exclusivos y válidos a partir de etiquetas de columna:

- A los elementos de fila y capa se les asignan nombres de variable genéricos (el prefijo *Var* seguido de un número secuencial).
- Los caracteres no permitidos en los nombres de variable (espacios, paréntesis, etc.) se eliminan. Por ejemplo, “Esta Etiqueta (Columna)” se convierte en una variable con el nombre *EstaEtiquetaColumna*.
- Si la etiqueta empieza por un carácter permitido en los nombres de variable, pero no permitido como primer carácter (por ejemplo, un número), se inserta “@” como prefijo. Por ejemplo, “2o” se convierte en una variable llamada *@2o*.

- Los caracteres de subrayado o los puntos al final de las tablas se eliminan de los nombres de variable resultantes. Los caracteres de subrayado al final de las variables generadas automáticamente *Comando\_*, *Subtipo\_* y *Etiqueta\_* no se eliminan.
- Si hay más de un elemento en la dimensión de columna, los nombres de variable se generan mediante la combinación de etiquetas de categoría con caracteres de subrayado entre dichas etiquetas. Las etiquetas de grupo no se incluyen. Por ejemplo, si *VarB* se anida bajo *VarA* en las columnas, obtendrá variables como *CatA1\_CatB1*, pero no *VarA\_CatA1\_VarB\_CatB1*.

Figura 48-14

Nombres de variable generados a partir de elementos de tabla

Variable de capa		Variable de columna			Total
		Entrada	(Salida)		
Máximos	Variable de fila	Sí	1	1	2
		No	0	1	1
	Total		1	2	3
Mínimos	Variable de fila	Sí	1	0	1
		No	1	1	2
	Total		2	1	3

Variable de capa		Máximos		Mínimos		Total
		Variable de columna		Variable de columna		
		Entrada	(Salida)	Entrada	(Salida)	Total
Variable de fila	Sí	1	1	1	0	1
	No	0	1	1	1	2
Total		1	2	2	1	3

## Estructura de tablas OXML

El formato XML de resultados (OXML) equivale a XML adaptado al esquema *resultados de SPSS*. Para obtener una descripción detallada del esquema, consulte Esquema de resultados en el sistema de ayuda.

- Los identificadores de comandos y subtipos de SGR se utilizan como valores de los atributos `command` y `subType` en OXML. A continuación se muestra un ejemplo:

```
<command text="Frequencies" command="Frequencies"...>
<pivotTable text="Gender" label="Gender" subType="Frequencies"...>
```

- Los valores de los atributos `command` y `subType` de SGR no se ven afectados por el idioma de los resultados o la configuración de presentación para los nombres de variable y etiquetas o para los valores y etiquetas de valor.
- XML distingue entre mayúsculas y minúsculas. El valor del atributo `subType` de “frequencies” *no* es igual al valor del atributo `subType` de “Frequencies”.

- Toda la información mostrada en la tabla se incluye en los valores de atributo de OXML. En el nivel de casillas individuales, OXML consta de “elementos” vacíos que contienen atributos, pero no incluye ningún “contenido” distinto del que se incluye en los valores de atributo.
- La estructura de tablas en OXML se representa por filas; los elementos que representan las columnas se anidan en las filas y las casillas individuales se anidan en los elementos de columna:

```

<pivotTable...>
<dimension axis='row'...>
  <dimension axis='column'...>
    <category...>
      <cell text='...' number='...' decimals='...' />
    </category>
  <category...>
    <cell text='...' number='...' decimals='...' />
  </category>
</dimension>
</dimension>
...
</pivotTable>

```

El ejemplo anterior es una representación simplificada de la estructura que muestra las relaciones descendentes/ascendentes de estos elementos. Sin embargo, el ejemplo no muestra necesariamente las relaciones parentales/filiales, ya que suele haber niveles de elementos anidados intercalados.

La siguiente figura muestra una tabla de frecuencias simple y la representación completa de XML con los resultados de dicha tabla.

Figura 48-15  
Tabla de frecuencias simple

Sexo					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Hombre	258	54,4	54,4	54,4
	Mujer	216	45,6	45,6	100,0
	Total	474	100,0	100,0	

Figura 48-16  
XML con los resultados para la tabla de frecuencias simple

```

<?xml version="1.0" encoding="UTF-8" ?>
<outputTreeoutputTree xmlns="http://xml.spss.com/spss/oms"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://xml.spss.com/spss/oms
  http://xml.spss.com/spss/oms/spss-output-1.0.xsd">
<command text="Frequencies" command="Frequencies"
  displayTableValues="label" displayOutlineValues="label"
  displayTableVariables="label" displayOutlineVariables="label">
<pivotTable text="Gender" label="Gender" subType="Frequencies"
  varName="gender" variable="true">
<dimension axis="row" text="Gender" label="Gender"

```

```

varName="gender" variable="true">
<group text="Valid">
<group hide="true" text="Dummy">
<category text="Female" label="Female" string="f"
varName="gender">
<dimension axis="column" text="Statistics">
<category text="Frequency">
<cell text="216" number="216"/>
</category>
<category text="Percent">
<cell text="45.6" number="45.569620253165" decimals="1"/>
</category>
<category text="Valid Percent">
<cell text="45.6" number="45.569620253165" decimals="1"/>
</category>
<category text="Cumulative Percent">
<cell text="45.6" number="45.569620253165" decimals="1"/>
</category>
</dimension>
</category>
<category text="Male" label="Male" string="m" varName="gender">
<dimension axis="column" text="Statistics">
<category text="Frequency">
<cell text="258" number="258"/>
</category>
<category text="Percent">
<cell text="54.4" number="54.430379746835" decimals="1"/>
</category>
<category text="Valid Percent">
<cell text="54.4" number="54.430379746835" decimals="1"/>
</category>
<category text="Cumulative Percent">
<cell text="100.0" number="100" decimals="1"/>
</category>
</dimension>
</category>
</group>
<category text="Total">
<dimension axis="column" text="Statistics">
<category text="Frequency">
<cell text="474" number="474"/>
</category>
<category text="Percent">
<cell text="100.0" number="100" decimals="1"/>
</category>
<category text="Valid Percent">
<cell text="100.0" number="100" decimals="1"/>
</category>
</dimension>
</category>
</group>
</dimension>

```

```

</pivotTable>
</command>
</outputTree>

```

Como puede observar, una tabla sencilla de dimensiones reducidas genera una cantidad considerable de XML. Esto se debe en parte a que el XML contiene cierta información no evidente en la tabla original, cierta información que puede no estar disponible en la tabla original y una cantidad determinada de redundancia.

- El contenido de la tabla tal y como aparece (o podría aparecer) en una tabla pivote en el Visor se incluye en los atributos de texto. A continuación se muestra un ejemplo:

```
<command text="Frequencies" command="Frequencies"...>
```

- Los atributos de texto se pueden ver afectados por el idioma de los resultados y la configuración que influyen en la presentación de los nombres de variable y etiquetas o valores y etiquetas de valor. En este ejemplo, el valor del atributo de `texto` difiere en función del idioma de los resultados, mientras que el valor del atributo de comando permanece igual independientemente del idioma de los resultados.
- Siempre que las variables o los valores se utilicen en las etiquetas de fila o columna, XML contiene un atributo de `texto` y uno o más valores de atributo adicionales. A continuación se muestra un ejemplo:

```

<dimension axis="row" text="Gender" label="Gender" varName="gender">
...<category text="Female" label="Female" string="f" varName="gender">

```

- Para una variable numérica, hay un atributo de `número` en lugar de un atributo de `cadena`. El atributo de `etiqueta` está presente sólo si la variable o los valores tienen etiquetas definidas.
- Los elementos `<cell>` que contienen valores de casilla para los números contienen el atributo de `texto` y uno o más valores de atributo adicionales. A continuación se muestra un ejemplo:

```
<cell text="45.6" number="45.569620253165" decimals="1"/>
```

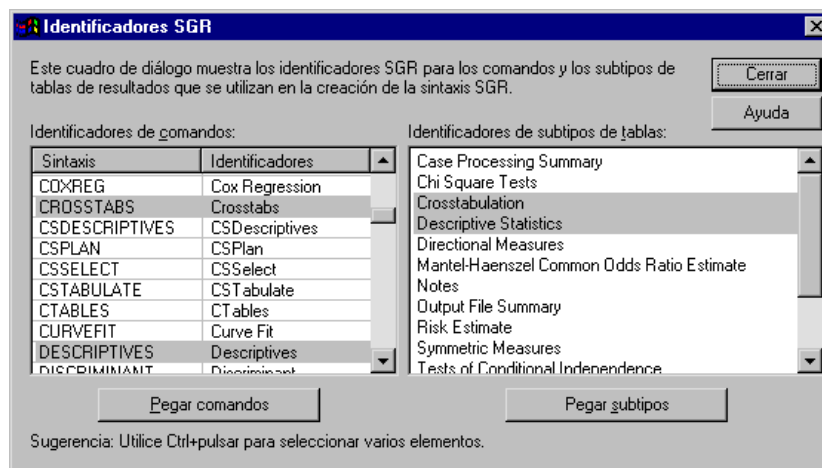
El atributo de `número` es el valor numérico real sin redondear, mientras que el atributo de `decimales` indica el número de decimales que se muestran en la tabla.

- Dado que las columnas se anidan en las filas, el elemento de categoría que identifica cada columna se repite para cada fila. Por ejemplo, dado que los estadísticos se muestran en las columnas, el elemento `<category text="Frequency">` aparece tres veces en XML: una vez para la fila de hombre, una vez para la fila de mujer y una vez para la fila total.

## Identificadores SGR

El objetivo del cuadro de diálogo Identificadores SGR es ofrecerle asistencia en la escritura de la sintaxis del comando OMS. Se puede utilizar este cuadro de diálogo para pegar los identificadores de subtipos y comandos seleccionados en una ventana de sintaxis de comandos.

Figura 48-17  
Cuadro de diálogo Identificadores SGR



### Para utilizar el cuadro de diálogo Identificadores SGR

- ▶ Elija en los menús:
  - Utilidades
  - Identificadores SGR...
- ▶ Seleccione uno o varios identificadores de comandos o de subtipos. (Mantenga pulsada la tecla Ctrl si desea seleccionar varios identificadores en cada lista.)
- ▶ Pulse en Pegar comandos y/o Pegar subtipos.
  - La lista de subtipos disponibles depende de los comandos seleccionados en ese momento. Si se seleccionan varios comandos, la lista de subtipos disponibles es la unión de todos los subtipos que están disponibles para cualquiera de los comandos seleccionados. Si no se selecciona ningún comando, en la lista aparecerán todos los subtipos.
  - Los identificadores se pegan en la posición actual del cursor dentro de la ventana de sintaxis de comandos designada. Si no hay abierta ninguna ventana de sintaxis de comandos, se abrirá automáticamente una nueva ventana de sintaxis.
  - Cuando se pega un identificador de subtipos y/o comandos, éste aparece entre comillas, ya que la sintaxis del comando OMS exige que así sea.
  - Las listas de identificadores de las palabras clave COMMANDS y SUBTYPES deben ir entre corchetes, por ejemplo:

```
/IF COMMANDS=['Crosstabs' 'Descriptives']
  SUBTYPES=['Crosstabulation' 'Descriptive Statistics']
```

### ***Copia de identificadores SGR desde los titulares del Visor***

Puede copiar y pegar identificadores de subtipos y comandos SGR desde el panel de titulares del Visor.

- ▶ Pulse con el botón derecho del ratón en la entrada del titular del elemento en el panel de titulares.
- ▶ Seleccione Copiar identificador de comandos de SGR o Copiar subtipo de tablas de SGR.

Este método presenta una diferencia respecto al del cuadro de diálogo Identificadores de SGR: el identificador copiado no se pega automáticamente en una ventana de sintaxis de comandos. Sólo tiene que copiar el identificador en el Portapapeles y, a continuación, podrá pegarlo donde desee. Como los valores de los identificadores de subtipos y comandos son idénticos a los correspondientes valores de los atributos de subtipos y comandos del formato XML de resultados (OXML), este método de copiar y pegar puede resultar muy útil para escribir transformaciones XSLT.

### ***Copia de etiquetas SGR***

En vez de identificadores, puede copiar etiquetas para utilizarlas con la palabra clave LABELS. Las etiquetas se pueden utilizar para diferenciar varios gráficos o varias tablas del mismo tipo en las que el texto del titular refleja algún atributo del objeto de resultados concreto como las etiquetas o los nombres de las variables. Hay, no obstante, ciertos factores que pueden afectar al texto de la etiqueta:

- Si el procesamiento de segmentación del archivo está activado, es posible que se añada a la etiqueta una identificación con el grupo de segmentación del archivo.
- Las etiquetas que incluyen información acerca de variables o valores se ven afectadas por la configuración de la presentación de nombres de variables/etiquetas y valores/etiquetas de valor del panel de titulares (menú Edición, Opciones, pestaña Etiquetas de los resultados).
- Las etiquetas se ven afectadas por el ajuste actual del idioma de los resultados (menú Edición, Opciones, pestaña General).

### ***Para copiar etiquetas SGR***

- ▶ Pulse con el botón derecho del ratón en la entrada del titular del elemento en el panel de titulares.
- ▶ Seleccione Copiar etiqueta de SGR.

Al igual que ocurría con los identificadores de subtipos y comandos, las etiquetas deben ir entre comillas y toda la lista debe ir entre corchetes, por ejemplo:

```
/IF LABELS=['Employment Category' 'Education Level']
```



- acceso a un servidor, 71
- Access (Microsoft), 20
- adición de etiquetas de grupo, 239
- agregación de datos, 189
  - agregación de funciones, 192
  - nombres y etiquetas de variable, 192
- agrupación, 127
- agrupación de categorías, 127
- agrupación de filas o columnas, 239
- ajuste, 246
  - control de la anchura de la columna para el texto ajustado, 246
  - etiquetas de variable y de valor, 86
- alfa de Cronbach
  - en Análisis de fiabilidad, 444, 446
- algoritmos, 12
- alineación, 88, 219, 481
  - en el Editor de datos, 88
  - resultados, 219, 481
- almacenamiento de archivos, 45–46
  - archivos de datos, 45–46
  - archivos de datos de SPSS, 45
  - consultas del archivo de base de datos, 31
  - control de ubicaciones de archivos por defecto, 492
- almacenamiento de gráficos, 224, 230–231
  - archivos BMP, 224, 231
  - archivos EMF, 224
  - archivos EPS, 224, 231
  - archivos JPEG, 224, 230
  - archivos PICT, 224
  - archivos PNG, 231
  - archivos PostScript, 231
  - archivos TIFF, 231
  - metarchivos, 224
- almacenamiento de resultados, 224, 227, 229
  - en formato PDF, 224, 227
  - formato de texto, 224, 229
  - formato Excel, 224, 226
  - formato HTML, 224
  - Formato PowerPoint, 224, 227
  - formato Word, 224, 226
  - HTML, 224, 226
- almacenamiento en la caché, 68
  - archivo activo, 68
- análisis alfa, 380
- análisis de componentes principales, 377, 380
- análisis de conglomerados
  - Análisis de conglomerados de K-medias, 401
  - análisis de conglomerados jerárquico, 395
  - eficacia, 403
- Análisis de conglomerados de K-medias
  - almacenamiento de información de conglomerados, 404
  - conceptos básicos, 401
  - conglomerado de pertenencia, 404
  - criterios de convergencia, 403
  - distancias entre conglomerados, 404
  - eficacia, 403
  - ejemplos, 401
  - estadísticos, 401, 404
  - funciones adicionales del comando, 405
  - iteraciones, 403
  - métodos, 401
  - valores perdidos, 404
- Análisis de conglomerados en dos fases, 387
  - almacenamiento en el archivo de trabajo, 393
  - almacenamiento en un archivo externo, 393
  - estadísticos, 393
  - gráficos, 392
  - opciones, 390
- análisis de conglomerados jerárquico, 395
  - almacenamiento de nuevas variables, 399
  - casos de conglomerado, 395
  - conglomerado de pertenencia, 398–399
  - dendrogramas, 399
  - diagramas de témpanos, 399
  - ejemplo, 395
  - estadísticos, 395, 398
  - funciones adicionales del comando, 400
  - historial de conglomerado, 398
  - matrices de distancias, 398
  - medidas de distancia, 397
  - medidas de similitud, 397
  - métodos de conglomeración, 397
  - orientación de los gráficos, 399
  - transformación de medidas, 397
  - transformación de valores, 397
  - variables de conglomerado, 395
- análisis de datos, 10
  - pasos básicos, 10
- Análisis de fiabilidad, 444
  - coeficiente de correlación intraclase, 446
  - correlaciones y covarianzas inter-elementos, 446
  - descriptivos, 446
  - ejemplo, 444
  - estadísticos, 444, 446
  - funciones adicionales del comando, 448
  - Kuder-Richardson 20, 446
  - Prueba de aditividad de Tukey, 446
  - tabla de ANOVA, 446
  - $\gamma$ cuadrado de Hotelling , 446

- análisis de la varianza
  - en ANOVA de un factor, 311
  - en Estimación curvilínea, 360
  - en Medias, 295
  - en regresión lineal, 350
- análisis de respuestas múltiples
  - Frecuencias de respuestas múltiples, 427
  - tabla de contingencia, 429
  - Tablas de contingencia de respuestas múltiples, 429
  - tablas de frecuencias, 427
- análisis de series temporales
  - predicción, 362
  - predicción de casos, 362
- Análisis discriminante, 369
  - almacenamiento de variables de clasificación, 376
  - coeficientes de la función, 372
  - criterios, 373
  - definición de rangos, 371
  - Distancia de Mahalanobis, 373
  - ejemplo, 369
  - estadísticos, 369, 372
  - estadísticos descriptivos, 372
  - exportación de información del modelo, 376
  - funciones adicionales del comando, 376
  - gráficos, 374
  - lambda de Wilks, 373
  - matrices, 372
  - matriz de covarianzas, 374
  - métodos de inclusión por pasos, 369
  - métodos discriminantes, 373
  - opciones de presentación, 373–374
  - probabilidades previas, 374
  - selección de casos, 371
  - V de Rao , 373
  - valores perdidos, 374
  - variables de agrupación, 369
  - variables independientes, 369
- Análisis factorial, 377
  - conceptos básicos, 377
  - convergencia, 380–381
  - descriptivos, 379
  - ejemplo, 377
  - estadísticos, 377, 379
  - formato de presentación de los coeficientes, 383
  - funciones adicionales del comando, 383
  - gráficos de saturaciones, 381
  - métodos de extracción, 380
  - métodos de rotación, 381
  - puntuaciones factoriales, 382
  - selección de casos, 378
  - valores perdidos, 383
- análisis imagen, 380
- ancho de columna, 88, 246, 257, 490
  - control de la anchura máxima, 246
  - control de la anchura para el texto ajustado, 246
  - control de la anchura por defecto, 490
  - en el Editor de datos, 88
  - tablas pivote, 257
- años, 482
  - valores de dos dígitos, 482
- ANOVA
  - en ANOVA de un factor, 311
  - en Medias, 295
  - en MLG Univariante, 318
  - modelo, 320
- ANOVA de un factor, 311
  - comparaciones múltiples, 313
  - contrastes, 312
  - contrastes polinómicos, 312
  - contrastes post hoc, 313
  - estadísticos, 316
  - funciones adicionales del comando, 317
  - opciones, 316
  - valores perdidos, 316
  - variables del factor, 311
- apertura de archivos, 15–19, 32
  - archivos de datos, 15–16
  - archivos de datos de texto, 32
  - archivos de dBASE, 15, 18
  - archivos de Excel, 15, 17
  - archivos de hoja de cálculo, 15, 17
  - archivos de Lotus 1-2-3, 15
  - archivos de Stata, 18
  - archivos delimitados por tabuladores, 15
  - archivos SYSTAT, 15
  - control de ubicaciones de archivos por defecto, 492
- archivo activo, 67–68
  - almacenamiento en la caché, 68
  - archivo activo virtual, 67
  - creación de un archivo activo temporal, 68
- archivo activo temporal, 68
- archivo activo virtual, 67
- archivo de diario, 492
- archivo de modelo
  - carga de modelos guardados para puntuar datos, 176
- archivos, 221
  - adición de un archivo de texto al Visor, 221
  - apertura, 15
- archivos BMP, 224, 231
  - exportación de gráficos, 224, 231
- archivos de datos, 15–16, 32, 44–46, 52, 68, 75, 199
  - adición de comentarios, 474
  - almacenamiento, 45–46
  - almacenamiento de resultados como archivos de datos
    - con formato SPSS, 511
  - almacenamiento de subconjuntos de variables, 52
  - apertura, 15–16
  - Dimensions, 41
  - información sobre el archivo, 44
  - información sobre el diccionario, 44
  - mejora del rendimiento para archivos grandes, 68
  - mrInterview, 41
  - protección, 66
  - Quancept, 41

- Quanvert, 41
- reestructuración, 199
- servidores remotos, 75
- texto, 32
- transposición, 182
- varios archivos de datos abiertos, 102, 479
- volteado, 182
- archivos de dBASE, 15, 18, 45–46
  - almacenamiento, 45–46
  - lectura, 15, 18
- archivos de Excel, 15, 17, 46, 496
  - adición de elementos de menú para enviar datos a Excel, 496
  - almacenamiento, 45–46
  - almacenamiento de etiquetas de valor en lugar de valores, 45
  - apertura, 15, 17
- archivos de hoja de cálculo, 15, 17, 48
  - apertura, 17
  - escritura de nombres de variable, 48
  - lectura de nombres de variables, 17
  - rangos de lectura, 17
- archivos de Lotus 1-2-3, 15, 45–46, 496
  - adición de elementos de menú para enviar datos a Lotus, 496
  - almacenamiento, 45–46
  - apertura, 15
- archivos de SAS
  - almacenamiento, 45
  - apertura, 15
- archivos de sintaxis de comandos, 265
- archivos de sintaxis del comando Unicode, 265
- archivos de Stata, 18
  - almacenamiento, 45
  - apertura, 18
  - lectura, 15
- archivos delimitados por comas, 32
- archivos delimitados por tabuladores, 15, 17, 32, 45–46, 48
  - almacenamiento, 45–46
  - apertura, 15
  - escritura de nombres de variable, 48
  - lectura de nombres de variables, 17
- archivos EPS, 224, 231
  - exportación de gráficos, 224, 231
- archivos JPEG, 224, 230
  - exportación de gráficos, 224, 230
- archivos PNG, 224, 231
  - exportación de gráficos, 224, 231
- archivos portátiles
  - nombres de variable, 46
- archivos PostScript (encapsulado), 224, 231
  - exportación de gráficos, 224, 231
- archivos spp
  - conversión a archivos spj, 510
- archivos SYSTAT, 15
  - apertura, 15
- archivos TIFF, 231
  - exportación de gráficos, 224, 231
- Asesor estadístico, 10
- asignación de memoria
  - en Análisis de conglomerados en dos fases, 390
- asignación de rangos a los casos, 148
  - percentiles, 149
  - puntuaciones de Savage, 149
  - rangos fraccionales, 149
  - valores empatados, 150
- asimetría
  - en Cubos OLAP, 299
  - en descriptivos, 273
  - en el Informe de estadísticos en columnas, 440
  - en el Informe de estadísticos en filas, 435
  - en explorar, 277
  - en Frecuencias, 268
  - en Medias, 295
  - en Resumir, 290
- asociación lineal por lineal
  - en Tablas de contingencia, 283
- Aspectos de tabla, 244–245
  - aplicación, 244
  - creación, 245
- atributos
  - atributos de variable personalizados, 89
- atributos de variable, 88–89
  - copia y pegado, 88–89
  - personalizados, 89
- atributos de variable personalizados, 89
- atributos personalizados, 89
- autoprosesos, 493
- autovalores
  - en Análisis factorial, 379–380
  - en regresión lineal, 350
- ayuda en pantalla, 12
  - Asesor estadístico, 10
- barra de estado, 6
- barras de herramientas, 497, 499–500
  - creación, 497, 499
  - creación de nuevas herramientas, 500
  - personalización, 497, 499
  - presentación en distintas ventanas, 499
  - presentación y ocultación, 497
- bases de datos, 19–20, 22, 24–25, 27, 29, 31
  - actualización, 52
  - adición de nuevos campos a una tabla, 61
  - adición de nuevos registros (casos) a una tabla, 62
  - almacenamiento, 52
  - almacenamiento de consultas, 31
  - cláusula Where, 25
  - consultas de parámetros, 25, 27
  - conversión de cadenas en variables numéricas, 29
  - creación de relaciones, 24
  - creación de una tabla nueva, 63
  - definición de variables, 29

- especificación de criterios, 25
- expresiones condicionales, 25
- lectura, 19–20, 22
- Microsoft Access, 20
- muestreo aleatorio, 25
- Pedir el valor al usuario, 27
- selección de campos de datos, 22
- selección de un origen de datos, 20
- sintaxis de SQL, 31
- sustitución de los valores de los campos existentes, 60
- sustitución de una tabla, 63
- uniones entre tablas, 24
- verificación de los resultados, 31
- bondad de ajuste
  - en regresión ordinal, 356
- Bonferroni
  - en ANOVA de un factor, 313
  - en MLG, 324
- bordes, 250, 257
  - presentación de bordes ocultos, 257
- botón Aceptar, 8
- botón Ayuda, 8
- botón Cancelar, 8
- botón Pegar, 8
- botón Restablecer, 8
- buscar y reemplazar
  - documentos del visor, 222
- C de Dunnett
  - en ANOVA de un factor, 313
  - en MLG, 324
- cálculo de variables, 136
  - cálculo de variables de cadena nuevas, 138
- cambio del nombre de los conjuntos de datos, 105
- cambio del orden de filas y columnas, 238
- capas, 232, 240, 242, 246, 251
  - creación, 240
  - en las tablas pivote, 240
  - en Tablas de contingencia, 283
  - impresión, 232, 246, 251
  - presentación, 240, 242
- casillas de las tablas pivote, 248, 256–257
  - anchuras, 257
  - formatos, 248
  - ocultación, 243
  - presentación, 243
- casos, 97, 199
  - búsqueda de duplicados, 124
  - búsqueda en el Editor de datos, 98
  - inserción de nuevos casos, 97
  - ordenación, 180
  - ponderación, 198
  - reestructuración en variables, 199
  - selección de subconjuntos, 194, 196–197
- casos duplicados (registros)
  - búsqueda y filtrado, 124
- casos filtrados, 100
  - en el Editor de datos, 100
- categoría de referencia
  - en MLG, 322
- categorización, 127
- Categorizador visual, 127
- CCI. Véase coeficiente de correlación intraclase, 446
- centrado de resultados, 219, 481
- chi-cuadrado, 406
  - asociación lineal por lineal, 283
  - corrección por continuidad de Yates, 283
  - en Tablas de contingencia, 283
  - estadísticos, 408
  - opciones, 408
  - para la independencia, 283
  - Pearson, 283
  - prueba exacta de Fisher, 283
  - prueba para una muestra, 406
  - rango esperado, 408
  - razón de verosimilitud, 283
  - valores esperados, 408
  - valores perdidos, 408
- chi-cuadrado de la razón de verosimilitud
  - en regresión ordinal, 356
  - en Tablas de contingencia, 283
- chi-cuadrado de Pearson
  - en regresión ordinal, 356
  - en Tablas de contingencia, 283
- clasificación
  - en Curva COR..., 459
- codificación local, 265
- coeficiente alfa
  - en Análisis de fiabilidad, 444, 446
- coeficiente de contingencia
  - en Tablas de contingencia, 283
- coeficiente de correlación de los rangos
  - en correlaciones bivariadas, 331
- coeficiente de correlación de Spearman
  - en correlaciones bivariadas, 331
  - en Tablas de contingencia, 283
- coeficiente de correlación intraclase (CCI)
  - en Análisis de fiabilidad, 446
- coeficiente de correlación  $r$ 
  - en correlaciones bivariadas, 331
  - en Tablas de contingencia, 283
- coeficiente de dispersión (CDD)
  - en Estadísticos de la razón, 457
- coeficiente de incertidumbre
  - en Tablas de contingencia, 283
- coeficiente de variación (CDV)
  - en Estadísticos de la razón, 457
- coeficientes beta
  - en regresión lineal, 350
- coeficientes de regresión
  - en regresión lineal, 350
- color de fondo, 253

- colores en las tablas pivote, 250
  - bordes, 250
- columna total
  - en informes, 441
- columnas, 257–258
  - cambio de la anchura en tablas pivote, 257
  - selección en tablas pivote, 258
- comparación de grupos
  - en Cubos OLAP, 302
- comparación de variables
  - en Cubos OLAP, 302
- comparaciones múltiples
  - en ANOVA de un factor, 313
- comparaciones múltiples post hoc, 313
- conglomeración
  - selección de procedimientos, 385
- conjuntos de datos
  - cambio del nombre, 105
- conjuntos de respuestas múltiples
  - categorías múltiples, 114
  - definición, 114
  - dicotomías múltiples, 114
- conjuntos de variables, 474–475
  - definición, 474
  - utilización, 475
- construcción de términos, 320, 359
- contrastes
  - en ANOVA de un factor, 312
  - en MLG, 322
- contrastes de desviación
  - en MLG, 322
- contrastes de diferencia
  - en MLG, 322
- Contrastes de Helmert
  - en MLG, 322
- contrastes de linealidad
  - en Medias, 295
- contrastes polinómicos
  - en ANOVA de un factor, 312
  - en MLG, 322
- contrastes repetidos
  - en MLG, 322
- contrastes simples
  - en MLG, 322
- control de página
  - en el informe de estadísticos en columnas, 442
  - en informes de estadísticos en filas, 436
- convergencia
  - en Análisis de conglomerados de K-medias, 403
  - en Análisis factorial, 380–381
- corrección por continuidad de Yates
  - en Tablas de contingencia, 283
- correlación de Pearson
  - en correlaciones bivariadas, 331
- Correlación de Pearson
  - en Tablas de contingencia, 283
- correlaciones
  - de orden cero, 336
    - en correlaciones bivariadas, 331
    - en Correlaciones parciales, 335
    - en Tablas de contingencia, 283
- correlaciones bivariadas
  - coeficientes de correlación, 331
  - estadísticos, 333
  - funciones adicionales del comando, 333
  - nivel de significación, 331
  - opciones, 333
  - valores perdidos, 333
- correlaciones de orden cero
  - en Correlaciones parciales, 336
- correlaciones parciales
  - en regresión lineal, 350
- Correlaciones parciales, 335
  - correlaciones de orden cero, 336
  - estadísticos, 336
  - funciones adicionales del comando, 337
  - opciones, 336
  - valores perdidos, 336
- cuadros de diálogo, 8, 474–475, 479
  - controles, 8
  - definición de conjuntos de variables, 474
  - iconos de variable, 9
  - información sobre la variable, 9
  - orden de presentación de variables, 479
  - presentación de etiquetas de variable, 7, 479
  - presentación de nombres de variable, 7, 479
  - reordenación de listas de destino, 477
  - selección de variables, 8
  - utilización de conjuntos de variables, 475
  - variables, 7
- cuartiles
  - en Frecuencias, 268
- Cubos OLAP, 298
  - estadísticos, 299
  - títulos, 303
- curtosis
  - en Cubos OLAP, 299
  - en descriptivos, 273
  - en el Informe de estadísticos en columnas, 440
  - en el Informe de estadísticos en filas, 435
  - en explorar, 277
  - en Frecuencias, 268
  - en Medias, 295
  - en Resumir, 290
- Curva COR, 459
  - estadísticos y gráficos, 460
- d*
  - en Tablas de contingencia, 283
- d* de Somers
  - en Tablas de contingencia, 283
- DATA LIST, 67
  - frente al comando GET DATA, 67

- datos categóricos, 111
  - conversión de datos de intervalo en categorías discretas, 127
- Datos de Dimensions, 41
  - almacenamiento, 65
- datos de serie temporal
  - creación de nuevas variables de serie temporal, 171
  - definición de variables de datos, 170
  - funciones de transformación, 172
  - sustitución de valores perdidos, 174
  - transformaciones de los datos, 169
- datos delimitados por espacios, 32
- datos ponderados, 215
  - y de datos reestructurados, 215
- definición de variables, 81, 83, 85–89, 108
  - aplicación de un diccionario de datos, 117
  - copia y pegado de atributos, 88–89
  - etiquetas de valor, 86, 108
  - etiquetas de variable, 85
  - plantillas, 88–89
  - tipos de datos, 83
  - valores perdidos, 87
- definir conjuntos de respuestas múltiples, 426
  - categorías, 426
  - dicotomías, 426
  - etiquetas del conjunto, 426
  - nombres del conjunto, 426
- dendrogramas
  - en Análisis de conglomerados jerárquico, 399
- descomposición jerárquica, 321
- descriptivos, 272
  - almacenamiento de puntuaciones  $z$ , 272
  - estadísticos, 273
  - funciones adicionales del comando, 274
  - orden de presentación, 273
- desplazamiento de filas y columnas, 238
- desviación absoluta promedio (DAP)
  - en Estadísticos de la razón, 457
- desviación típica
  - en Cubos OLAP, 299
  - en descriptivos, 273
  - en el Informe de estadísticos en columnas, 440
  - en el Informe de estadísticos en filas, 435
  - en Estadísticos de la razón, 457
  - en explorar, 277
  - en Frecuencias, 268
  - en Medias, 295
  - en MLG Univariante, 328
  - en Resumir, 290
- DfAjuste
  - en regresión lineal, 347
- DfBeta
  - en regresión lineal, 347
- diagnósticos de colinealidad
  - en regresión lineal, 350
- diagnósticos por caso
  - en regresión lineal, 350
- diagramas de caja
  - comparación de niveles del factor, 278
  - comparación de variables, 278
  - en explorar, 278
- diagramas de dispersión
  - en regresión lineal, 346
- diagramas de dispersión por nivel
  - en explorar, 278
  - en MLG Univariante, 328
- diagramas de témpanos
  - en Análisis de conglomerados jerárquico, 399
- diario de la sesión, 492
- diccionario, 44
- diccionario de datos
  - aplicar desde otro archivo, 117
- diferencia honestamente significativa de Tukey
  - en ANOVA de un factor, 313
  - en MLG, 324
- diferencia menos significativa
  - en MLG, 324
- diferencia mínima significativa
  - en ANOVA de un factor, 313
- diferencial relativo al precio (DRP)
  - en Estadísticos de la razón, 457
- diferencias entre grupos
  - en Cubos OLAP, 302
- diferencias entre variables
  - en Cubos OLAP, 302
- directorio temporal, 492
  - definir ubicación en modo local, 492
  - variable de entorno SPSSTMPDIR, 492
- distancia chi-cuadrado
  - en Distancias, 340
- distancia de bloques
  - en Distancias, 340
- distancia de Chebychev
  - en Distancias, 340
- distancia de Cook
  - en regresión lineal, 347
- Distancia de Cook
  - en MLG, 326
- distancia de Mahalanobis
  - en regresión lineal, 347
- Distancia de Mahalanobis
  - en Análisis discriminante, 373
- distancia de Minkowski
  - en Distancias, 340
- Distancia euclídea
  - en Distancias, 340
- distancia euclídea al cuadrado
  - en Distancias, 340
- Distancias, 338
  - cálculo de distancias entre casos, 338
  - cálculo de distancias entre variables, 338
  - ejemplo, 338
  - estadísticos, 338
  - funciones adicionales del comando, 342

- medidas de disimilaridad, 340
- medidas de similaridad, 341
- transformación de medidas, 340–341
- transformación de valores, 340–341
- división
  - división entre columnas del informe, 441
- divisor de paneles
  - Editor de datos, 100
- divisor de ventanas
  - Editor de datos, 100
- DMS de Fisher
  - en MLG, 324
- edición de datos, 95–96
- Editor de datos, 79, 81, 88, 94–98, 100–101, 496
  - alineación, 88
  - ancho de columna, 88
  - cambiar tipo de datos, 98
  - casos filtrados, 100
  - definición de variables, 81
  - desplazamiento de variables, 98
  - edición de datos, 95–96
  - envío de datos a otras aplicaciones, 496
  - impresión, 101
  - inserción de nuevas variables, 97
  - inserción de nuevos casos, 97
  - introducción de datos, 94
  - introducción de datos no numéricos, 95
  - introducción de datos numéricos, 94
  - opciones de presentación, 100
  - restricciones de los valores de datos, 95
  - varias vistas/paneles, 100
  - varios archivos de datos abiertos, 102, 479
  - Vista de datos, 79
  - Vista de variables, 80
- Editor de gráficos, 466
  - propiedades, 467
- eliminación de etiquetas de grupo, 239
- eliminación de resultados, 218
- eliminación de varios comandos EXECUTE en archivos de sintaxis, 266
- eliminación hacia atrás
  - en regresión lineal, 344
- encabezados, 233
- enlace
  - en regresión ordinal, 354
- entrada de datos, 94
- error típico
  - en Curva COR..., 460
  - en descriptivos, 273
  - en explorar, 277
  - en Frecuencias, 268
  - en MLG, 326, 328
- error típico de la asimetría
  - en Cubos OLAP, 299
  - en Medias, 295
  - en Resumir, 290
- error típico de la curtosis
  - en Cubos OLAP, 299
  - en Medias, 295
  - en Resumir, 290
- error típico de la media
  - en Cubos OLAP, 299
  - en Medias, 295
  - en Resumir, 290
- escala, 83
  - en Análisis de fiabilidad, 444
  - en escalamiento multidimensional, 449
  - nivel de medida, 83, 111
- escalamiento
  - tablas pivote, 246, 251
- escalamiento multidimensional, 449
  - condicionalidad, 452
  - creación de matrices de distancias, 451
  - criterios, 453
  - definición de la forma de los datos, 451
  - dimensiones, 452
  - ejemplo, 449
  - estadísticos, 449
  - funciones adicionales del comando, 454
  - medidas de distancia, 451
  - modelos de escalamiento, 452
  - niveles de medida, 452
  - opciones de presentación, 453
  - transformación de valores, 451
- espacio en disco, 67–68
  - temporal, 67–68
- espacio temporal en disco, 67–68
- estadístico de Brown-Forsythe
  - en ANOVA de un factor, 316
- estadístico de Cochran
  - en Tablas de contingencia, 283
- estadístico de Mantel-Haenszel
  - en Tablas de contingencia, 283
- estadístico de Welch
  - en ANOVA de un factor, 316
- estadístico Durbin-Watson
  - en regresión lineal, 350
- estadístico *R*
  - en Medias, 295
  - en regresión lineal, 350
- Estadísticos de la razón, 455
  - estadísticos, 457
- estadísticos descriptivos
  - en Análisis de conglomerados en dos fases, 393
  - en descriptivos, 272
  - en Estadísticos de la razón, 457
  - en explorar, 277
  - en Frecuencias, 268
  - en MLG Univariante, 328
  - en Resumir, 290
- Estimación curvilínea, 360
  - almacenamiento de intervalos de pronóstico, 362
  - almacenamiento de residuos, 362

- almacenamiento de valores pronosticados, 362
- análisis de la varianza, 360
- inclusión de constante, 360
- modelos, 362
- predicción, 362
- estimaciones de Blom, 149
- estimaciones de los parámetros
  - en MLG Univariante, 328
  - en regresión ordinal, 356
- estimaciones de potencia
  - en MLG Univariante, 328
- estimaciones de proporción
  - en la asignación de rangos a los casos, 149
- estimaciones de Rankit, 149
- estimaciones de tamaño de efecto
  - en MLG Univariante, 328
- estimaciones de Tukey, 149
- estimaciones de Van der Waerden, 149
- estimador bponderado de Tukey
  - en explorar, 277
- estimador en onda de Andrews
  - en explorar, 277
- estimador-M de Huber
  - en explorar, 277
- estimador-M redescendente de Hampel
  - en explorar, 277
- estimadores robustos centrales
  - en explorar, 277
- estudio de control de casos
  - Prueba T para muestras relacionadas, 307
- estudio de pares relacionados
  - en Prueba T para muestras relacionadas, 307
- eta
  - en Medias, 295
  - en Tablas de contingencia, 283
- eta-cuadrado
  - en Medias, 295
  - en MLG Univariante, 328
- etiquetas, 239
  - eliminación, 239
  - frente a nombres de subtipos en SGR, 516
  - inserción de etiquetas de grupo, 239
- etiquetas de grupo, 239
- etiquetas de valor, 86, 95, 100, 108, 485
  - almacenamiento en archivos de Excel, 45
  - aplicación a varias variables, 113
  - copia, 113
  - en archivos de datos fusionados, 186
  - en el Editor de datos, 100
  - en el panel de titulares, 485
  - en las tablas pivote, 485
  - inserción de saltos de línea, 86
  - uso para entrada de datos, 95
- etiquetas de variable, 85, 479, 485
  - de los cuadros de diálogo, 7, 479
  - en archivos de datos fusionados, 186
  - en el panel de titulares, 485
  - en las tablas pivote, 485
  - inserción de saltos de línea, 86
- exclusión de resultados del Visor con SGR, 522
- EXECUTE (comando)
  - pegado desde cuadros de diálogo, 266
- explorar, 276
  - estadísticos, 277
  - funciones adicionales del comando, 280
  - gráficos, 278
  - opciones, 279
  - transformaciones de potencia, 279
  - valores perdidos, 279
- exportación de datos, 496
  - adición de elementos de menú para exportar datos, 496
- exportación de gráficos, 224, 230–231, 502
  - producción automatizada, 502
- exportación de resultados, 224, 227, 229
  - en formato PDF, 224, 227
  - formato Excel, 224, 226
  - formato HTML, 224
  - Formato PowerPoint, 224
  - formato Word, 224, 226
  - HTML, 226
  - SGR, 511
- exportar datos, 45
- F* múltiple de Ryan-Einot-Gabriel-Welsch
  - en ANOVA de un factor, 313
  - en MLG, 324
- factor de inflación de la varianza
  - en regresión lineal, 350
- factorización de ejes principales, 380
- fiabilidad de dos mitades
  - en Análisis de fiabilidad, 444, 446
- fiabilidad de Spearman-Brown
  - en Análisis de fiabilidad, 446
- filas, 258
  - selección en tablas pivote, 258
- formato
  - columnas en informes, 435
- formato CSV
  - almacenamiento de datos, 46
  - lectura de datos, 32
- Formato de archivo de datos de SPSS
  - envío de resultados a un archivo de datos, 517, 523
- formato de archivo SAV
  - envío de resultados a un archivo de datos de SPSS, 517, 523
- formato de cadena, 83
- formato de COMA, 83, 85
- formato de DÓLAR, 83, 85
- formato de PUNTO, 83, 85
- formato Excel
  - exportación de resultados, 224, 226
- formato fijo, 32
- formato libre, 32
- formato numérico, 83, 85



- Formato PowerPoint
  - exportación de resultados, 224
- formato Word
  - exportación de resultados, 224, 226
  - tablas anchas, 224
- formatos de entrada, 85
- formatos de fecha
  - años de dos dígitos, 482
- formatos de moneda, 484
- formatos de moneda personalizados, 83, 484
- formatos de presentación, 85
- Frecuencias, 267
  - estadísticos, 268
  - formatos, 271
  - gráficos, 270
  - orden de presentación, 271
  - supresión de tablas, 271
- frecuencias acumuladas
  - en regresión ordinal, 356
- frecuencias de los conglomerados
  - en Análisis de conglomerados en dos fases, 393
- Frecuencias de respuestas múltiples, 427
  - valores perdidos, 427
- frecuencias esperadas
  - en regresión ordinal, 356
- frecuencias observadas
  - en regresión ordinal, 356
- fuentes, 100, 221, 253
  - en el Editor de datos, 100
  - en el panel de titulares, 221
- función de adelanto, 172
- función de diferencia, 172
- función de diferencia estacional, 172
- función de media móvil anterior, 172
- función de media móvil centrada, 172
- función de medianas móviles, 172
- función de suavizado, 172
- función de suma acumulada, 172
- funciones, 139
  - tratamiento de los valores perdidos, 139
- fusión de archivos de datos
  - archivos con casos distintos, 184
  - archivos con distintas variables, 187
  - cambio de nombre de las variables, 186
  - información sobre el diccionario, 186
  
- gamma
  - en Tablas de contingencia, 283
- gamma de Goodman y Kruskal
  - en Tablas de contingencia, 283
- generación de gráficos, 462
- Generador de gráficos, 462
  - galería, 463
- GET DATA, 67
  - frente al comando DATA LIST, 67
  - frente al comando GET CAPTURE, 67
- gráfico de tabla, 260
  - gráficos, 218, 224, 260, 462, 487
    - conceptos básicos, 462
    - copia en otras aplicaciones, 224
    - creación, 462
    - creación a partir de tablas pivote, 260
    - en Curva COR..., 459
    - etiquetas de caso, 360
    - exportación, 224
    - Generador de gráficos, 462
    - ocultación, 218
    - paneles ajustados, 470
    - plantillas, 470, 487
    - relación de aspecto, 487
    - tamaño, 470
    - valores perdidos, 470
  - gráficos de barras
    - en Frecuencias, 270
  - gráficos de importancia
    - en Análisis de conglomerados en dos fases, 392
  - gráficos de la importancia de las variables
    - en Análisis de conglomerados en dos fases, 392
  - gráficos de los residuos
    - en MLG Univariante, 328
  - gráficos de perfil
    - en MLG, 323
  - gráficos de probabilidad normal
    - en explorar, 278
    - en regresión lineal, 346
  - gráficos de saturaciones
    - en Análisis factorial, 381
  - gráficos de sectores
    - en Frecuencias, 270
  - gráficos de tallo y hojas
    - en explorar, 278
  - gráficos interactivos, 224
    - copia en otras aplicaciones, 224
  - gráficos normales sin tendencia
    - en explorar, 278
  - gráficos parciales
    - en regresión lineal, 346
- GT2 de Hochberg
  - en ANOVA de un factor, 313
  - en MLG, 324
  
- H* de Kruskal-Wallis
  - en Pruebas para dos muestras independientes, 420
- histogramas
  - en explorar, 278
  - en Frecuencias, 270
  - en regresión lineal, 346
- historial de iteraciones
  - en regresión ordinal, 356
- HTML, 224, 226
  - exportación de resultados, 224, 226
  
- iconos
  - de los cuadros de diálogo, 9

- identificadores de comandos, 515
- importación de datos, 15, 19
- impresión, 101, 232–233, 235, 246, 251, 259
  - capas, 232, 246, 251
  - control de las rupturas de tabla, 259
  - datos, 101
  - encabezados y pies, 233
  - escalamiento de tablas, 246, 251
  - espacio entre los elementos de resultados, 235
  - gráficos, 232
  - números de páginas, 235
  - presentación preliminar, 232
  - resultados de texto, 232
  - tablas pivote, 232
  - tamaño del gráfico, 235
- índice de concentración
  - en Estadísticos de la razón, 457
- información sobre el archivo, 44
- información sobre la variable, 473
- informe de estadísticos en columnas, 438
- Informe de estadísticos en columnas, 438
  - columnas totales, 441
  - control de página, 442
  - diseño de página, 437
  - formato de columnas, 435
  - funciones adicionales del comando, 443
  - numeración de páginas, 442
  - subtotales, 442
  - total final, 442
  - valores perdidos, 442
- Informe de estadísticos en filas, 433
  - columnas de datos, 433
  - columnas de ruptura, 433
  - control de página, 436
  - diseño de página, 437
  - espaciado de ruptura, 436
  - formato de columnas, 435
  - funciones adicionales del comando, 443
  - numeración de páginas, 436
  - ordenación de secuencias, 433
  - pies, 438
  - títulos, 438
  - valores perdidos, 436
  - variables en los títulos, 438
- informes
  - columnas totales, 441
  - comparación de columnas, 441
  - división de valores de las columnas, 441
  - informe de estadísticos en columnas, 438
  - informes de estadísticos en filas, 433
  - multiplicación de valores de las columnas, 441
  - totales compuestos, 441
- inserción de etiquetas de grupo, 239
- intervalos de confianza
  - almacenamiento en regresión lineal, 347
  - en ANOVA de un factor, 316
  - en Curva COR..., 460
  - en explorar, 277
  - en MLG, 322, 328
  - en Prueba T para muestras relacionadas, 308
  - en Prueba t para una muestra, 310
  - en Pruebas t para muestras independientes, 306
  - en regresión lineal, 350
- intervalos de pronóstico
  - almacenamiento en Estimación curvilínea, 362
  - almacenamiento en regresión lineal, 347
- introducción de datos, 94–95
  - no numéricos, 95
  - numérica, 94
  - uso de etiquetas de valor, 95
- iteraciones
  - en Análisis de conglomerados de K-medias, 403
  - en Análisis factorial, 380–381
- justificación, 219, 481
  - resultados, 219, 481
- kappa
  - en Tablas de contingencia, 283
- kappa de Cohen
  - en Tablas de contingencia, 283
- KR20
  - en Análisis de fiabilidad, 446
- Kuder-Richardson 20 (KR20)
  - en Análisis de fiabilidad, 446
- LAG (función), 172
- lambda
  - en Tablas de contingencia, 283
- lambda de Goodman y Kruskal
  - en Tablas de contingencia, 283
- lambda de Wilks
  - en Análisis discriminante, 373
- lengua
  - cambio del idioma de los resultados, 479
- lenguaje de comandos, 261
- líneas de cuadrícula, 257
  - tablas pivote, 257
- listado de casos, 288
- listas de destino, 477
- listas de variables, 477
  - reordenación de listas de destino, 477
- matriz de configuración
  - en Análisis factorial, 377
- matriz de correlaciones
  - en Análisis discriminante, 372
  - en Análisis factorial, 377, 379
  - en regresión ordinal, 356
- matriz de covarianzas
  - en Análisis discriminante, 372, 374

- en MLG, 326
- en regresión lineal, 350
- en regresión ordinal, 356
- matriz de transformación
  - en Análisis factorial, 377
- máxima verosimilitud
  - en Análisis factorial, 380
- máximo
  - comparación de columnas del informe, 441
  - en Cubos OLAP, 299
  - en descriptivos, 273
  - en Estadísticos de la razón, 457
  - en explorar, 277
  - en Frecuencias, 268
  - en Medias, 295
  - en Resumir, 290
- media
  - de varias columnas del informe, 441
  - en ANOVA de un factor, 316
  - en Cubos OLAP, 299
  - en descriptivos, 273
  - en el Informe de estadísticos en columnas, 440
  - en el Informe de estadísticos en filas, 435
  - en Estadísticos de la razón, 457
  - en explorar, 277
  - en Frecuencias, 268
  - en Medias, 295
  - en Resumir, 290
  - subgrupo, 293, 298
- media armónica
  - en Cubos OLAP, 299
  - en Medias, 295
  - en Resumir, 290
- media geométrica
  - en Cubos OLAP, 299
  - en Medias, 295
  - en Resumir, 290
- media ponderada
  - en Estadísticos de la razón, 457
- media recortada
  - en explorar, 277
- mediana
  - en Cubos OLAP, 299
  - en Estadísticos de la razón, 457
  - en explorar, 277
  - en Frecuencias, 268
  - en Medias, 295
  - en Resumir, 290
- mediana agrupada
  - en Cubos OLAP, 299
  - en Medias, 295
  - en Resumir, 290
- Medias, 293
  - estadísticos, 295
  - opciones, 295
- medias de grupo, 293, 298
- medias de subgrupo, 293, 298
- medias marginales estimadas
  - en MLG Univariante, 328
- medias observadas
  - en MLG Univariante, 328
- medida de diferencia de configuración
  - en Distancias, 340
- medida de diferencia de tamaño
  - en Distancias, 340
- medida de disimilaridad de Lance y Williams, 340
  - en Distancias, 340
- medida de distancia de phi cuadrado
  - en Distancias, 340
- medidas de dispersión
  - en descriptivos, 273
  - en Estadísticos de la razón, 457
  - en explorar, 277
  - en Frecuencias, 268
- medidas de distancia
  - en Análisis de conglomerados jerárquico, 397
  - en Distancias, 340
- medidas de distribución
  - en descriptivos, 273
  - en Frecuencias, 268
- medidas de similitud
  - en Análisis de conglomerados jerárquico, 397
  - en Distancias, 341
- medidas de tendencia central
  - en Estadísticos de la razón, 457
  - en explorar, 277
  - en Frecuencias, 268
- memoria, 479
- menús, 496
  - personalización, 496
- metarchivos, 224
  - exportación de gráficos, 224
- métodos de selección, 258
  - selección de filas y columnas en tablas pivote, 258
- Microsoft Access, 20
- mínimo
  - comparación de columnas del informe, 441
  - en Cubos OLAP, 299
  - en descriptivos, 273
  - en Estadísticos de la razón, 457
  - en explorar, 277
  - en Frecuencias, 268
  - en Medias, 295
  - en Resumir, 290
- mínimos cuadrados generalizados
  - en Análisis factorial, 380
- mínimos cuadrados no ponderados
  - en Análisis factorial, 380
- mínimos cuadrados ponderados
  - en regresión lineal, 343
- MLG
  - almacenamiento de matrices, 326
  - almacenamiento de variables, 326
  - contrastes post hoc, 324

- gráficos de perfil, 323
- modelo, 320
- suma de cuadrados, 320
- MLG Univariante, 318, 329
  - contrastes, 322
  - diagnósticos, 328
  - medias marginales estimadas, 328
  - opciones, 328
  - presentación, 328
- moda
  - en Frecuencias, 268
- modelo compuesto
  - en Estimación curvilínea, 362
- modelo cuadrático
  - en Estimación curvilínea, 362
- modelo cúbico
  - en Estimación curvilínea, 362
- modelo de crecimiento
  - en Estimación curvilínea, 362
- modelo de curva S
  - en Estimación curvilínea, 362
- modelo de escala
  - en regresión ordinal, 358
- modelo de Guttman
  - en Análisis de fiabilidad, 444, 446
- modelo de potencia
  - en Estimación curvilínea, 362
- modelo de ubicación
  - en regresión ordinal, 357
- modelo estrictamente paralelo
  - en Análisis de fiabilidad, 444, 446
- modelo exponencial
  - en Estimación curvilínea, 362
- modelo inverso
  - en Estimación curvilínea, 362
- modelo lineal
  - en Estimación curvilínea, 362
- modelo logarítmico
  - en Estimación curvilínea, 362
- modelo logístico
  - en Estimación curvilínea, 362
- modelo paralelo
  - en Análisis de fiabilidad, 444, 446
- modelos factoriales completos
  - en MLG, 320
- modelos personalizados
  - en MLG, 320
- modificadores de la línea de comando, 509
  - Trabajos de producción, 509
- modo distribuido, 71–72, 75, 77
  - acceso a un archivo de datos, 75
  - procedimientos disponibles, 77
  - rutas relativas, 77
- mrInterview, 41
- muestra aleatoria, 25
  - bases de datos, 25
  - selección, 196
  - semilla de aleatorización, 140
- muestras relacionadas, 418, 423
- muestreo
  - muestra aleatoria, 196
- multiplicación
  - multiplicación entre columnas del informe, 441
- Newman-Keuls
  - en MLG, 324
- nivel de medida, 83, 111
  - definición, 83
  - iconos de los cuadros de diálogo, 9
- nombres de variable, 82, 479
  - ajuste de los nombres largos de variable en los resultados, 82
  - archivos portátiles, 46
  - de los cuadros de diálogo, 7, 479
  - generados por SGR, 530
  - nombres de variable de casos mixtos, 82
  - reglas, 82
  - truncado de nombres de variable largos en versiones anteriores, 46
- nominal, 83
  - nivel de medida, 83, 111
- notación científica, 83, 479
  - supresión en resultados, 479
- notas al pie, 247, 255–257
  - gráficos, 469
  - marcadores, 247, 256
  - renumerar, 257
- numeración de páginas, 235
  - en el informe de estadísticos en columnas, 442
  - en informes de estadísticos en filas, 436
- número de casos
  - en Cubos OLAP, 299
  - en Medias, 295
  - en Resumir, 290
- número máximo de ramas
  - en Análisis de conglomerados en dos fases, 390
- números de puerto, 72
- ocultación, 218, 243–244, 256, 497
  - barras de herramientas, 497
  - etiquetas de dimensión, 244
  - filas y columnas, 243
  - notas al pie, 256
  - resultados de un procedimiento, 218
  - textos al pie, 256
  - títulos, 244
- ocultación (exclusión) de resultados del Visor con SGR, 522
- opciones, 479, 481–482, 484–485, 487, 490, 492–493
  - años de dos dígitos, 482
  - aspecto de tablas pivote, 490
  - datos, 482
  - directorio temporal, 492
  - etiquetas de los resultados, 485

- generales, 479
- gráficos, 487
- moneda, 484
- procesos, 493
- Visor, 481
- Vista de variables, 484
- opciones del gráfico, 487
- orden de presentación, 238
- ordenación
  - variables, 181
- ordenación de casos, 180
- ordenación de variables, 181
- ordinal, 83
  - nivel de medida, 83, 111
- ortografía, 94
  - diccionario, 482
- OXML, 536
  
- pares de variables, 199
  - creación, 199
- pasos básicos, 10
- PDF
  - exportación de resultados, 224, 227
- percentiles
  - en explorar, 277
  - en Frecuencias, 268
- phi
  - en Tablas de contingencia, 283
- pies, 233
- pivotado
  - control con SGR para resultados exportados, 528
- plantillas, 88–89, 470, 487
  - definición de variables, 88–89
  - en los gráficos, 487
  - gráficos, 470
  - uso de un archivo de datos externo como plantilla, 117
- PLUM
  - en regresión ordinal, 353
- ponderación de casos, 198
  - ponderaciones fraccionarias en Tablas de contingencia, 198
- porcentajes
  - en Tablas de contingencia, 286
- porcentajes de fila
  - en Tablas de contingencia, 286
- porcentajes de la columna
  - en Tablas de contingencia, 286
- porcentajes totales
  - en Tablas de contingencia, 286
- PowerPoint, 227
  - exportación de resultados como PowerPoint, 227
- predicción
  - en Estimación curvilínea, 362
- preparar página, 233, 235
  - encabezados y pies, 233
  - tamaño del gráfico, 235
- presentación, 218, 243–244, 256, 497
  - barras de herramientas, 497
  - etiquetas de dimensión, 244
  - filas o columnas, 243
  - notas al pie, 256
  - resultados, 218
  - títulos, 244
- primera
  - en Cubos OLAP, 299
  - en Medias, 295
  - en Resumir, 290
- procesamiento de segmentación del archivo, 193
- procesos, 496, 500
  - adición a menús, 496
  - ejecución mediante los botones de la barra de herramientas, 500
  - lenguaje por defecto, 493
- producción automatizada, 502
- profundidad del árbol
  - en Análisis de conglomerados en dos fases, 390
- programación con lenguaje de comandos, 261
- propiedades, 246
  - tablas, 246
  - tablas pivote, 246
- propiedades de casilla, 253–254
- Proximidades
  - en Análisis de conglomerados jerárquico, 395
- prueba *b* de Tukey
  - en ANOVA de un factor, 313
  - en MLG, 324
- Prueba binomial, 409
  - dicotomías, 409
  - estadísticos, 410
  - funciones adicionales del comando, 411
  - opciones, 410
  - valores perdidos, 410
- Prueba de aditividad de Tukey
  - en Análisis de fiabilidad, 444, 446
- Prueba de comparación por parejas de Gabriel
  - en ANOVA de un factor, 313
  - en MLG, 324
- Prueba de comparación por parejas de Games y Howell
  - en ANOVA de un factor, 313
  - en MLG, 324
- prueba de esfericidad de Bartlett
  - en Análisis factorial, 379
- prueba de Friedman
  - en pruebas para varias muestras relacionadas, 423
- prueba de homogeneidad marginal
  - en Pruebas para dos muestras relacionadas, 418
- Prueba de Kolmogorov-Smirnov para una muestra, 413
  - distribución de contraste, 413
  - estadísticos, 414
  - funciones adicionales del comando, 415
  - opciones, 414
  - valores perdidos, 414

- prueba de la mediana
  - en Pruebas para dos muestras independientes, 420
- prueba de Levene
  - en ANOVA de un factor, 316
  - en explorar, 278
  - en MLG Univariante, 328
- prueba de Lilliefors
  - en explorar, 278
- prueba de líneas paralelas
  - en regresión ordinal, 356
- prueba de los signos
  - en Pruebas para dos muestras relacionadas, 418
- prueba de McNemar
  - en Pruebas para dos muestras relacionadas, 418
  - en Tablas de contingencia, 283
- Prueba de rachas
  - estadísticos, 412
  - funciones adicionales del comando, 413
  - opciones, 412
  - puntos de corte, 411–412
  - valores perdidos, 412
- prueba de rangos múltiples de Duncan
  - en ANOVA de un factor, 313
  - en MLG, 324
- prueba de reacciones extremas de Moses
  - en Pruebas para dos muestras independientes, 416
- prueba de Scheffé
  - en MLG, 324
- prueba de Shapiro-Wilk
  - en explorar, 278
- prueba de Wilcoxon de los rangos con signo
  - en Pruebas para dos muestras relacionadas, 418
- prueba exacta de Fisher
  - en Tablas de contingencia, 283
- Prueba M de Box
  - en Análisis discriminante, 372
- prueba *t*
  - en MLG Univariante, 328
  - en Prueba T para muestras relacionadas, 307
  - en Prueba *t* para una muestra, 309
  - en Pruebas *t* para muestras independientes, 304
- prueba *t* de Dunnett
  - en ANOVA de un factor, 313
  - en MLG, 324
- prueba *t* de Sidak
  - en ANOVA de un factor, 313
  - en MLG, 324
- prueba *t* de Student, 304
- prueba *t* de Waller-Duncan
  - en ANOVA de un factor, 313
  - en MLG, 324
- prueba *t dependiente*
  - en Prueba T para muestras relacionadas, 307
- prueba *t* para dos muestras
  - en Pruebas *t* para muestras independientes, 304
- Prueba T para muestras independientes, 304
  - definición de grupos, 306
  - intervalos de confianza, 306
  - opciones, 306
  - valores perdidos, 306
  - variables de agrupación, 306
  - variables de cadena, 306
- Prueba T para muestras relacionadas, 307
  - opciones, 308
  - selección de variables relacionadas, 307
  - valores perdidos, 308
- Prueba T para una muestra, 309
  - funciones adicionales del comando, 310
  - intervalos de confianza, 310
  - opciones, 310
  - valores perdidos, 310
- pruebas de homogeneidad de las varianzas
  - en ANOVA de un factor, 316
  - en MLG Univariante, 328
- pruebas de normalidad
  - en explorar, 278
- pruebas no paramétricas
  - chi-cuadrado, 406
  - Prueba de Kolmogorov-Smirnov para una muestra, 413
  - Prueba de rachas, 411
  - Pruebas para dos muestras independientes, 415
  - Pruebas para dos muestras relacionadas, 418
  - Pruebas para varias muestras independientes, 420
  - Pruebas para varias muestras relacionadas, 423
- Pruebas para dos muestras independientes, 415
  - definición de grupos, 417
  - estadísticos, 417
  - funciones adicionales del comando, 418
  - opciones, 417
  - tipos de pruebas, 416
  - valores perdidos, 417
  - variables de agrupación, 417
- Pruebas para dos muestras relacionadas, 418
  - estadísticos, 420
  - funciones adicionales del comando, 420
  - opciones, 420
  - tipos de pruebas, 419
  - valores perdidos, 420
- pruebas para la independencia
  - chi-cuadrado, 283
- Pruebas para varias muestras independientes, 420
  - definición de rango, 422
  - estadísticos, 422
  - funciones adicionales del comando, 422
  - opciones, 422
  - tipos de pruebas, 421
  - valores perdidos, 422
  - variables de agrupación, 422
- Pruebas para varias muestras relacionadas, 423
  - estadísticos, 424
  - funciones adicionales del comando, 424
  - tipos de pruebas, 423
- puntuación
  - carga de modelos guardados, 176

- modelos admitidos para exportar y puntuar, 175
- visualización de modelos cargados, 179
- puntuaciones de Savage, 149
- puntuaciones factoriales, 382
- puntuaciones factoriales de Anderson-Rubin, 382
- puntuaciones factoriales de Bartlett, 382
- puntuaciones normales
  - en la asignación de rangos a los casos, 149
- puntuaciones  $z$ 
  - almacenamiento como variables, 272
  - en descriptivos, 272
  - en la asignación de rangos a los casos, 149
- $Q$  de Cochran
  - en pruebas para varias muestras relacionadas, 423
- Quancept, 41
- Quanvert, 41
- $R$  múltiple
  - en regresión lineal, 350
- R-E-G-W  $F$ 
  - en ANOVA de un factor, 313
  - en MLG, 324
- R-E-G-W  $Q$ 
  - en ANOVA de un factor, 313
  - en MLG, 324
- $R^2$ 
  - Cambio en  $R^2$ , 350
  - en Medias, 295
  - en regresión lineal, 350
- $R^2$  ajustado
  - en regresión lineal, 350
- $R^2$  de Cox y Snell
  - en regresión ordinal, 356
- $R^2$  de McFadden
  - en regresión ordinal, 356
- $R^2$  de Nagelkerke
  - en regresión ordinal, 356
- Rachas de Wald-Wolfowitz
  - en Pruebas para dos muestras independientes, 416
- rango
  - en Cubos OLAP, 299
  - en descriptivos, 273
  - en Estadísticos de la razón, 457
  - en Frecuencias, 268
  - en Medias, 295
  - en Resumir, 290
- rango múltiple de Ryan-Einot-Gabriel-Welsch
  - en ANOVA de un factor, 313
  - en MLG, 324
- razón entre covarianzas
  - en regresión lineal, 347
- recodificación de los valores, 127, 143–144, 146–147, 151
- recuento de apariciones, 141
- recuento esperado
  - en Tablas de contingencia, 286
- recuento observado
  - en Tablas de contingencia, 286
- reestructuración de los datos, 199, 202–204, 206–211, 213–215
  - conceptos básicos, 199
  - creación de una única variable de índice para variables a casos, 208
  - creación de variables de índice para variables a casos, 206
  - creación de varias variables de índice para variables a casos, 209
  - ejemplo de casos a variables, 203
  - ejemplo de dos índices para variables a casos, 208
  - ejemplo de un índice para variables a casos, 207
  - ejemplo de variables a casos, 202
  - grupos de variables para variables a casos, 203
  - opciones para casos a variables, 214
  - opciones para variables a casos, 210
  - ordenación de los datos para reestructurar casos a variables, 213
  - selección de datos para reestructurar casos a variables, 211
  - selección de datos para reestructurar variables a casos, 204
  - tipos de reestructuración, 199
  - y datos ponderados, 215
- regresión
  - gráficos, 346
  - regresión lineal, 343
  - regresión múltiple, 343
- Regresión de mínimos cuadrados parciales, 364
  - exportar variables, 367
  - modelo, 366
- regresión lineal, 343
  - almacenamiento de nuevas variables, 347
  - bloques, 343
  - estadísticos, 350
  - exportación de información del modelo, 347
  - funciones adicionales del comando, 352
  - gráficos, 346
  - métodos de selección de variables, 344, 351
  - ponderaciones, 343
  - residuos, 347
  - valores perdidos, 351
  - variable de selección, 346
- regresión múltiple
  - en regresión lineal, 343
- regresión ordinal, 353
  - enlace, 354
  - estadísticos, 353
  - funciones adicionales del comando, 359
  - modelo de escala, 358
  - modelo de ubicación, 357
  - opciones, 354
- relación de aspecto, 487
- rendimiento, 68
  - caché de datos, 68

- residuos
  - almacenamiento en Estimación curvilínea, 362
  - almacenamiento en regresión lineal, 347
  - en Tablas de contingencia, 286
- residuos de Pearson
  - en regresión ordinal, 356
- residuos eliminados
  - en MLG, 326
  - en regresión lineal, 347
- residuos estudentizados
  - en regresión lineal, 347
- residuos no tipificados
  - en MLG, 326
- residuos tipificados
  - en MLG, 326
  - en regresión lineal, 347
- Respuesta múltiple
  - funciones adicionales del comando, 432
- resultados, 217–219, 223–224, 236, 481
  - alineación, 219, 481
  - almacenamiento, 236
  - cambio del idioma de los resultados, 479
  - centrado, 219, 481
  - copia, 218
  - copia en otras aplicaciones, 224
  - desplazamiento, 218
  - eliminación, 218
  - exportación, 224
  - ocultación, 218
  - pegado en otras aplicaciones, 223
  - presentación, 218
  - Visor, 217
- resultados de gráficos, 224
- Resumir, 288
  - estadísticos, 290
  - opciones, 290
- rho
  - en correlaciones bivariadas, 331
  - en Tablas de contingencia, 283
- riesgo
  - en Tablas de contingencia, 283
- riesgo relativo
  - en Tablas de contingencia, 283
- rotación de etiquetas, 239
- rotación equamax
  - en Análisis factorial, 381
- rotación oblimin directa
  - en Análisis factorial, 381
- rotación quartimax
  - en Análisis factorial, 381
- rotación varimax
  - en Análisis factorial, 381
- rupturas de tabla, 259
- saltos de línea
  - etiquetas de variable y de valor, 86
- Scheffé test
  - en ANOVA de un factor, 313
- selección de casos, 194
  - muestra aleatoria, 196
  - rango de casos, 197
  - rango de fechas, 197
  - rango de horas, 197
  - según criterios de selección, 196
- selección hacia delante
  - en regresión lineal, 344
- selección por pasos
  - en regresión lineal, 344
- seleccionar casos, 194
- semilla de aleatorización, 140
- separación de tablas, 259
  - control de las rupturas de tabla, 259
- servidores, 71–72
  - acceso, 71
  - adición, 72
  - edición, 72
  - nombres, 72
  - números de puerto, 72
- servidores remotos, 71–72, 75, 77
  - acceso, 71
  - acceso a un archivo de datos, 75
  - adición, 72
  - edición, 72
  - procedimientos disponibles, 77
  - rutas relativas, 77
- SGR, 511, 535
  - control de pivotes de tabla, 517, 528
  - exclusión de resultados del Visor, 522
  - Formato de archivo de datos de SPSS, 517, 523
  - formato de archivo SAV, 517, 523
  - formato de texto, 517
  - identificadores de comandos, 515
  - nombres de variable en los archivos SAV, 530
  - subtipos de tabla, 515
  - tipos de objetos de resultados, 514
  - Uso de XSLT con OXML, 536
  - XML, 517, 531
- sintaxis, 261, 265, 499, 502
  - acceso a *SPSS Command Syntax Reference*, 12
  - anotaciones de los resultados, 264
  - archivo de diario, 266
  - archivos de sintaxis del comando Unicode, 265
  - ejecución, 265
  - ejecución de sintaxis de comandos mediante botones de las barras de herramientas, 499
  - pegado, 263
  - Reglas de los trabajos de producción, 502
  - reglas de sintaxis, 261
  - sintaxis de comandos, 261, 265, 496, 500, 502
  - acceso a *SPSS Command Syntax Reference*, 12
  - adición a menú, 496
- S-stress
  - en escalamiento multidimensional, 449



- anotaciones de los resultados, 264
- archivo de diario, 266
- ejecución, 265
- ejecución mediante los botones de la barra de herramientas, 500
- pegado, 263
- Reglas de los trabajos de producción, 502
- reglas de sintaxis, 261
- Sistema de gestión de resultados (SGR), 511, 535
- sistema de medida, 479
- stress
  - en escalamiento multidimensional, 449
- Student-Newman-Keuls
  - en ANOVA de un factor, 313
  - en MLG, 324
- suavizado T4253H, 172
- subconjuntos de casos
  - muestra aleatoria, 196
  - selección, 194, 196–197
- subtipos, 515
  - frente a etiquetas, 516
- subtipos de tabla, 515
  - frente a etiquetas, 516
- subtítulos
  - gráficos, 469
- subtotales
  - en el informe de estadísticos en columnas, 442
- suma
  - en Cubos OLAP, 299
  - en descriptivos, 273
  - en Frecuencias, 268
  - en Medias, 295
  - en Resumir, 290
- suma de cuadrados, 321
  - en MLG, 320
- sustitución de valores perdidos
  - interpolación lineal, 175
  - media de la serie, 175
  - media de los puntos adyacentes, 175
  - mediana de los puntos adyacentes, 175
  - tendencia lineal, 175
- T2 de Tamhane
  - en ANOVA de un factor, 313
  - en MLG, 324
- T3 de Dunnett
  - en ANOVA de un factor, 313
  - en MLG, 324
- tabla de claves, 187
- tabla de contingencia
  - en Tablas de contingencia, 281
  - respuesta múltiple, 429
- tablas, 259
  - alineación, 254
  - color de fondo, 253
  - control de las rupturas de tabla, 259
  - fuentes, 253
  - márgenes, 254
  - propiedades de casilla, 253–254
- tablas anchas
  - pegado en Microsoft Word, 223
- tablas de contingencia, 281
- Tablas de contingencia, 281
  - capas, 283
  - estadísticos, 283
  - formatos, 287
  - gráficos de barras agrupadas, 283
  - ponderaciones fraccionarias, 198
  - presentación de casillas, 286
  - supresión de tablas, 281
  - variables de control, 283
- Tablas de contingencia de respuestas múltiples, 429
  - definición de rangos de valores, 430
  - emparejamiento de las variables entre los conjuntos de respuestas, 431
  - porcentajes basados en casos, 431
  - porcentajes basados en respuestas, 431
  - porcentajes de casilla, 431
  - valores perdidos, 431
- tablas de frecuencias
  - en explorar, 277
  - en Frecuencias, 267
- tablas pivote, 218, 223–224, 232, 237–240, 243–244, 246–248, 250–251, 257–260, 490
  - agrupación de filas o columnas, 239
  - ajustes en la anchura de columna por defecto, 490
  - alineación, 254
  - anchuras de casillas, 257
  - aspecto por defecto para las tablas nuevas, 490
  - bordes, 250
  - cambio del aspecto, 244
  - cambio del orden de presentación, 238
  - capas, 240
  - color de fondo, 253
  - control de las rupturas de tabla, 259
  - copia en otras aplicaciones, 224
  - creación de gráficos a partir de tablas, 260
  - desagrupación de filas o columnas, 239
  - desplazamiento de filas y columnas, 238
  - edición, 237
  - eliminación de etiquetas de grupo, 239
  - escalamiento para ajustarse a la página, 246, 251
  - exportación como HTML, 224
  - formatos de casilla, 248
  - fuentes, 253
  - gestión, 237
  - impresión de capas, 232
  - impresión de tablas grandes, 259
  - inserción de etiquetas de grupo, 239
  - líneas de cuadrícula, 257
  - márgenes, 254
  - notas al pie, 255–256
  - ocultación, 218
  - pegado como tablas, 223

- pegado en otras aplicaciones, 223
- pivotado, 237
- presentación de bordes ocultos, 257
- presentación y ocultación de casillas, 243
- propiedades, 246
- propiedades de casilla, 253–254
- propiedades de las notas al pie, 247
- propiedades generales, 246
- rotación de etiquetas, 239
- selección de filas y columnas, 258
- texto de continuación, 251
- textos al pie, 255–256
- transposición de filas y columnas, 239
- utilizando iconos, 237
- tamaños, 221
  - en titulares, 221
- tau de Goodman y Kruskal
  - en Tablas de contingencia, 283
- tau de Kruskal
  - en Tablas de contingencia, 283
- tau-*b*
  - en Tablas de contingencia, 283
- Tau-*b* de Kendall
  - en correlaciones bivariadas, 331
  - en Tablas de contingencia, 283
- tau-*c*
  - en Tablas de contingencia, 283
- Tau-*c* de Kendall , 283
  - en Tablas de contingencia, 283
- T<sup>2</sup>cuadrado de Hotelling
  - en Análisis de fiabilidad, 444, 446
- términos de interacción, 320, 359
- texto, 32, 221, 224, 229
  - adición al Visor, 221
  - adición de un archivo de texto al Visor, 221
  - archivos de datos, 32
  - exportación de resultados como texto, 224, 229
- texto de continuación, 251
  - para tablas pivote, 251
- texto de etiqueta vertical, 239
- textos al pie, 255–256
- tipificación
  - en Análisis de conglomerados en dos fases, 390
- tipos de datos, 83, 85, 98, 484
  - cambio, 98
  - definición, 83
  - formatos de entrada, 85
  - formatos de presentación, 85
  - moneda personalizada, 83, 484
- tipos de objetos de resultados
  - en SGR, 514
- titulares, 219–220
  - cambio de los niveles, 220
  - contracción, 220
  - en el Visor, 219
  - expansión, 220
- títulos, 221
  - adición al Visor, 221
  - en Cubos OLAP, 303
  - gráficos, 469
- tolerancia
  - en regresión lineal, 350
- totales finales
  - en el informe de estadísticos en columnas, 442
- Trabajos de producción, 502, 507–509
  - archivos de resultados, 502
  - conversión de los archivos de la unidad de producción, 510
  - ejecución de varios trabajos de producción, 509
  - exportación de gráficos, 502
  - modificadores de la línea de comando, 509
  - programación de trabajos de producción, 509
  - reglas de sintaxis, 502
  - sustitución de valores en archivos de sintaxis, 507
- transformaciones condicionales, 138
- transformaciones de archivos, 199
  - agregación de datos, 189
  - fusión de archivos de datos, 184, 187
  - ordenación de casos, 180
  - ponderación de casos, 198
  - procesamiento de segmentación del archivo, 193
  - reestructuración de los datos, 199
  - transposición de variables y casos, 182
- transformaciones de los datos, 482
  - asignación de rangos a los casos, 148
  - cálculo de variables, 136
  - funciones, 139
  - recodificación de los valores, 143–144, 146–147, 151
  - retraso de la ejecución, 482
  - transformaciones condicionales, 138
  - variables de cadena, 138
  - variables de series temporales, 169, 171
- transposición de filas y columnas, 239
- transposición de variables y casos, 182
- tratamiento del ruido
  - en Análisis de conglomerados en dos fases, 390
- U* de Mann-Whitney
  - en Pruebas para dos muestras independientes, 416
- ubicaciones de archivos
  - control de ubicaciones de archivos por defecto, 492
  - ubicaciones de archivos por defecto, 492
- última
  - en Cubos OLAP, 299
  - en Medias, 295
  - en Resumir, 290
- umbral inicial
  - en Análisis de conglomerados en dos fases, 390
- Unicode, 15, 45
- Unidad de producción, 479
  - conversión de archivos a trabajos de producción, 510
  - uso de la sintaxis de comandos de archivos de diario, 479
- unión exterior, 24

- unión interior, 24
- V
- en Tablas de contingencia, 283
  - V de Cramér
    - en Tablas de contingencia, 283
  - V de Rao
    - en Análisis discriminante, 373
  - valores atípicos
    - en Análisis de conglomerados en dos fases, 390
    - en explorar, 277
    - en regresión lineal, 346
  - valores de influencia
    - en MLG, 326
    - en regresión lineal, 347
  - valores definidos por el usuario como perdidos, 87
  - valores extremos
    - en explorar, 277
  - valores perdidos, 87, 470
    - definición, 87
    - en Análisis factorial, 383
    - en ANOVA de un factor, 316
    - en correlaciones bivariadas, 333
    - en Correlaciones parciales, 336
    - en Curva COR..., 460
    - en el informe de estadísticos en columnas, 442
    - en el Informe de estadísticos en filas, 436
    - en explorar, 279
    - en la prueba de chi-cuadrado, 408
    - en las frecuencias de respuestas múltiples, 427
    - en las funciones, 139
    - en las tablas de contingencia de respuestas múltiples, 431
    - en Prueba binomial, 410
    - en Prueba de Kolmogorov-Smirnov para una muestra, 414
    - en Prueba de rachas, 412
    - en Prueba T para muestras relacionadas, 308
    - en Prueba t para una muestra, 310
    - en Pruebas para dos muestras independientes, 417
    - en Pruebas para dos muestras relacionadas, 420
    - en Pruebas para varias muestras independientes, 422
    - en Pruebas t para muestras independientes, 306
    - en regresión lineal, 351
      - gráficos, 470
    - reemplazo en datos de serie temporal, 174
    - variables de cadena, 87
  - valores pronosticados
    - almacenamiento en Estimación curvilínea, 362
    - almacenamiento en regresión lineal, 347
  - valores pronosticados ponderados
    - en MLG, 326
  - valores tipificados
    - en descriptivos, 272
  - variable de entorno SPSSTMPDIR, 492
  - variable de selección
    - en regresión lineal, 346
  - variables, 8, 81, 97–98, 199, 473–474, 479
    - búsqueda en el Editor de datos, 98
    - cambio de nombre de los archivos de datos fusionados, 186
    - de los cuadros de diálogo, 7
    - definición, 81
    - definición de conjuntos de variables, 474
    - desplazamiento, 98
    - información sobre la definición, 473
    - información sobre variables en los cuadros de diálogo, 9
    - inserción de nuevas variables, 97
    - orden de presentación de los cuadros de diálogo, 479
    - ordenación, 181
    - recodificación, 143–144, 146–147, 151
    - reestructuración en casos, 199
    - selección en los cuadros de diálogo, 8
  - variables de agrupación, 199
    - creación, 199
  - variables de cadena, 87, 95
    - cálculo de variables de cadena nuevas, 138
    - de los cuadros de diálogo, 7
    - fragmentación de cadenas largas en versiones anteriores, 46
    - introducción de datos, 95
    - recodificación a valores enteros consecutivos, 151
    - valores perdidos, 87
  - variables de control
    - en Tablas de contingencia, 283
  - variables de entorno, 492
    - SPSSTMPDIR, 492
  - variables de escala
    - agrupación para la generación de variables categóricas, 127
  - variables de fecha
    - definición para datos de serie temporal, 170
  - variables de formato de fecha, 83, 85, 482
    - adición o sustracción de variables de fecha/hora, 154
    - creación de variable de fecha/hora a partir de conjunto de variables, 154
    - creación de variable de fecha/hora a partir de una cadena, 154
    - extracción de parte de variable de fecha/hora, 154
  - variables de segmentación
    - en Agregar datos, 189
  - varianza
    - en Cubos OLAP, 299
    - en descriptivos, 273
    - en el Informe de estadísticos en columnas, 440
    - en el Informe de estadísticos en filas, 435
    - en explorar, 277
    - en Frecuencias, 268
    - en Medias, 295
    - en Resumir, 290
  - varias vistas/paneles
    - Editor de datos, 100
  - varios archivos de datos abiertos, 102, 479
    - supresión, 106

- velocidad, 68
  - caché de datos, 68
- ventana activa, 6
- ventana designada, 6
- ventanas, 4
  - ventana activa, 6
  - ventana designada, 6
- ventanas de ayuda, 12
- visor
  - buscar y reemplazar información, 222
  - búsqueda y sustitución de información, 222
- Visor, 217–221, 235–236, 481, 485
  - almacenamiento de documentos, 236
  - cambio de las fuentes de los titulares, 221
  - cambio de los niveles de titulares, 220
  - cambio de los tamaños de los titulares, 221
  - contracción de titulares, 220
  - desplazamiento de los resultados, 218
  - eliminación de resultados, 218
  - espacio entre los elementos de resultados, 235
  - exclusión de tipos de resultados con SGR, 522
  - expansión de titulares, 220
  - muestra de los valores de datos, 485
  - ocultación de resultados, 218
  - opciones de presentación, 481
  - paneles de resultados, 217
  - paneles de titulares, 217
  - presentación de etiquetas de valor, 485
  - presentación de etiquetas de variable, 485
  - presentación de nombres de variable, 485
  - titulares, 219
- Vista de datos, 79
- Vista de variables, 80
  - personalización, 93, 484
- visualización
  - textos al pie, 256
  
- W de Kendall
  - en pruebas para varias muestras relacionadas, 423
  
- XML
  - almacenamiento de resultados como XML, 511
  - envío de resultados a XML, 517
  - estructura de tablas en OXML, 531
  - Resultados OXML del SGR, 536
- XSLT
  - uso con OXML, 536
  
- Z de Kolmogorov-Smirnov
  - en Prueba de Kolmogorov-Smirnov para una muestra, 413
  - en Pruebas para dos muestras independientes, 416