

IBM SPSS Statistics 32 Guía breve



Nota

Antes de utilizar esta información y el producto al que da soporte, lea la información del apartado “Avisos” en la página 91.

Información del producto

Esta edición se aplica a la versión 32, release 0, modificación 0 de IBM® SPSS Estadística y a todas las versiones y modificaciones posteriores hasta que se indique lo contrario en nuevas ediciones.

© Copyright International Business Machines Corporation .

Contenido

Capítulo 1. Introducción.....	1
Archivos de muestra.....	1
Abrir un archivo de datos.....	1
Ejecutar un análisis.....	2
Creación de gráficos.....	4
Capítulo 2. Leyendo datos.....	7
Estructura básica de archivos de datos de IBM SPSS Estadística	7
Lectura de archivos de datos de IBM SPSS Estadística	7
Lectura de datos de Excel.....	8
Lectura de datos de una base de datos.....	11
Lectura de datos desde un archivo de texto.....	14
Capítulo 3. Uso del Editor de datos.....	19
Especificar datos numéricos.....	19
Especificar datos de cadena.....	20
Definición de datos.....	21
Añadir etiquetas de variable.....	21
Cambiar el tipo y el formato de la variable.....	22
Añadir etiquetas de valores.....	23
Manejo de datos perdidos.....	23
Valores perdidos de una variable numérica.....	24
Valores perdidos de una variable de cadena.....	24
Capítulo 4. Examinar estadísticos de resumen de variables individuales.....	27
Nivel de medición.....	27
Medidas de resumen de datos categóricos.....	27
Gráficos de datos categóricos.....	28
Medidas de resumen de variables de escala.....	29
Histogramas de variables de escala.....	30
Capítulo 5. Creación y edición de gráficos.....	33
Fundamentos de la creación de gráficos.....	33
Utilizar la galería del Creador de gráficos.....	33
Definición de variables y estadísticas.....	34
Adición de texto.....	36
Crear el gráfico.....	36
Capítulo 6. Cómo trabajar con la salida.....	39
Uso del Visor.....	39
Utilización del editor de tablas dinámicas.....	40
Acceso a las definiciones de resultados.....	40
Tablas dinámicas.....	41
Creación y visualización de capas.....	42
Edición de tablas.....	42
Ocultación de filas y columnas.....	43
Cambiar los formatos de visualización de datos.....	44
TableLooks.....	45
Utilización de formatos predefinidos.....	45
Personalización de estilos de Aspecto de tabla.....	46

Cambio de los formatos de tablas predeterminadas.....	48
Personalizar los valores de visualización inicial.....	48
Visualización de etiquetas de variable y valor.....	49
Uso de resultados en otras aplicaciones.....	50
Pegar resultados como tablas de Word.....	51
Cómo pegar resultados como texto.....	51
Exportar resultados a archivos de Microsoft Word, PowerPoint y Excel.....	52
Exportar resultados a PDF.....	57
Exportación de resultados a HTML.....	59
Capítulo 7. Trabajar con sintaxis.....	61
Pegar sintaxis.....	61
Editar la sintaxis.....	62
Abrir y ejecutar un archivo de sintaxis.....	63
Utilizar puntos de interrupción.....	63
Capítulo 8. Modificar los valores de los datos.....	65
Crear una variable categórica a partir de una variable de escala.....	65
Cálculo de nuevas variables.....	67
Utilizar funciones en expresiones.....	68
Utilizar expresiones condicionales.....	69
Trabajar con fechas y horas.....	69
Cálculo de la longitud de tiempo entre dos fechas.....	70
Añadir una duración a una fecha.....	71
Capítulo 9. Ordenación y selección de datos.....	73
Ordenación de datos.....	73
procesamiento de archivos segmentados.....	73
Ordenación de casos para el procesamiento de archivos segmentados.....	75
Activación y desactivación del procesamiento de archivos segmentados.....	75
Selección de subconjuntos de casos.....	75
Para seleccionar subconjuntos de casos basados en una expresión condicional.....	76
Para seleccionar una muestra aleatoria de casos.....	77
Selección de un rango de tiempo o un rango de casos.....	78
Tratamiento de casos no seleccionados.....	78
Estado de selección de casos.....	79
Capítulo 10. Archivos de muestra.....	81
Avisos.....	91
Marcas comerciales.....	92
Índice.....	95

Capítulo 1. Introducción

Esta guía le mostrará cómo utilizar muchas de las características disponibles. Está diseñado para proporcionar una guía práctica paso a paso. Todos los ejemplos que aparecen en los ejemplos vienen instalados con la aplicación para que pueda seguirlos a la hora de realizar el mismo análisis y obtener los mismos resultados que aparecen aquí.

Si quiere ver ejemplos detallados de diversas técnicas de análisis estadísticos, acceda a los Estudios de casos, disponibles desde el menú Ayuda.

Archivos de muestra

La mayoría de los ejemplos que se ofrecen utilizan el archivo de datos *demo.sav*. Este archivo de datos es un estudio ficticio de varios miles de personas que contiene información básica demográfica y de consumo.

Si utiliza la versión para estudiantes, su versión de *demo.sav* será una muestra representativa del archivo de datos original, reducido para ajustarse al límite de 1.500 casos. Los resultados que obtenga utilizando ese archivo de datos diferirán de los resultados que se muestran aquí.

Los archivos muestrales instalados con el producto se encuentran en el subdirectorio *Samples* del directorio de instalación. En el subdirectorio *Samples* hay una carpeta distinta para cada uno de los idiomas siguientes: Inglés, francés, alemán, italiano, japonés, coreano, polaco, chino simplificado, español y chino tradicional.

No todos los archivos muestrales están disponibles en todos los idiomas. Si un archivo muestral no está disponible en un idioma, esa carpeta de idioma contendrá una versión en inglés del archivo muestral.

Abrir un archivo de datos

Para abrir un archivo de datos:

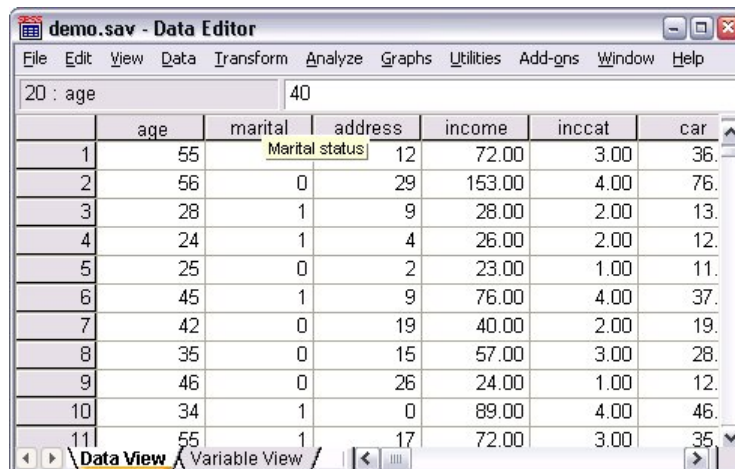
1. En los menús seleccione:

Archivo > Abrir > Datos...

Se muestra un recuadro de diálogo para abrir archivos.

De forma predeterminada, se muestran los archivos de datos de IBM SPSS Estadística (.*extensiónsav*).

Este ejemplo utiliza el archivo *demo.sav*.



	age	marital	address	income	inccat	car
1	55	12	72.00	3.00	36.	
2	56	0	29	153.00	4.00	76.
3	28	1	9	28.00	2.00	13.
4	24	1	4	26.00	2.00	12.
5	25	0	2	23.00	1.00	11.
6	45	1	9	76.00	4.00	37.
7	42	0	19	40.00	2.00	19.
8	35	0	15	57.00	3.00	28.
9	46	0	26	24.00	1.00	12.
10	34	1	0	89.00	4.00	46.
11	55	1	17	72.00	3.00	35.

Figura 1. Archivo *demo.sav* en el Editor de datos

El archivo de datos aparece en el Editor de datos. En el Visor de datos, si sitúa el cursor sobre un nombre de variable (las cabeceras de columnas), aparecerá una etiqueta de variable más descriptiva (si se ha definido para esa variable).

De forma predeterminada, aparecen los valores de los datos reales. Para visualizar las etiquetas:

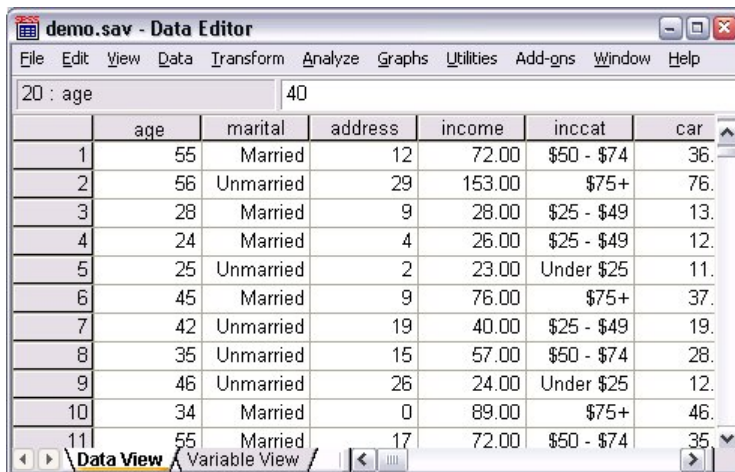
2. En los menús seleccione:

Vista > Etiquetas de valor



Figura 2. Botón Etiquetas de valor

De forma alternativa, puede utilizar el botón Etiquetas de valor de la barra de herramientas.



	age	marital	address	income	inccat	car
1	55	Married	12	72.00	\$50 - \$74	36.
2	56	Unmarried	29	153.00	\$75+	76.
3	28	Married	9	28.00	\$25 - \$49	13.
4	24	Married	4	26.00	\$25 - \$49	12.
5	25	Unmarried	2	23.00	Under \$25	11.
6	45	Married	9	76.00	\$75+	37.
7	42	Unmarried	19	40.00	\$25 - \$49	19.
8	35	Unmarried	15	57.00	\$50 - \$74	28.
9	46	Unmarried	26	24.00	Under \$25	12.
10	34	Married	0	89.00	\$75+	46.
11	55	Married	17	72.00	\$50 - \$74	35.

Figura 3. Etiquetas de valor en el Editor de datos

Ahora se muestran etiquetas de valor descriptivas para facilitar la interpretación de las respuestas.

Ejecutar un análisis

Si tiene alguna opción de complemento, el menú Analizar contiene una lista de categorías de informes y análisis estadísticos.

Para comenzar, crearemos una tabla de frecuencias simple (tabla de recuentos). Este ejemplo requiere la opción Statistics Base.

1. En los menús seleccione:

Analizar > Estadísticas descriptivas > Frecuencias...

Se visualiza el cuadro de diálogo Frecuencias.

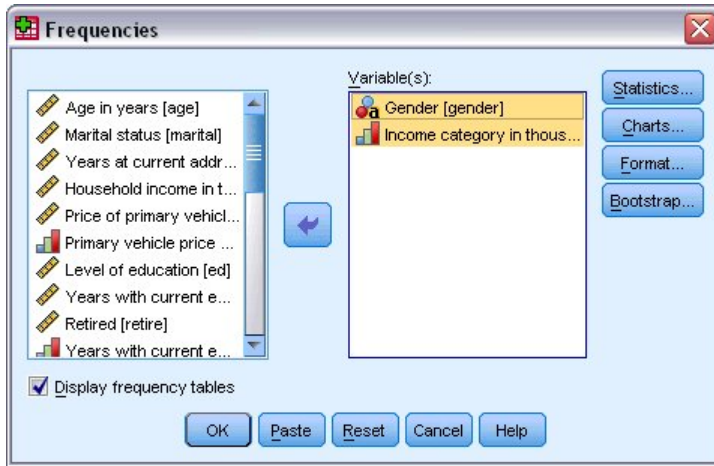


Figura 4. Cuadro de diálogo Frecuencias

Un icono junto a cada variable proporciona información sobre el tipo de datos y el nivel de medición.

	Numérico	Serie	Fecha	Hora
Escala (Continuo)		n/a		
Ordinal				
Nominal				

Si el nombre o la etiqueta de la variable aparecen truncadas en la lista, puede ver el texto completo del nombre o la etiqueta situando el cursor sobre ella. El nombre de la variable *cating* aparecerá entre corchetes después de la etiqueta de variable descriptiva. *Categoría de ingresos (en miles)* es la etiqueta de la variable. Si no hubiera ninguna etiqueta de variable, solo aparecería el nombre de la variable en el recuadro de lista.

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.

En el cuadro de diálogo, seleccione las variables que desea analizar de la lista de origen de la izquierda y arrástrelas y suéltelas en la lista Variable(s) de la derecha. El botón **Aceptar**, que ejecuta el análisis, está inhabilitado hasta que se coloca al menos una variable en la lista Variable(s).

En muchos cuadros de diálogo puede obtener más información si pulsa con el botón derecho del ratón en cualquier nombre de variable de la lista y selecciona **Información de variable** en el menú emergente.

2. Pulse *Sexo [gender]* en la lista de variables de origen y arrastre la variable a la lista de variables de destino.
3. Pulse *Categoría de ingresos en miles [inccat]* en la lista de origen y arrástrela a la lista de destino.

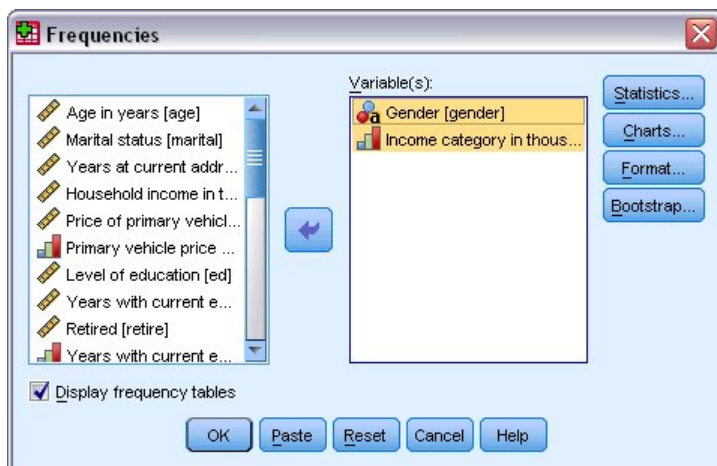


Figura 5. Variables seleccionadas para el análisis

4. Pulse **Aceptar** para ejecutar el procedimiento.

Los resultados se muestran en la ventana del Visor.

	Frequency	Percent	Valid Percent	Cumul: Perce
Valid Female	3179	49.7	49.7	
Male	3221	50.3	50.3	
Total	6400	100.0	100.0	

	Frequency	Percent	Valid Percent	Cum Pe
Valid Under \$25	1174	18.3	18.3	
\$25 - \$49	2388	37.3	37.3	
\$50 - \$74	1120	17.5	17.5	
\$75+	1718	26.8	26.8	
Total	6400	100.0	100.0	

Figura 6. Tabla de frecuencias de las categorías de ingresos

Creación de gráficos

Aunque algunos procedimientos estadísticos pueden crear gráficos, también puede utilizar el menú Gráficos para crear gráficos.

Por ejemplo, puede crear un gráfico que muestre la relación entre el servicio de teléfono inalámbrico y la propiedad de PDA (asistente digital personal).

1. En los menús seleccione:

Gráficos > Creador de gráficos...

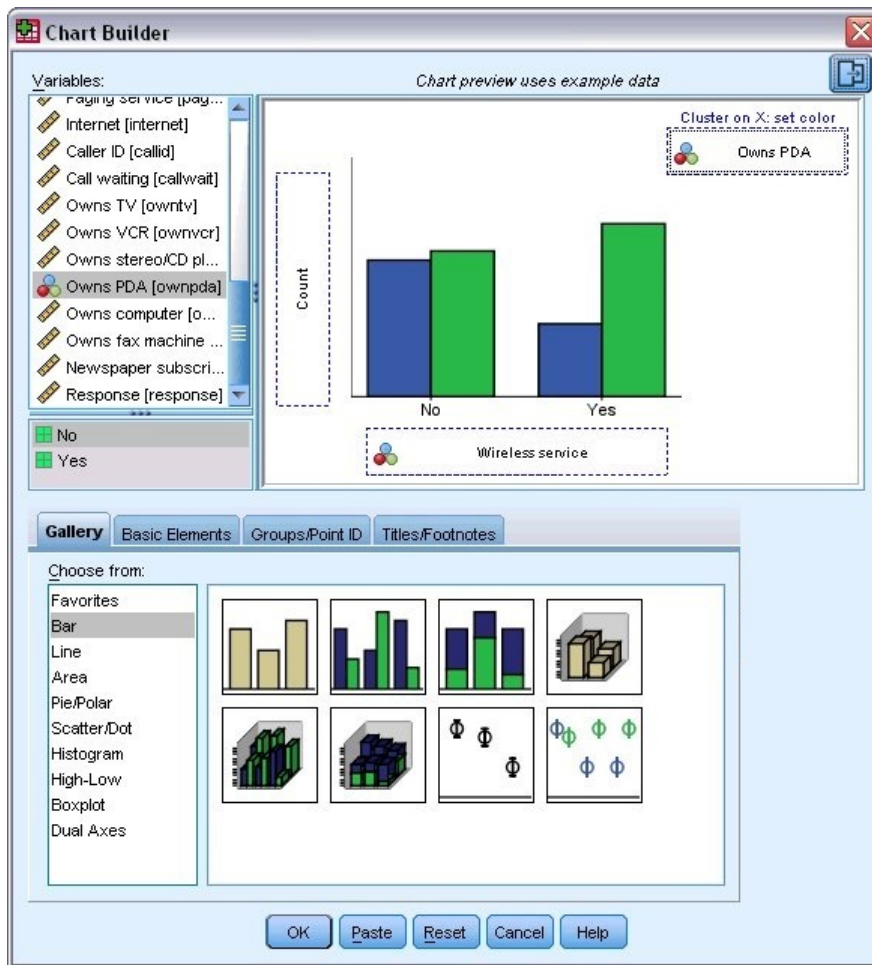


Figura 7. Cuadro de diálogo Generador de gráficos con las zonas de colocación rellenas

2. Pulse el separador **Galería** (si no está seleccionado).
3. Pulse **Barra** (si no está seleccionado).
4. Arrastre el icono Barra agrupada al lienzo, que es el área grande situada encima de la Galería.
5. Desplácese hacia abajo en la lista Variables, pulse con el botón derecho del ratón *Servicio inalámbrico [wireless]* y, a continuación, seleccione **Nominal** como nivel de medición.
6. Arrastre la variable *Servicio inalámbrico [wireless]* al eje x.
7. Pulse con el botón derecho del ratón *Posee PDA [ownpda]* y seleccione **Nominal** como nivel de medición.
8. Arrastre la variable *Posee PDA [ownpda]* a la zona de colocación del clúster en la esquina superior derecha del lienzo.
9. Pulse en **Aceptar** para crear el gráfico.

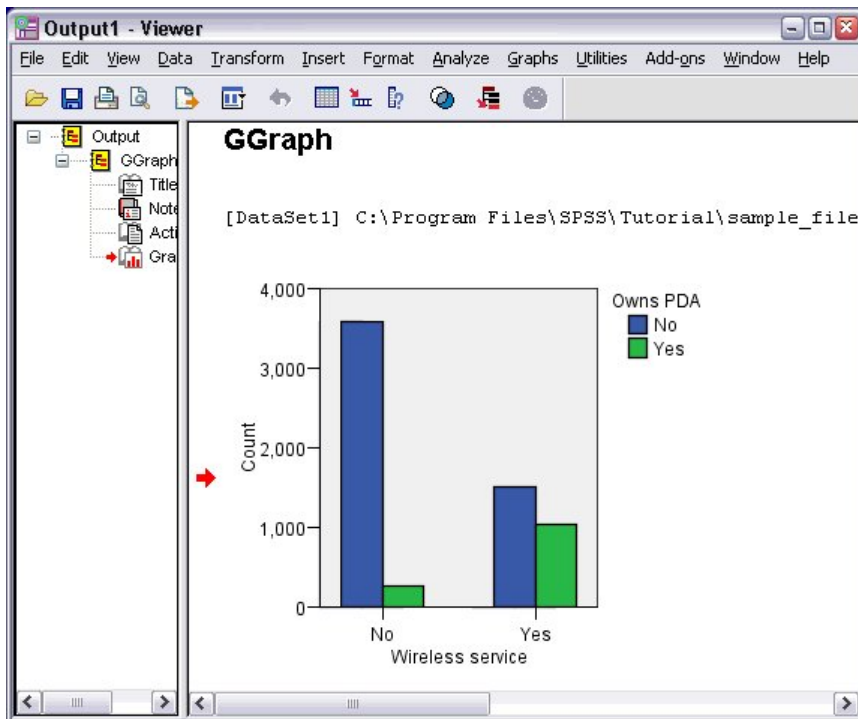


Figura 8. Gráfico de barras en la ventana Visor

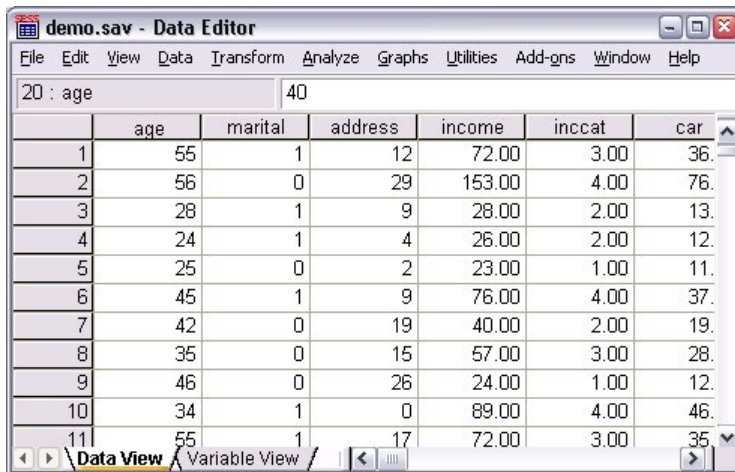
El gráfico de barras aparecerá en el Visor. El gráfico muestra que las personas con servicio de telefonía inalámbrica son mucho más propensas a tener PDA que las personas sin servicio inalámbrico.

Puede editar los gráficos y las tablas pulsando dos veces en ellos en el panel de contenidos de la ventana Visor y, a continuación, puede copiar y pegar los resultados en otras aplicaciones. Esos temas se tratarán más adelante.

Capítulo 2. Leyendo datos

Los datos se pueden introducir directamente, o importarse desde diferentes orígenes. Los procesos para leer datos almacenados en archivos de datos de IBM SPSS Estadística ; aplicaciones de hoja de cálculo, como Microsoft Excel; aplicaciones de base de datos, como Microsoft Access; y archivos de texto se describen en este capítulo.

Estructura básica de archivos de datos de IBM SPSS Estadística



	age	marital	address	income	inccat	car
1	55	1	12	72.00	3.00	36.
2	56	0	29	153.00	4.00	76.
3	28	1	9	28.00	2.00	13.
4	24	1	4	26.00	2.00	12.
5	25	0	2	23.00	1.00	11.
6	45	1	9	76.00	4.00	37.
7	42	0	19	40.00	2.00	19.
8	35	0	15	57.00	3.00	28.
9	46	0	26	24.00	1.00	12.
10	34	1	0	89.00	4.00	46.
11	55	1	17	72.00	3.00	35.

Figura 9. Editor de datos

Los archivos de datos IBM SPSS Estadística están organizados por casos (filas) y variables (columnas). En este archivo de datos, los casos representan a los encuestados individuales que responden a una encuesta. Las variables representan respuestas a cada pregunta que se hace en la encuesta.

Lectura de archivos de datos de IBM SPSS Estadística

Los archivos de datos de IBM SPSS Estadística , que tienen una extensión de archivo *.sav*, contienen los datos guardados.

1. En los menús seleccione:

Archivo > Abrir > Datos...

2. Acceda a *demo.sav* y ábralo. Consulte el tema [Capítulo 10, “Archivos de muestra”](#), en la página 81 para obtener más información.

Los datos se muestran ahora en el Editor de datos.

The screenshot shows the SPSS Data Editor window for a file named 'demo.sav'. The window title bar includes standard OS controls (minimize, maximize, close) and the text 'demo.sav - Data Editor'. Below the title bar is a menu bar with options: File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Add-ons, Window, and Help. The main area displays a data table with 11 rows and 6 columns. The columns are labeled 'age', 'marital', 'address', 'income', 'inccat', and 'car'. The first row shows values: 1, 55, 1, 12, 72.00, 3.00, 36. The second row: 2, 56, 0, 29, 153.00, 4.00, 76. The third row: 3, 28, 1, 9, 28.00, 2.00, 13. The fourth row: 4, 24, 1, 4, 26.00, 2.00, 12. The fifth row: 5, 25, 0, 2, 23.00, 1.00, 11. The sixth row: 6, 45, 1, 9, 76.00, 4.00, 37. The seventh row: 7, 42, 0, 19, 40.00, 2.00, 19. The eighth row: 8, 35, 0, 15, 57.00, 3.00, 28. The ninth row: 9, 46, 0, 26, 24.00, 1.00, 12. The tenth row: 10, 34, 1, 0, 89.00, 4.00, 46. The eleventh row: 11, 55, 1, 17, 72.00, 3.00, 35. At the bottom of the window, there are navigation buttons and a label 'Data View' which is currently selected.

	age	marital	address	income	inccat	car
1	55	1	12	72.00	3.00	36.
2	56	0	29	153.00	4.00	76.
3	28	1	9	28.00	2.00	13.
4	24	1	4	26.00	2.00	12.
5	25	0	2	23.00	1.00	11.
6	45	1	9	76.00	4.00	37.
7	42	0	19	40.00	2.00	19.
8	35	0	15	57.00	3.00	28.
9	46	0	26	24.00	1.00	12.
10	34	1	0	89.00	4.00	46.
11	55	1	17	72.00	3.00	35.

Figura 10. Archivo de datos abierto

Lectura de datos de Excel

En lugar de escribir todos los datos directamente en el Editor de datos, puede leer datos desde aplicaciones como Microsoft Excel. También se pueden leer cabeceras de columna como nombres de variable.

1. En los menús seleccione:

Archivo > Importar datos > Excel

2. Vaya a la carpeta Samples\English y seleccione demo.xlsx.

El cuadro de diálogo **Leer archivo Excel** ofrece una vista previa del archivo de datos. Se muestra el contenido de la primera hoja del archivo. Si el archivo tiene varias hojas, puede seleccionar la hoja de la lista.

Puede ver que algunos de los valores de cadena para *Sexo* tienen espacios iniciales. Algunos de los valores de *MaritalStatus* se visualizan como puntos (.).

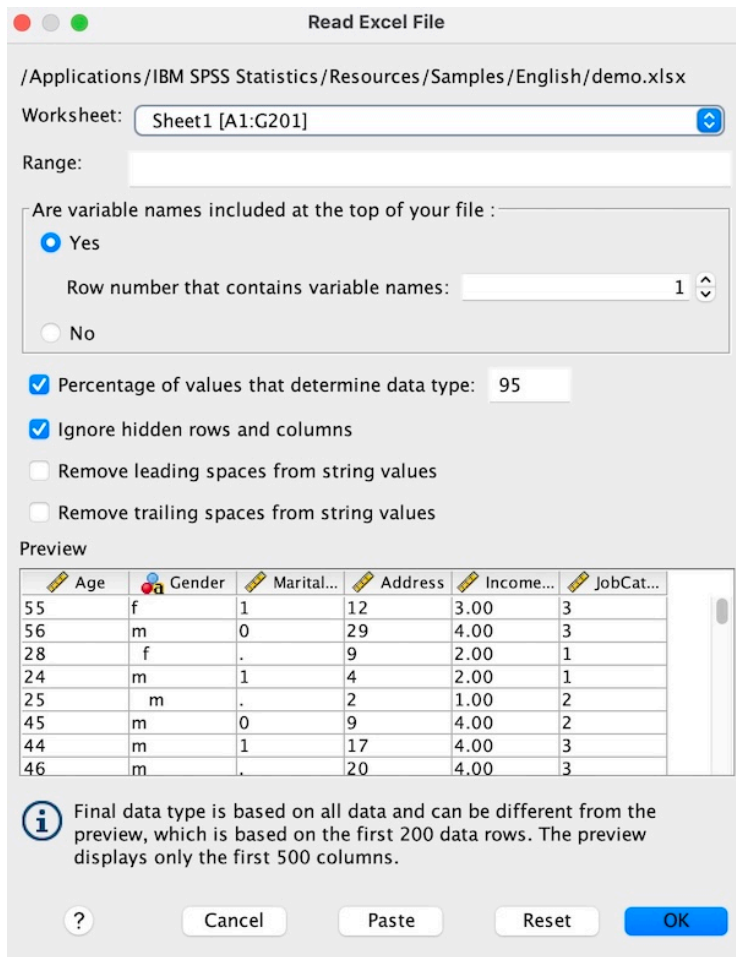
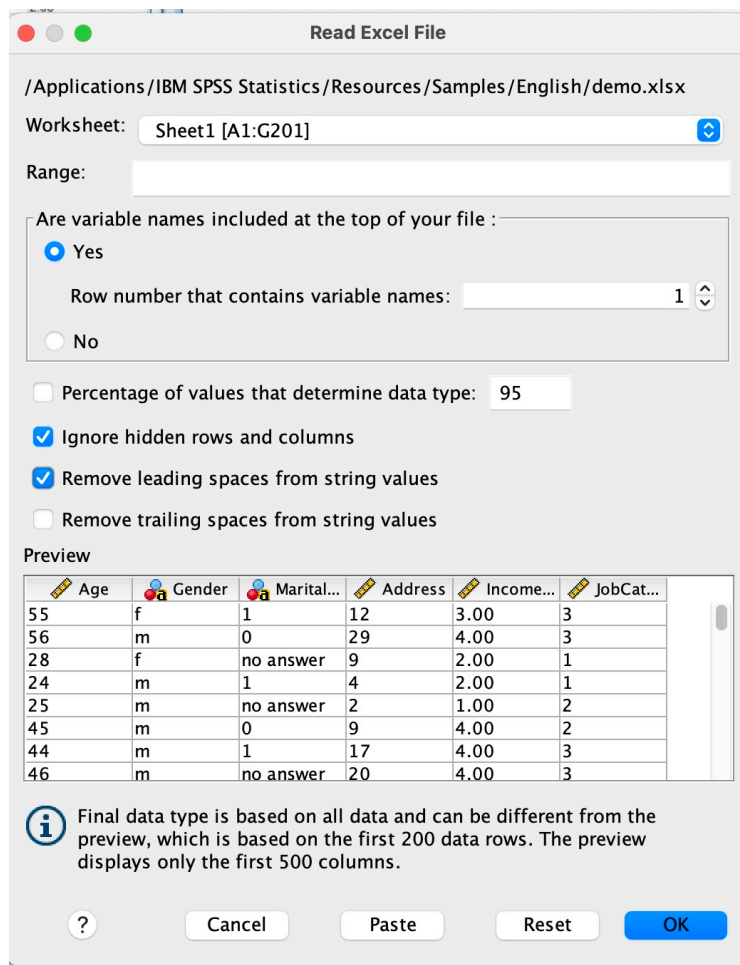


Figura 11. Cuadro de diálogo Leer archivo Excel

- Asegúrese de seleccionar **Sí** para **Son nombres de variables incluidos en la parte superior de su archivo**. El valor por defecto para el **número de Fila que contiene nombres de variables** es 1. Si los encabezados de las columnas no se ajustan a las reglas de nombre de variable, se convierten a nombres de variable válidos. Los encabezados de columna originales se guardan como etiquetas de variable.
- Seleccione **Eliminar espacios iniciales de valores de cadena**.
- Deseleccione **Porcentaje de valores que determinan tipo de datos**.



El valor de cadena "sin respuesta" se muestra ahora en las celdas que no estaban en el sistema. Si no hay parámetro de porcentaje de valores y la columna contiene una combinación de tipos de datos, la variable se lee como un tipo de datos de cadena. Los valores se mantienen, pero los valores numéricos se tratan como valores de cadena.

6. Seleccione (marque) **Porcentaje de valores que determinan tipo de datos** para tratar *MaritalStatus* como una variable numérica.
7. Pulse **Aceptar** para leer el archivo Excel.

Los datos aparecerán en el Editor de datos, con los encabezados de las columnas utilizados como nombres de variables. Al no poder contener espacios los nombres de variables, se eliminan los espacios de los encabezados de las columnas originales. Por ejemplo, el encabezado de columna "Estado civil" se convierte en la variable *MaritalStatus*. La cabecera de columna original se retiene como una etiqueta de variable.

Visible: 6 of 6 Variables

	Age	Gender	Marital Status	Address	Income Category	Job Category	var	var
1	55	f	1	12	3.00	3		
2	56	m	0	29	4.00	3		
3	28	f	no answer	9	2.00	1		
4	24	m	1	4	2.00	1		
5	25	m	no answer	2	1.00	2		
6	45	m	0	9	4.00	2		
7	44	m	1	17	4.00	3		
8	46	m	no answer	20	4.00	3		
9	41	m	no answer	10	2.00	2		
10	29	f	no answer	4	1.00	2		
11	34	m	0	0	4.00	2		
12	55	f	0	17	3.00	1		
13	28	m	0	9	3.00	1		
14	21	f	1	2	1.00	1		

Overview Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:ON Classic

Figura 12. Datos de Excel importados

Información relacionada

“Archivos de muestra” en la página 81

Lectura de datos de una base de datos

Los datos procedentes de bases de datos se importan con gran facilidad mediante el Asistente para bases de datos. Se puede leer directamente cualquier base de datos que utilice controladores de ODBC (Open Database Connectivity), una vez instalados los controladores. En el CD de instalación, se suministran controladores ODBC para diferentes formatos de bases de datos. Se pueden obtener controladores adicionales de otros fabricantes. En este ejemplo se describe una de las aplicaciones de base de datos más comunes, Microsoft Access.

Nota: este ejemplo es específico para Microsoft Windows y requiere un controlador ODBC para Access. El controlador ODBC de Microsoft Access solo funciona con la versión 32-bit de IBM SPSS Estadística. Los pasos son similares en otras plataformas, pero pueden requerir un controlador ODBC de terceros para Access.

1. En los menús seleccione:

Archivo > Importar datos > Base de datos > Nueva consulta...

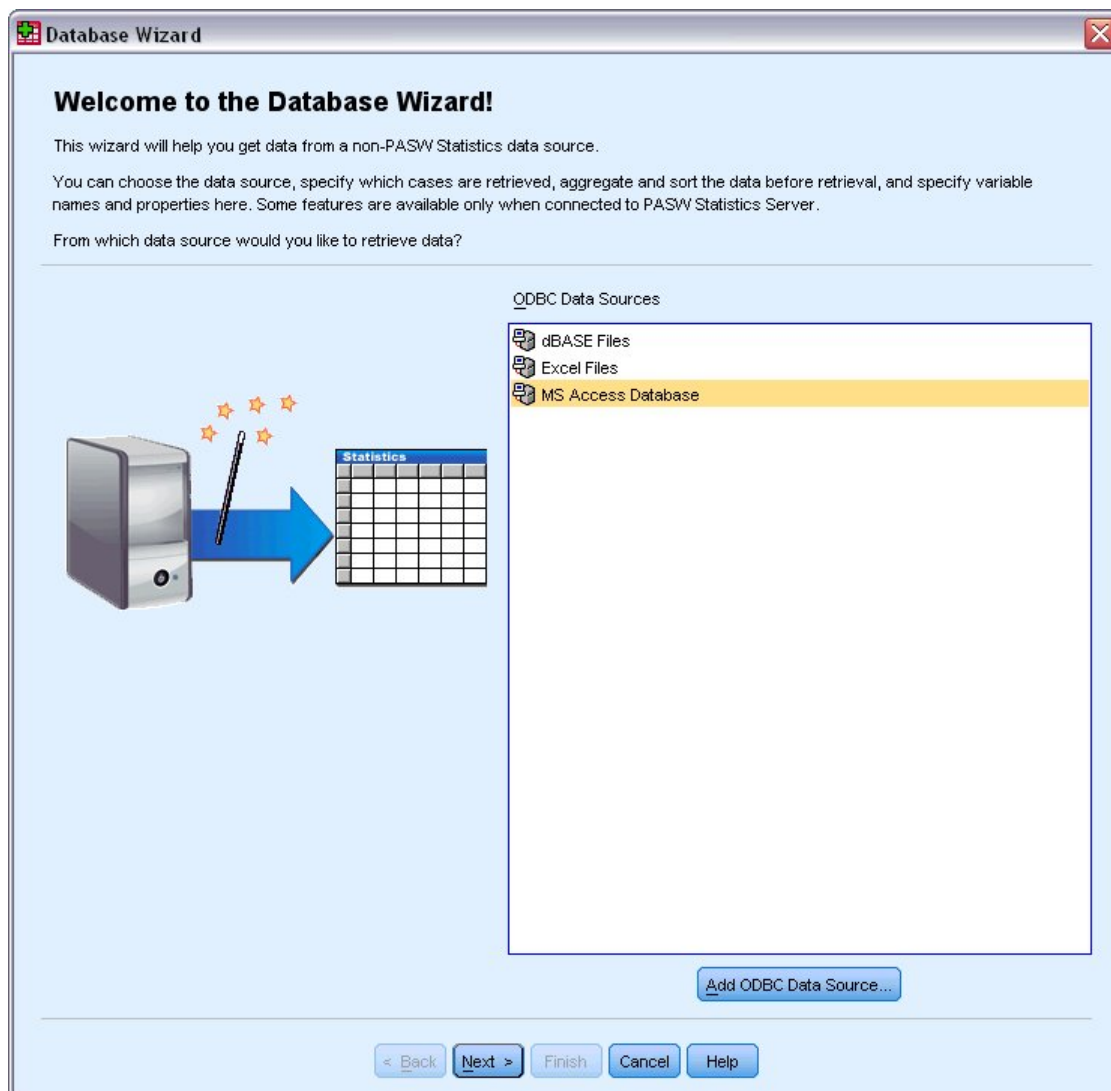


Figura 13. Cuadro de diálogo Bienvenido al asistente para bases de datos

2. Seleccione **MS Access Database** de la lista de orígenes de datos y pulse en **Siguiente**.

Nota: dependiendo de la instalación, es posible que también aparezca una lista de los orígenes de datos OLEDB en la parte izquierda del asistente (sólo sistemas operativos Windows), pero este ejemplo utiliza la lista de orígenes de datos ODBC que aparece a la derecha.

3. Pulse **Examinar** para ir al archivo de base de datos de Access que desea abrir.
4. Abrir *demo.mdb*. Consulte el tema [Capítulo 10, “Archivos de muestra”](#), en la [página 81](#) para obtener más información.
5. Pulse en **Aceptar** en el cuadro de diálogo de acceso.

En el paso siguiente, puede especificar las tablas y las variables que desea importar.

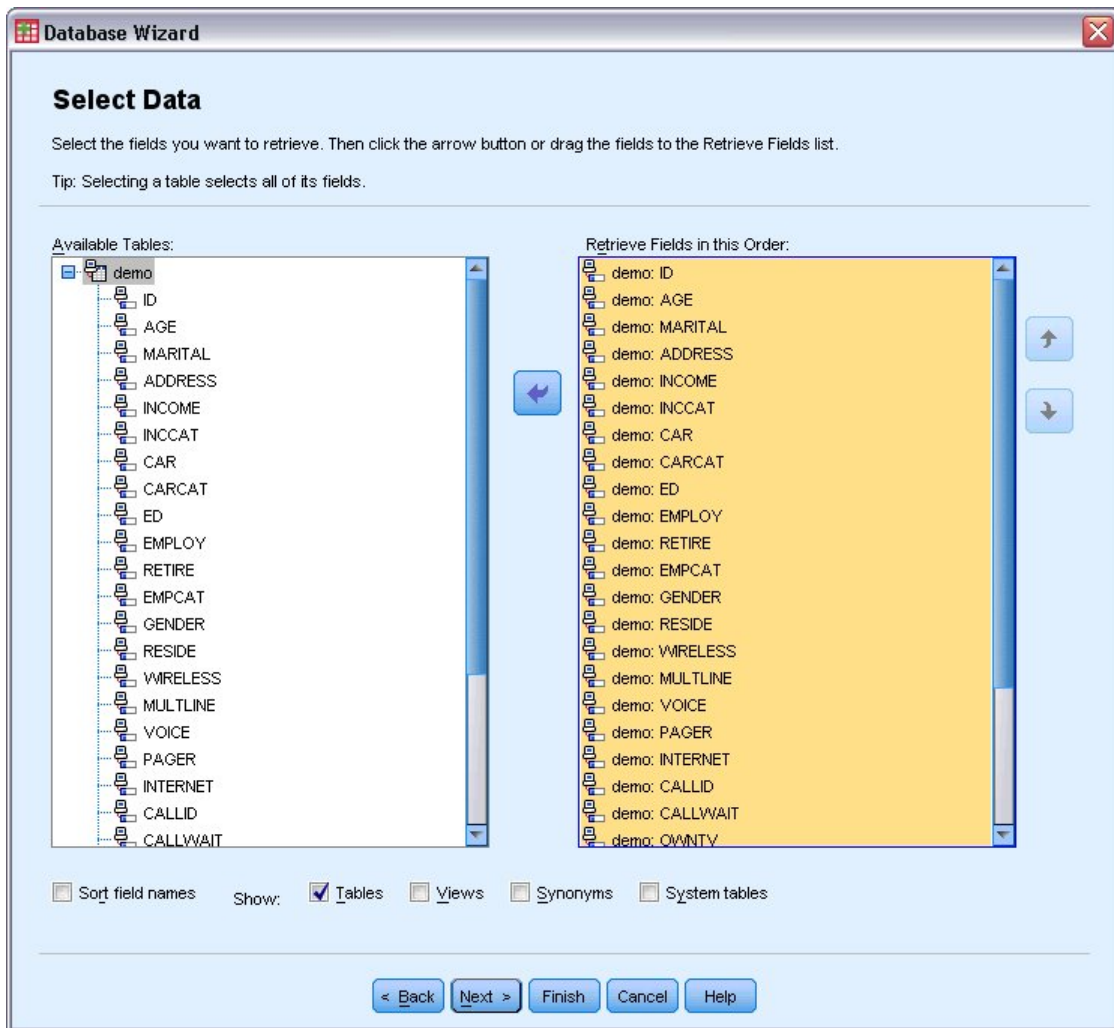


Figura 14. Paso Seleccionar datos

6. Arrastre toda la tabla **demo** a la lista Recuperar campos en este orden.
7. Pulse **Siguiente**.

En el siguiente paso, se pueden seleccionar los registros (casos) que desea importar.

Si no desea importar todos los casos, puede importar un subconjunto de casos (por ejemplo, los varones mayores de 30) o una muestra aleatoria de casos del origen de datos. Para orígenes de datos grandes, es posible que desee limitar el número de casos a una muestra pequeña y representativa para reducir el tiempo de proceso.

8. Pulse en **Siguiente** para continuar.

Los nombres de campo se utilizan para crear nombres de variable. Si es preciso, los nombres se convierten en nombres de variable válidos. Los nombres de campos originales se reservan como etiquetas de variable. También puede cambiar los nombres de variable antes de importar la base de datos.

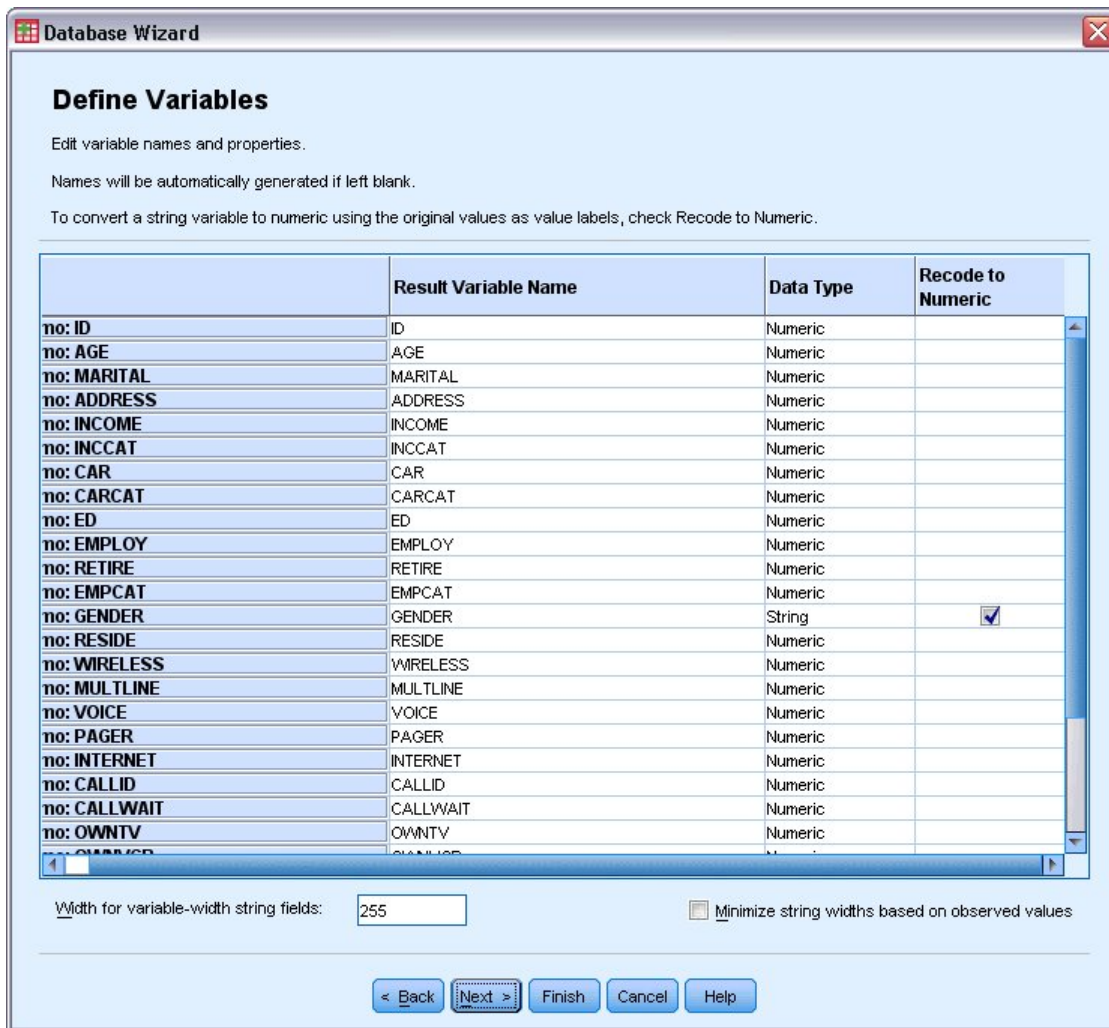


Figura 15. Paso Definir las variables

9. Pulse en la casilla **Recodificar como numérica** en el campo Género. Esta opción convierte las variables de cadena en variables de entero y conserva el valor original como la etiqueta de valor para la nueva variable.
10. Pulse en **Siguiente** para continuar.
 En el cuadro de diálogo Resultados aparecerá la sentencia SQL creada a partir de las selecciones realizadas en el Asistente para bases de datos. Esta sentencia se puede ejecutar ahora o guardar en un archivo para su uso posterior.
11. Pulse **Finalizar** para importar los datos.

En el editor de datos, ya están disponibles todos los datos de la base de datos de Access seleccionados para su importación.

Lectura de datos desde un archivo de texto

Los archivos de texto representan otro origen habitual de datos. Muchos programas de hojas de cálculo y de bases de datos pueden guardar su contenido en uno de los varios formatos de archivo de texto. Los archivos delimitados por comas o por tabuladores hacen referencia a filas de datos que utilizan comas o tabuladores para indicar cada variable. En este ejemplo, los datos están delimitados por tabulaciones.

1. En los menús seleccione:
Archivo > Importar datos > Datos de texto...
2. Vaya a la carpeta Samples\English y seleccione demo.txt.

El Asistente para la importación de texto le guiará a través del proceso para definir cómo se interpreta el archivo de texto especificado.

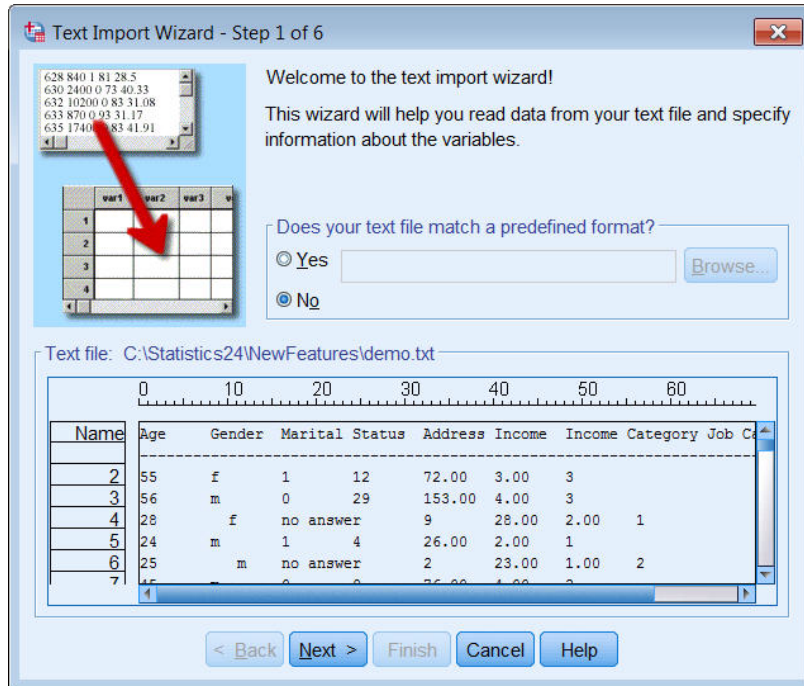


Figura 16. Asistente para la importación de texto: paso 1 de 6

3. En el Paso 1, se puede seleccionar un formato predefinido o crear uno nuevo en el asistente. Seleccione **No**.

4. Pulse en **Siguiente** para continuar.

Como ya se ha comentado, este archivo utiliza un formato delimitado por tabulaciones. Además, los nombres de variable se definen en la línea superior de este archivo.

5. En el paso 2 del asistente, seleccione **Delimitadas** para indicar que el archivo tiene una estructura de formato delimitado.
6. Seleccione **Sí** para indicar que el archivo incluye nombres de variable en la parte superior del archivo.
7. Pulse en **Siguiente** para continuar.
8. En paso 3, introduzca 2 para el número de línea donde comienza el primer caso de datos (ya que los nombres de variables están en la primera línea).
9. Mantenga los valores predeterminados del resto de este paso y pulse en **Siguiente** para continuar.

La vista previa de datos del paso 4 proporciona una forma rápida de asegurarse de que el archivo se lea correctamente

10. Seleccione **Tabulador** y desactive las demás opciones para delimitadores. **Espacio** está seleccionado de forma predeterminada porque el archivo contiene espacios. Para este archivo, los espacios son parte de los valores de datos, no delimitadores. Debe dejar de seleccionar **Espacio** para leer el archivo correctamente.
11. Seleccione **Eliminar espacios iniciales de valores de cadena**. Los espacios al principio de los valores de una cadena afectan al modo en que se avalúan los valores de cadena en las expresiones. En este archivo, algunos de los valores para Sexo tienen espacios iniciales que no forman parte del valor. Si no elimina dichos espacios, el valor " f" se trata de forma distinta que el valor "f".

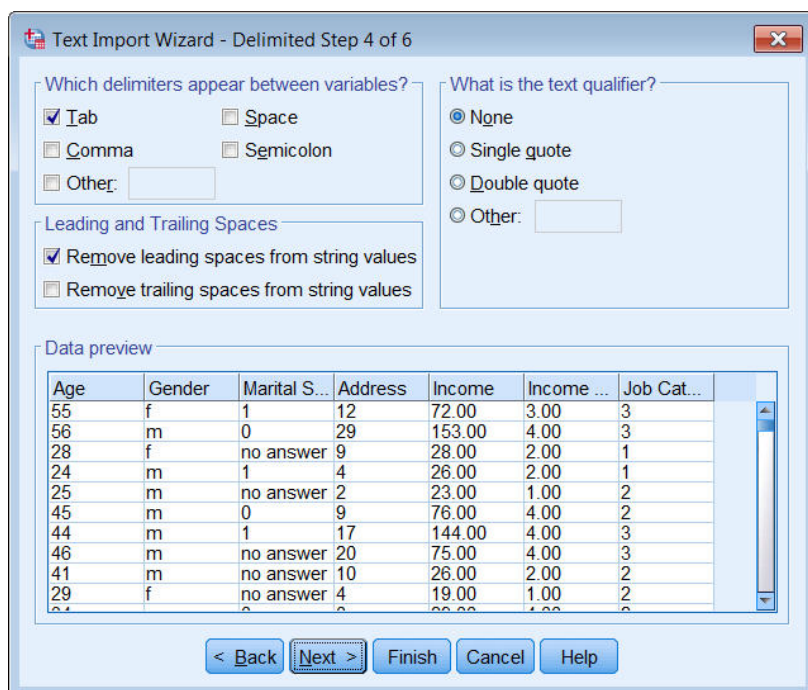


Figura 17. Asistente para la importación de texto: paso 4 de 6

12. Pulse en **Siguiente** para continuar.

Puesto que los nombres de las variables se modifican para adaptarse a las normas de denominación, el paso 5 ofrece la oportunidad de editar cualquier nombre no deseado.

También se pueden definir aquí los tipos de datos. Por ejemplo, puede cambiar *Ingresos* por el formato de moneda de dólar.

Para cambiar un tipo de datos:

13. En **Vista previa de datos**, seleccione *Ingresos*.

14. Seleccione **Dólar** en la lista desplegable Formato de datos.

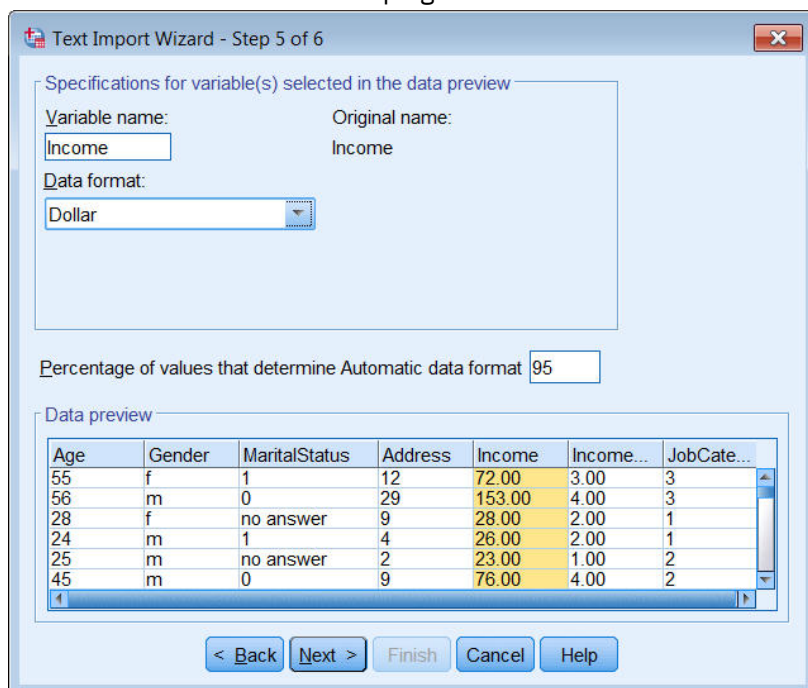


Figura 18. Cambio del tipo de datos

La variable *MaritalStatus* contiene valores de cadena y también numéricos. Menos del cinco por ciento de los valores son cadenas. Con el valor predeterminado de 95% para **Porcentaje de valores que determinan el formato de datos automático**, la variable se trata como numérica y los valores de cadena se establecen como perdidos del sistema. Si no hay datos que cumplan el valor de porcentaje, la variable se trata como una variable de cadena. Si cambia el valor a 100, se mantienen todos los valores pero todos los valores numéricos se tratan como cadenas.

15. Pulse en **Siguiente** para continuar.
16. Mantenga las selecciones predeterminadas del último paso y pulse en **Finalizar** para importar los datos.

Capítulo 3. Uso del Editor de datos

En el Editor de datos se muestra el contenido del archivo de datos activo. La información del Editor de datos consta de variables y casos.

- En la vista de datos, las columnas representan variables y las filas representan casos (observaciones).
- En la vista de variables, cada fila es una variable y cada columna es un atributo asociado a esa variable.

Las variables se utilizan para representar los diferentes tipos de datos que haya recopilado. Una analogía común es la de una encuesta. La respuesta a cada pregunta de una encuesta equivale a una variable. Las variables se proporcionan en muchos tipos diferentes, incluyendo números, cadenas, moneda y fechas.

Especificar datos numéricos

Se pueden especificar datos en el editor de datos, lo que puede ser útil para archivos de datos pequeños o para realizar pequeñas ediciones en archivos de datos más grandes.

1. Pulse en la pestaña **Vista de variables**, que se encuentra en la parte inferior de la ventana Editor de datos.

Debe definir las variables que va a utilizar. En este caso, sólo se necesitan tres variables: *edad*, *estado civil* e *ingresos*.

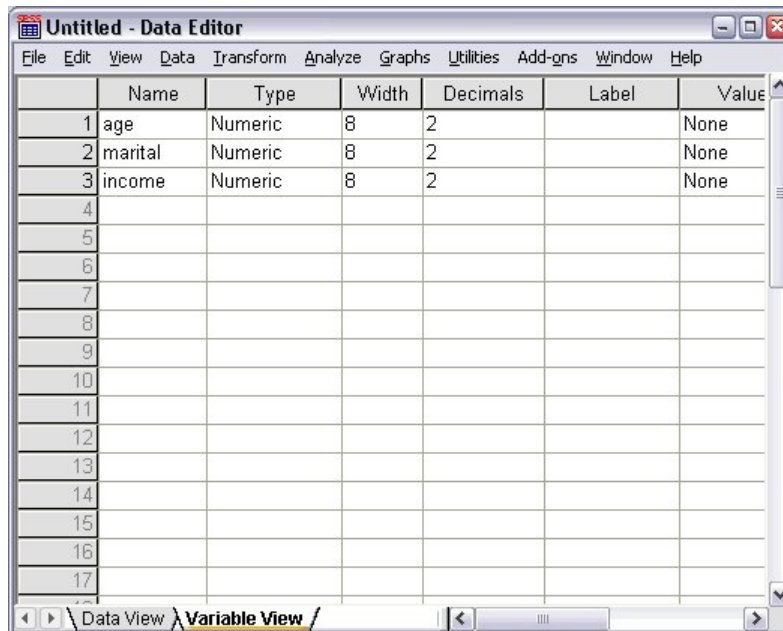


Figura 19. Nombres de variables en la Vista de variables

2. En la primera fila de la primera columna, escriba *age*.
3. En la segunda fila, escriba *marital*.
4. En la tercera fila, escriba *income*.

A las nuevas variables se proporciona automáticamente un tipo de datos numérico.

Si no introduce nombres de variable, se crearán nombres exclusivos automáticamente. Sin embargo, estos nombres no son descriptivos y no se recomiendan para archivos de datos grandes.

5. Pulse en la pestaña **Vista de datos** para seguir introduciendo los datos.

Los nombres que ha especificado en la vista de variables ahora son las cabeceras de las tres primeras columnas de la vista de datos.

Empiece a especificar datos en la primera fila, empezando en la primera columna.

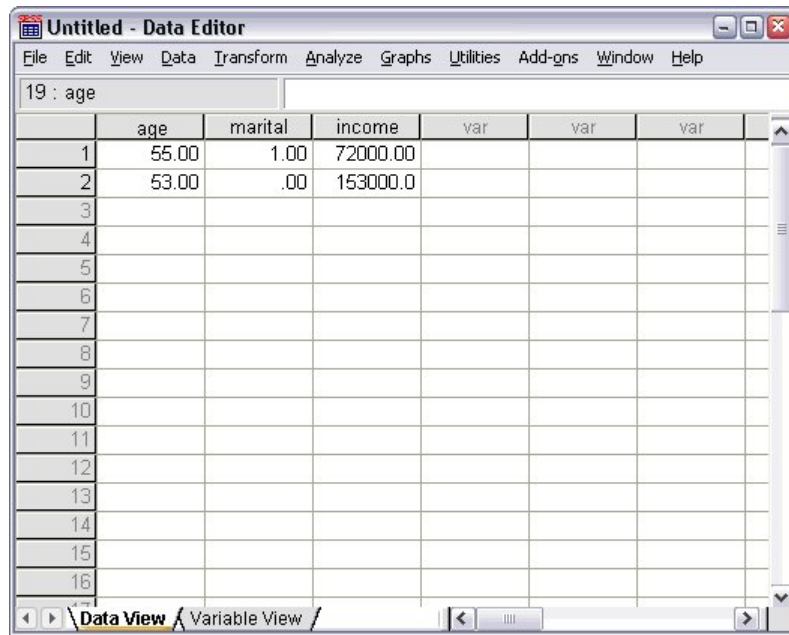


Figura 20. Valores introducidos en Vista de datos

6. En la columna *edad*, escriba 55.
7. En la columna *ecivil*, escriba 1.
8. En la columna *ingresos*, escriba 72000.
9. Mueva el cursor a la segunda fila de la primera columna para añadir los datos del siguiente sujeto.
10. En la columna *edad*, escriba 53.
11. En la columna *ecivil*, escriba 0.
12. En la columna *ingresos*, escriba 153000.

En la actualidad, en las columnas de la *edad* y el *estado civil* aparecen separaciones decimales, aunque sus valores sean números enteros. Para ocultar los separadores decimales en estas variables:

13. Pulse en la pestaña **Vista de variables**, que se encuentra en la parte inferior de la ventana Editor de datos.
14. En la columna *Decimales* de la fila *edad*, escriba 0 para ocultar el decimal.
15. En la columna *Decimales* de la fila *ecivil*, escriba 0 para ocultar el decimal.

Especificar datos de cadena

Los datos no numéricos, como las cadenas de texto, también se pueden especificar en el Editor de datos.

1. Pulse en la pestaña **Vista de variables**, que se encuentra en la parte inferior de la ventana Editor de datos.
2. En la primera celda de la primera fila vacía, escriba *sex* para el nombre de la variable.
3. Pulse la celda *Tipo* junto a la entrada.
4. Pulse el botón del lado derecho de la celda *Tipo* para abrir el recuadro de diálogo Tipo de variable.
5. Seleccione **Cadena** para especificar el tipo de variable.
6. Pulse **Aceptar** para guardar la selección y volver al Editor de datos.

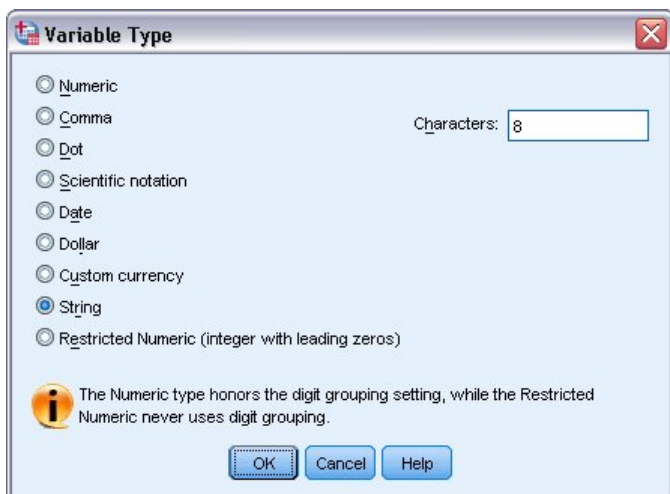


Figura 21. Cuadro de diálogo Tipo de variable

Definición de datos

Además de definir los tipos de datos, también puede definir etiquetas descriptivas de variable y etiquetas de valor para los nombres de variables y los valores de datos. Estas etiquetas descriptivas se utilizan en los informes y gráficos estadísticos.

Añadir etiquetas de variable

Las etiquetas están destinadas a proporcionar descripciones de las variables. Estas descripciones suelen ser versiones extendidas de los nombres de variable. Las etiquetas pueden ocupar hasta 255 bytes. Estas etiquetas se utilizan en la salida para identificar las diferentes variables.

1. Pulse en la pestaña **Vista de variables**, que se encuentra en la parte inferior de la ventana Editor de datos.
2. En la columna *Etiqueta* de la fila de *edad*, escriba Respondent ' s Age.
3. En la columna *Etiqueta* de la fila *marital*, escriba Marital Status.
4. En la columna *Etiqueta* de la fila de *ingresos*, escriba Household Income.
5. En la columna *Etiqueta* de la fila de *sexo*, escriba Gender.

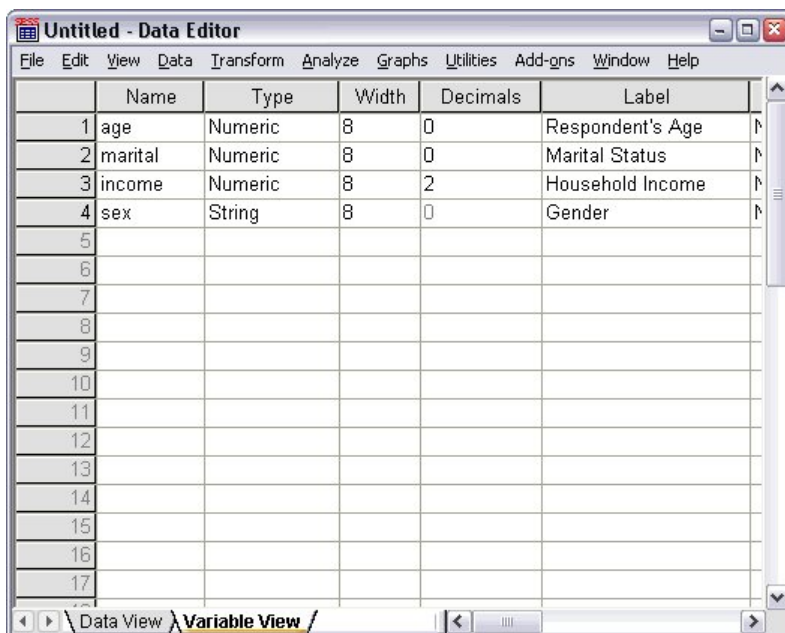


Figura 22. Etiquetas de variable introducidas en Vista de variables

Cambiar el tipo y el formato de la variable

En la columna *Tipo* aparecen los tipos de datos actuales de cada variable. Los tipos de datos más comunes son los numéricos y las cadenas, pero se admiten otros muchos formatos. En el archivo de datos actual, la variable *Ingresos* se define como un tipo numérico.

1. Pulse la celda *Tipo* de la fila *Ingresos* y, a continuación, pulse el botón situado en la parte derecha de la celda para abrir el recuadro de diálogo Tipo de variable.
2. Seleccione **Dólar**.

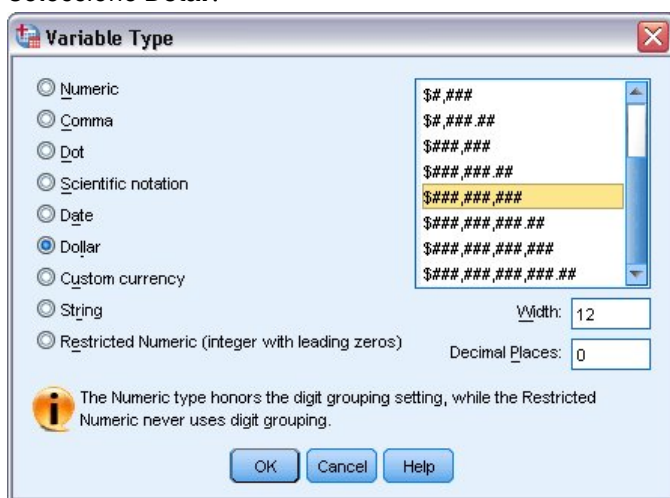


Figura 23. Cuadro de diálogo Tipo de variable

Se muestran las opciones de formato para el tipo de datos seleccionado actualmente.

3. Para el formato de la moneda en este ejemplo, seleccione **\$###,###,###**.
4. Pulse en **Aceptar** para guardar los cambios.

Añadir etiquetas de valores

Las etiquetas de valor proporcionan un método para correlacionar los valores de variable con una etiqueta de texto. En este ejemplo, hay dos valores aceptables para la variable *ecivil*. Un valor de 0 significa que el sujeto está soltero, y un valor de 1 significa que está casado.

1. Pulse en la casilla *Valores* de la fila *ecivil* y, a continuación, pulse el botón que hay a la derecha de la casilla para abrir el cuadro de diálogo Etiquetas de valor.

El **valor** es el valor numérico real.

La **etiqueta de valor** es la etiqueta de cadena que se aplica al valor numérico especificado.

2. Escriba 0 en el campo Valor.
3. Escriba Single en el campo Etiqueta.
4. Pulse **Añadir** para añadir esta etiqueta a la lista.

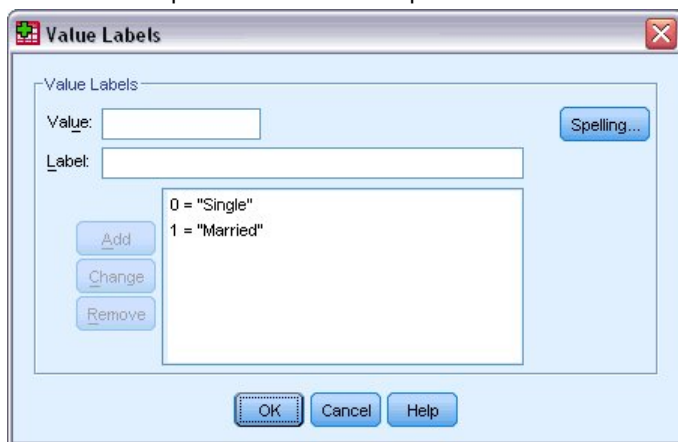


Figura 24. Cuadro de diálogo Etiquetas de valor

5. Escriba 1 en el campo Valor y Married en el campo Etiqueta.
6. Pulse en **Añadir** y a continuación, pulse en **Aceptar** para guardar los cambios y volver al Editor de datos.

Estas etiquetas también se pueden visualizar en la vista de datos, lo que puede hacer que sus datos sean más legibles.

7. Pulse el separador **Vista de datos** en la parte inferior de la ventana del Editor de datos.
8. En los menús seleccione:

Vista > Etiquetas de valor

Las etiquetas aparecerán ahora en una lista cuando introduzca valores en el Editor de datos. Esta configuración tiene el beneficio de sugerir una respuesta válida y de proporcionar una respuesta más descriptiva.

Si el elemento de menú Etiquetas de valor ya está activo (con una marca de selección junto a él), al seleccionar **Etiquetas de valor** se *desactivará* la visualización de las etiquetas de valor.

Manejo de datos perdidos

Los datos perdidos o no válidos son generalmente tan comunes que no se deben ignorar. Puede que las personas que han respondido a una encuesta se nieguen a contestar ciertas preguntas, no sepan la respuesta o contesten de forma inesperada. Si no filtra ni identifica estos datos, es posible que el análisis no proporcione resultados precisos.

Para datos numéricos, los campos de datos vacíos o los campos que contienen entradas no válidas se convierten en valores perdidos del sistema, identificables por un único periodo.

El motivo por el que se pierde un valor puede ser importante para el análisis. Por ejemplo, puede resultar útil distinguir entre los encuestados que se negaron a responder a una pregunta y los encuestados que no respondieron a una pregunta porque no era aplicable.

Valores perdidos de una variable numérica

1. Pulse en la pestaña **Vista de variables**, que se encuentra en la parte inferior de la ventana Editor de datos.
2. Pulse en la casilla *Perdido* de la fila *edad* y, a continuación, pulse el botón que hay a la derecha de la casilla para abrir el cuadro de diálogo Valores perdidos.

En este recuadro de diálogo, puede especificar hasta tres valores perdidos distintos, o puede especificar un rango de valores más un valor discreto adicional.



Figura 25. Cuadro de diálogo Valores perdidos

3. Seleccione **Valores perdidos discretos**.
4. Escriba 999 en el primer cuadro de texto u deje los otros dos cuadros de texto en blanco.
5. Pulse en **Aceptar** para guardar los cambios y volver a Editor de datos.
Ahora que se ha añadido el valor de datos perdidos, se puede aplicar una etiqueta a ese valor.
6. Pulse la celda *Valores* de la fila *age* y, a continuación, pulse el botón situado en la parte derecha de la celda para abrir el recuadro de diálogo Tipo de variable.
7. Escriba 999 en el campo Valor.
8. Escriba No Response en el campo Etiqueta.
9. Pulse **Añadir** para añadir esta etiqueta al archivo de datos.
10. Pulse en **Aceptar** para guardar los cambios y volver a Editor de datos.

Valores perdidos de una variable de cadena

Los valores perdidos en variables de cadena se tratan de forma parecida a los valores perdidos en variables numéricas. Sin embargo, a diferencia de las variables numéricas, los campos vacíos de las variables de cadena no se consideran perdidos del sistema. En su lugar, se interpretan como una cadena vacía.

1. Pulse en la pestaña **Vista de variables**, que se encuentra en la parte inferior de la ventana Editor de datos.
2. Pulse la celda *Perdidos* de la fila *sex* y, a continuación, pulse el botón situado en la parte derecha de la celda para abrir el recuadro de diálogo Valores perdidos.
3. Seleccione **Valores perdidos discretos**.
4. Escriba NR en el primer cuadro de texto.

Los valores perdidos de las variables de cadena distinguen entre mayúsculas y minúsculas. De este modo, un valor de *nr* no se considerará un valor perdido.

5. Pulse en **Aceptar** para guardar los cambios y volver a Editor de datos.

Ahora puede añadir una etiqueta para el valor perdido.

6. Pulse en la casilla *Valores* de la fila *sexo* y, a continuación, pulse el botón que hay a la derecha de la casilla para abrir el cuadro de diálogo Etiquetas de valor.
7. Introduzca NR en el campo Valor.
8. Escriba No Response en el campo Etiqueta.
9. Pulse **Añadir** para añadir esta etiqueta al proyecto.
10. Pulse en **Aceptar** para guardar los cambios y volver a Editor de datos.

Capítulo 4. Examinar estadísticos de resumen de variables individuales

Esta sección describe medidas de resumen sencillas y la influencia del nivel de medición de una variable en los tipos de estadísticos que se deben utilizar. Se utilizará el archivo de datos *demo.sav*. Consulte el tema [Capítulo 10, “Archivos de muestra”](#), en la [página 81](#) para obtener más información.

Nivel de medición

Las diferentes medidas de resumen son apropiadas para diferentes tipos de datos, dependiendo del nivel de medición:

Categorías. Datos con un número limitado de valores o categorías distintas (por ejemplo, género o estado civil). También se hace referencia a estos datos como **datos cualitativos**. Las variables categóricas pueden ser variables de cadena (alfanuméricas) o variables numéricas que utilizan códigos numéricos para representar a las categorías (por ejemplo, 0 = *Soltero/a* y 1 = *Casado/a*). Hay dos tipos básicos de datos categóricos:

- **Nominal.** Datos categóricos en los que las categorías no tienen un orden inherente. Por ejemplo, una categoría de trabajo de *ventas* no es superior o inferior a una categoría de trabajo de *marketing* o *investigación*.
- **Ordinal.** Datos categóricos en los que las categorías tienen un orden significativo, pero sin una distancia medible entre las categorías. Por ejemplo, hay un orden en los valores *alto*, *medio* y *bajo*, pero no se puede calcular la "distancia" entre los valores.

Escalas. Datos medidos en una escala de intervalo o de razón en los que los valores de los datos indican el orden de los valores y la distancia entre ellos. Por ejemplo, un salario de 72.195 \$ es superior a un salario de 52.398 \$ y la distancia entre los datos valores es 19.797 \$. También se denominan datos **cuantitativos** o **datos continuos**.

Medidas de resumen de datos categóricos

Para los datos categóricos, la medida de resumen más habitual es el número o el porcentaje de casos de cada categoría. El **modo** es la categoría que contiene el mayor número de casos. Para los datos ordinales, la **mediana** (el valor en el que la mitad de los casos se encuentran por encima y la otra mitad por debajo) también puede ser una medida de resumen útil si hay un gran número de categorías.

El procedimiento Frecuencias genera tablas de frecuencia que muestran tanto el número como el porcentaje de casos para cada valor observado de una variable.

1. En los menús seleccione:

Analizar > Estadísticas descriptivas > Descriptivos...

Nota: esta característica requiere la opción de Statistics Base.

2. Seleccione *Tiene PDA [pda]* y *Tiene TV [tv]* y muévelas hasta la lista Variable(s).

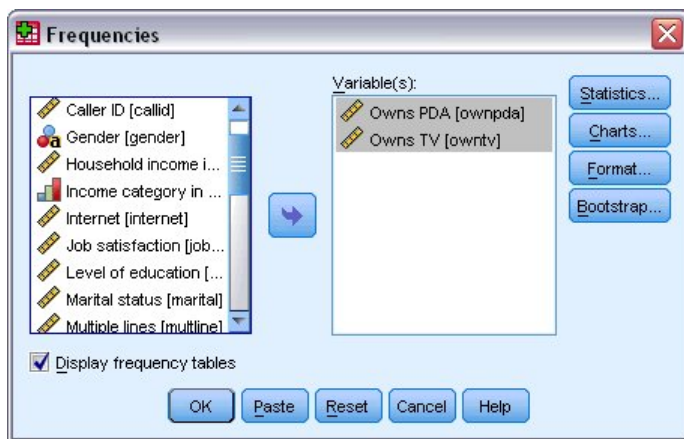


Figura 26. Variables categóricas seleccionadas para el análisis

3. Pulse **Aceptar** para ejecutar el procedimiento.

Frequency Table

Owns PDA

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	5093	79.6	79.6	79.6
	Yes	1307	20.4	20.4	100.0
Total		6400	100.0	100.0	

Owns TV

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	63	1.0	1.0	1.0
	Yes	6337	99.0	99.0	100.0
Total		6400	100.0	100.0	

Figura 27. Tablas de frecuencias

Las tablas de frecuencias aparecen en la ventana Visor. Las tablas de frecuencias revelan que sólo el 20,4% de las personas poseen una agenda digital, pero que casi todos poseen un televisor (99,0%). Estas podrían no ser revelaciones interesantes, aunque podría ser interesante saber más sobre el pequeño grupo de personas que no poseen televisores.

Gráficos de datos categóricos

Puede visualizar gráficamente la información en una tabla de frecuencia con un gráfico de barras o gráfico circular.

1. Vuelva a abrir el cuadro de diálogo Frecuencias. (Las dos variables deben estar seleccionadas aún).

Puede utilizar el botón Recuperación cuadros de diálogo de la barra de herramientas para regresar rápidamente a los procedimientos utilizados recientemente.



Figura 28. Botón Recuperar cuadros de diálogo

2. Pulse en **Gráficos**.
3. Seleccione **Gráficos de barras** y, a continuación, pulse en **Continuar**.
4. Pulse **Aceptar** en el recuadro de diálogo principal para ejecutar el procedimiento.

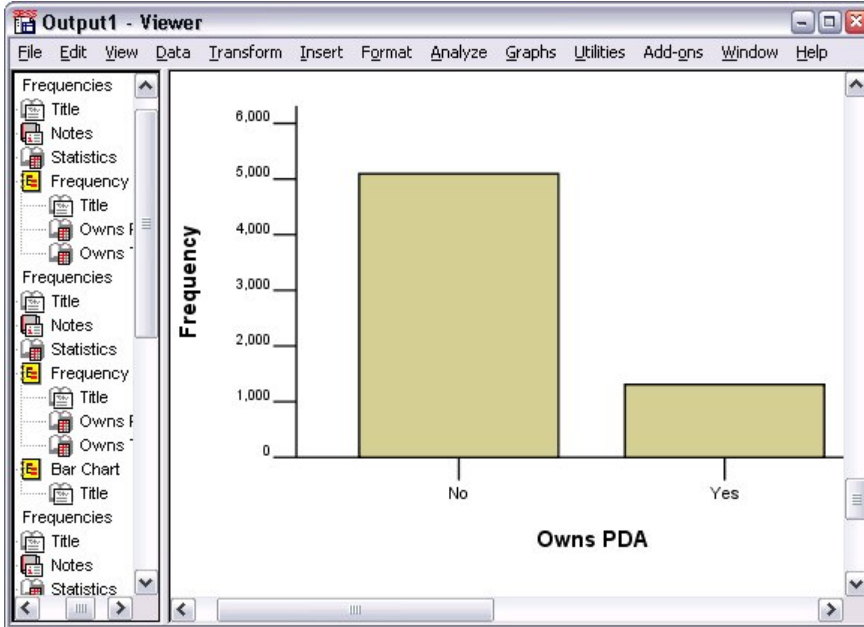


Figura 29. Gráfico de barras

Además de las tablas de frecuencia, la misma información se muestra ahora en forma de gráficos de barras, lo que facilita ver que la mayoría de las personas no poseen PDA, pero casi todo el mundo posee un televisor.

Medidas de resumen de variables de escala

Hay muchas medidas de resumen disponibles para las variables de escala, incluyendo:

- **Medidas de tendencia central.** Las medidas más comunes de tendencia central son la **media** (media aritmética) y la **mediana** (valor en el que la mitad de los casos se encuentran por encima y por debajo).
- **Medidas de dispersión.** Las estadísticas que miden la cantidad de variación o propagación en los datos incluyen la desviación estándar, el mínimo y el máximo.

1. Vuelva a abrir el cuadro de diálogo Frecuencias.
2. Pulse **Restablecer** para borrar los valores anteriores.
3. Seleccione *Ingresos del hogar en miles [ingres]* y muévela hasta la lista Variable(s).
4. Pulse en **Estadísticos**.
5. Seleccione **Media, Mediana, Std. desviación, Mínimo y Máximo**.
6. Pulse **Continuar**.
7. Anule la selección de **Mostrar tablas de frecuencias** en el cuadro de diálogo principal. (Por lo general, las tablas de frecuencias no son útiles para variables de escala ya que debe haber casi tantos valores distintos como casos en el archivo de datos).
8. Pulse **Aceptar** para ejecutar el procedimiento.

La tabla Estadísticas de frecuencias se visualiza en la ventana del Visor.

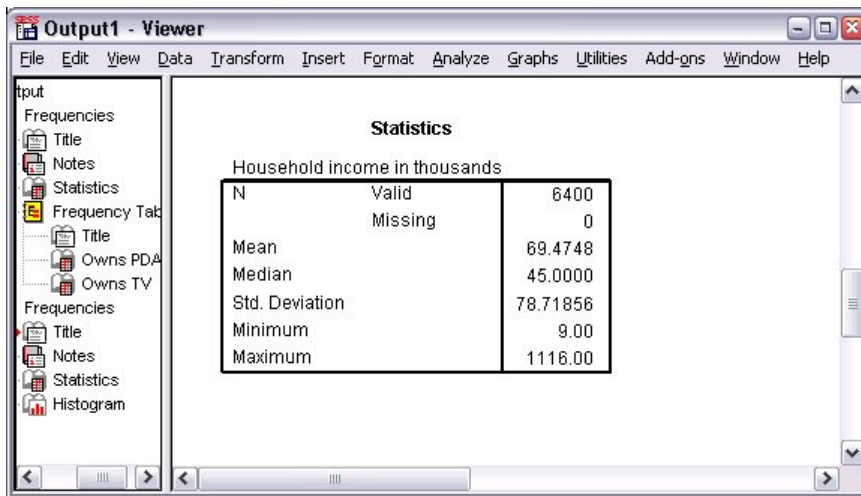


Figura 30. Tabla Frecuencias: Estadísticos

En este ejemplo, hay una importante diferencia entre la media y la mediana. La media es casi 25.000 mayor que la mediana, lo que indica que los valores no se distribuyen normalmente. Puede comprobar visualmente la distribución con un histograma.

Histogramas de variables de escala

1. Vuelva a abrir el cuadro de diálogo Frecuencias.
2. Pulse en **Gráficos**.
3. Seleccione **Histogramas** y **Con curva normal**.
4. Pulse **Continuar** y después pulse **Aceptar** en el recuadro de diálogo principal para ejecutar el procedimiento.

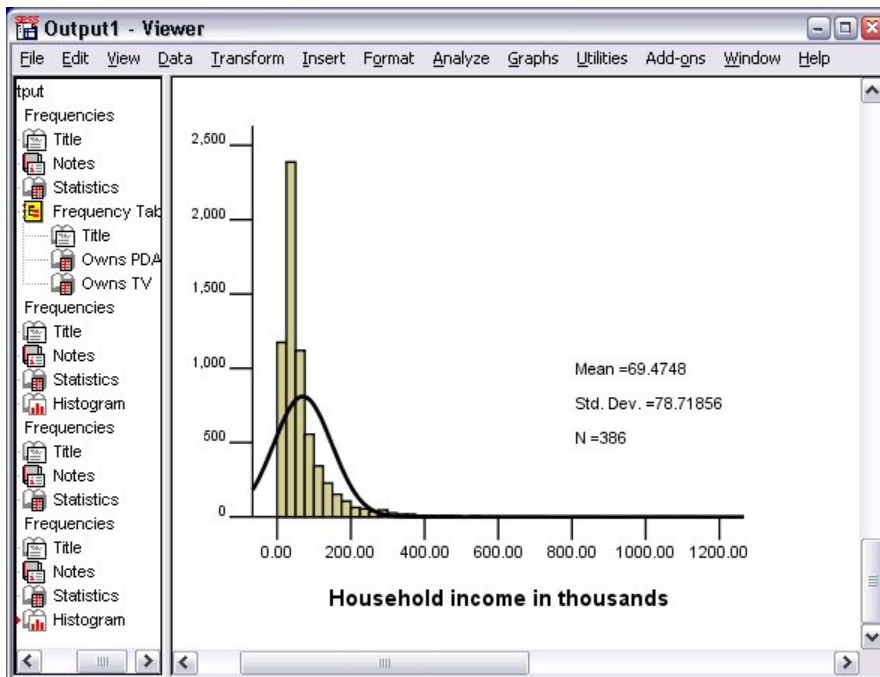


Figura 31. Histograma

La mayoría de los casos se agrupan en el extremo inferior de la escala, y la mayoría están por debajo de los 100.000. Sin embargo, hay unos pocos casos en el rango de 500.000 y superiores (demasiado pocos para que se visualicen sin modificar el histograma). Estos valores altos para solo unos pocos casos tienen

un efecto significativo en la media, pero poco o ningún efecto en la mediana, lo que hace que la mediana sea un mejor indicador de tendencia central en este ejemplo.

Capítulo 5. Creación y edición de gráficos

Puede crear y editar una gran variedad de tipos de gráficos. En este capítulo, crearemos y editaremos gráficos de barras. Puede aplicar los principios a cualquier tipo de gráfico.

Fundamentos de la creación de gráficos

Para mostrar el procedimiento básico de creación de gráficos, crearemos un gráfico de barras de los ingresos medios para los diferentes niveles de satisfacción laboral. En este ejemplo se utiliza el archivo de datos *demo.sav*. Consulte el tema Capítulo 10, “Archivos de muestra”, en la página 81 para obtener más información.

1. En los menús seleccione:

Gráficos > Creador de gráficos...

El recuadro de diálogo Creador de gráficos es una ventana interactiva que le permite obtener una vista previa del aspecto que tendrá un gráfico mientras lo crea.

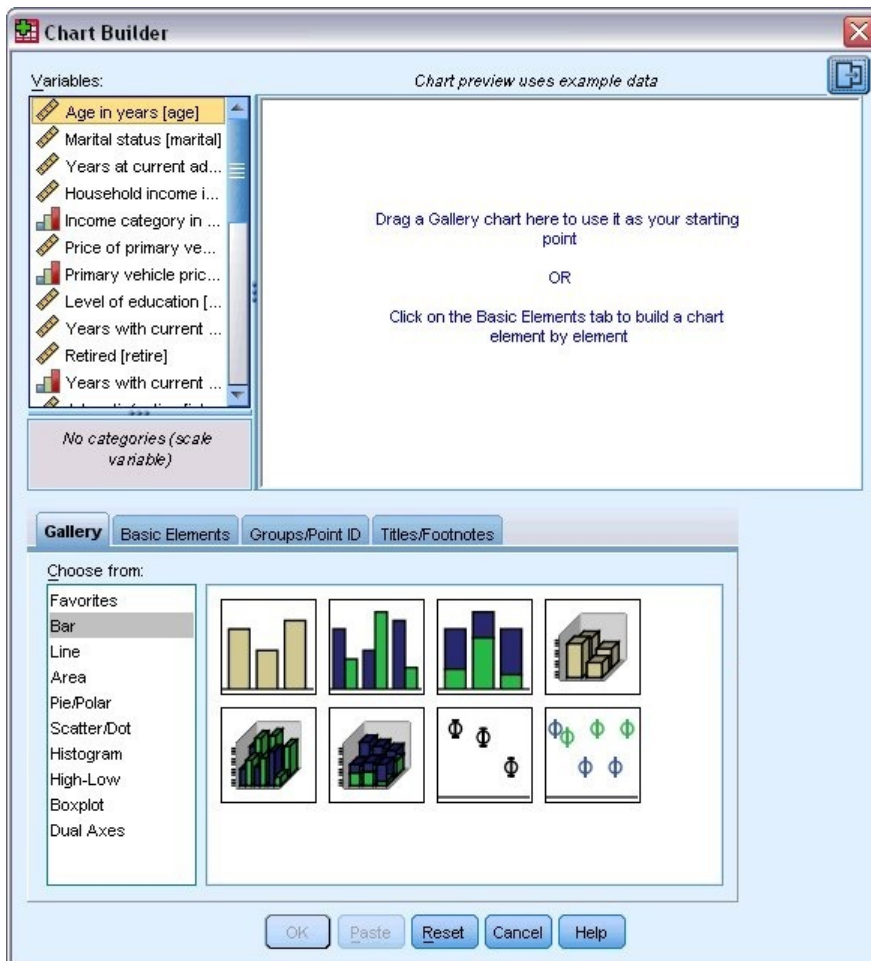


Figura 32. Cuadro de diálogo Generador de gráficos

Utilizar la galería del Creador de gráficos

1. Pulse en la pestaña **Galería** si no está seleccionada.

La galería incluye un gran número de gráficos predefinidos, organizados por tipo de gráfico. El separador Elementos básicos también proporciona elementos básicos (como ejes y elementos gráficos) para crear gráficos desde cero, pero es más fácil utilizar la Galería.

2. Pulse en la pestaña **Barra** si no está seleccionada.

Los iconos que representan los gráficos de barras que hay disponibles en la galería aparecen en el cuadro de diálogo. Las imágenes deberían proporcionar suficiente información para identificar el tipo de gráfico concreto. Si necesita más información, también puede visualizar una descripción de ayuda contextual del gráfico pausando el cursor sobre un icono.

3. Arrastre el icono correspondiente al gráfico de barras simple al "lienzo", que es la zona grande que hay encima de la galería. El generador de gráficos muestra una vista previa del gráfico en el lienzo. Observe que los datos utilizados para dibujar el gráfico no son los datos reales. Son datos de ejemplo.

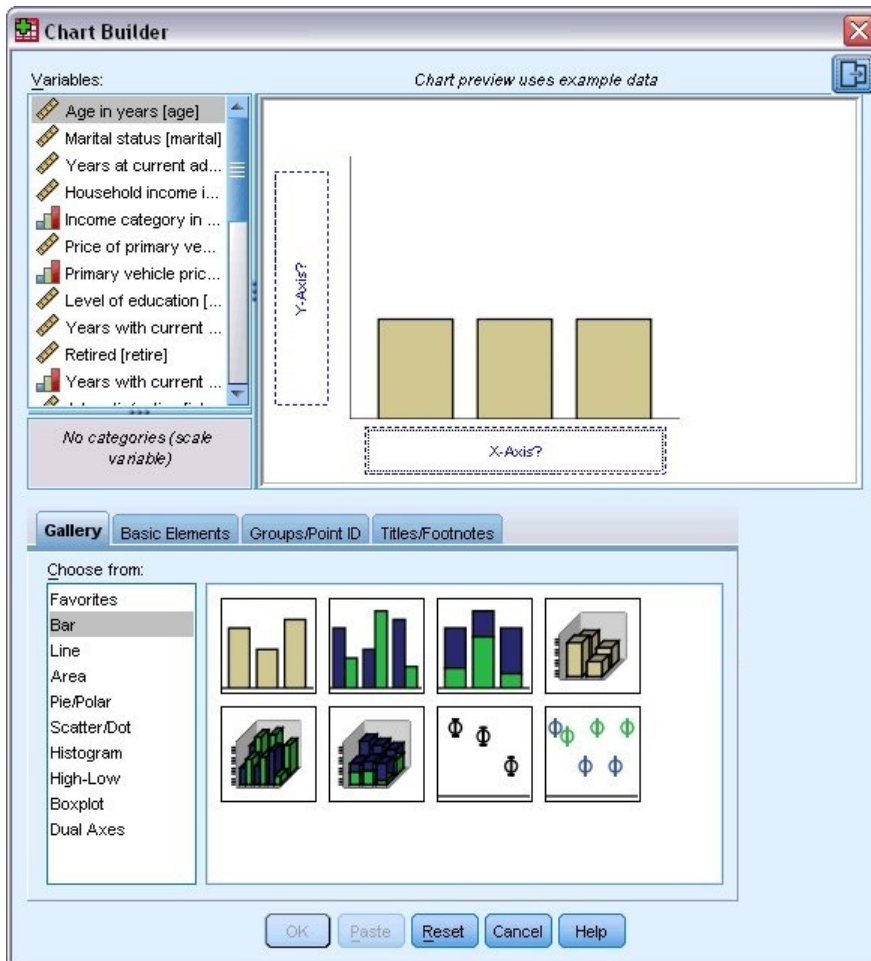


Figura 33. Gráfico de barras en el lienzo del generador de gráficos

Definición de variables y estadísticas

Aunque hay un gráfico en el lienzo, no está completo ya que no hay ninguna variable ni estadístico que controle la altura de las barras ni se ha especificado la variable que corresponde a cada barra. No se puede tener un gráfico sin variables ni estadísticos. Las variables se añaden arrastrándolas desde la lista Variables, que se encuentra a la izquierda del lienzo.

El nivel de medición de una variable es importante en el generador de gráficos. Vamos a utilizar la variable *Satisfacción laboral* en el eje X. Sin embargo, el icono (que tiene el aspecto de una regla) que hay junto a la variable indica que el nivel de medición se ha definido como escala. Para crear el gráfico correcto, deberá utilizar un nivel de medición categórico. En lugar de retroceder y cambiar el nivel de medición en la Vista de variables, puede cambiar el nivel de medición temporalmente en el Creador de gráficos.

1. Pulse con el botón derecho del ratón en *Satisfacción laboral* en la lista de variables y elija **Ordinal**. El nivel de medición Ordinal es adecuado ya que las categorías de *Satisfacción laboral* se pueden ordenar por nivel de satisfacción. Tenga en cuenta que el icono cambia después de cambiar el nivel de medición.
2. Ahora arrastre *Satisfacción laboral* de la lista de variables a la zona de colocación del eje X.
La zona de colocación del eje y de forma predeterminada tiene el estadístico *Recuento*. Si desea utilizar otro estadístico (como el porcentaje o la media), puede cambiarlo fácilmente. No utilizará ninguna de estas estadísticas en este ejemplo, pero revisaremos el proceso en caso de que necesite cambiar esta estadística en otro momento.
3. Pulse la pestaña **Propiedades** que aparece en la barra lateral del Generador de gráficos. (Si la barra lateral no se visualiza, pulse el botón de la esquina superior derecha del Generador de gráficos para que se muestre la barra lateral.)



Figura 34. Propiedades del elemento

Propiedades del elemento permite cambiar las propiedades de los diferentes elementos del gráfico. Entre éstos se incluyen los elementos gráficos (como las barras del gráfico de barras) y los ejes del gráfico. Seleccione uno de los elementos de la lista Editar propiedades de para cambiar las propiedades asociadas a dicho elemento. Observe también la X roja que aparece en la parte derecha de la lista. Este botón permite eliminar un elemento gráfico del lienzo. Dado que **Bar1** está seleccionado, las propiedades mostradas se aplican a los elementos gráficos, específicamente al elemento gráfico de la barra.

La lista desplegable Estadístico muestra los estadísticos concretos que están disponibles. Normalmente, para todos los tipos de gráficos están disponibles los mismos estadísticos. Tenga en cuenta que algunas estadísticas requieren que la zona de colocación del eje y contenga una variable.

4. Arrastre *Ingresos del hogar en miles* de la lista de variables a la zona de colocación del eje y. Al ser la variable del eje Y escalar y la variable del eje X escalar (ordinal es un tipo de nivel de medición categórico), en la zona de colocación del eje Y aparece de forma predeterminada el estadístico *Media*. Estas son las variables y las estadísticas que desea, por lo que no es necesario cambiar las propiedades del elemento.

Adición de texto

También puede añadir títulos y notas al pie de página al gráfico.

1. Pulse en la pestaña **Títulos/notas al pie**.
2. Seleccione **Título 1**.

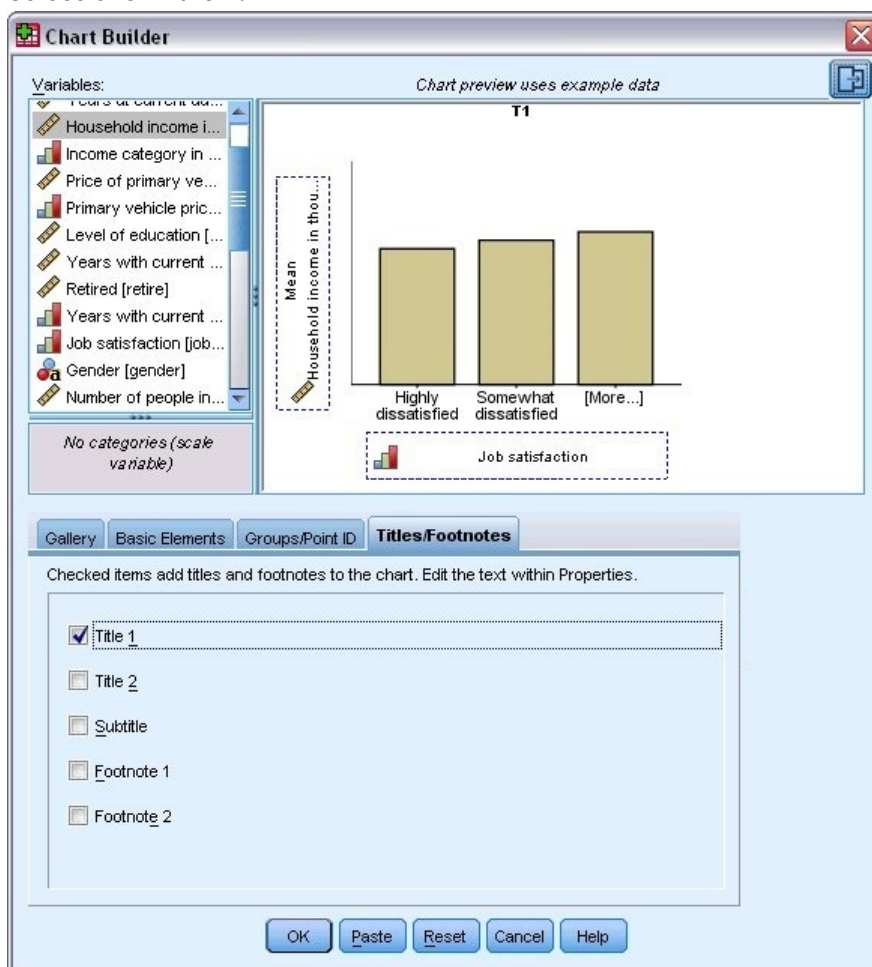


Figura 35. Título 1 mostrado en el lienzo

El título aparece en el lienzo con la etiqueta **T1**.

3. En la pestaña **Propiedades del elemento**, seleccione **Título 1** en la lista Editar propiedades de.
4. En el cuadro de texto Contenido, escriba Income by Job Satisfaction. Este es el texto que mostrará el título.

Crear el gráfico

1. Pulse **Aceptar** para crear el gráfico de barras.

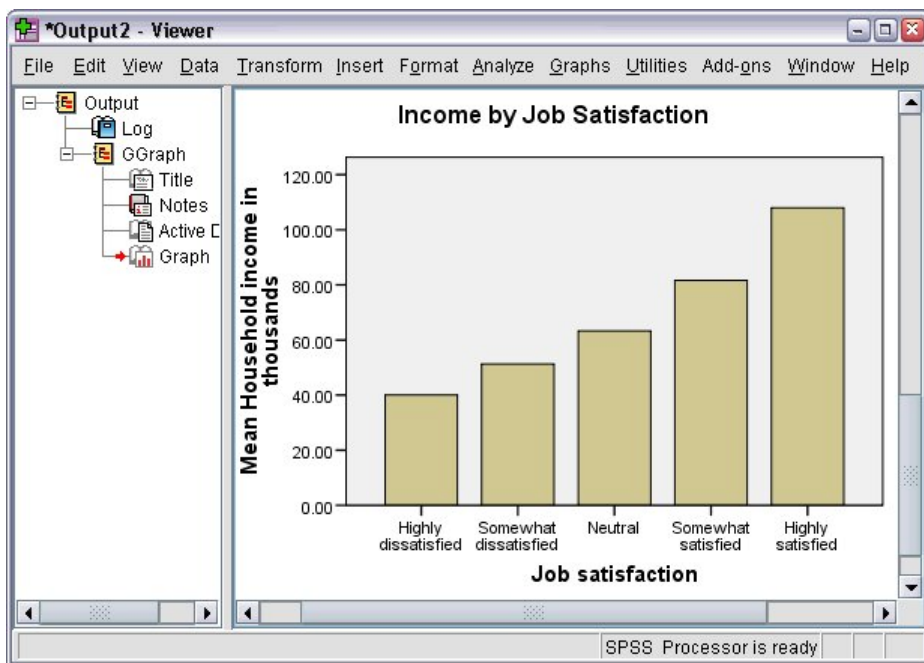


Figura 36. Gráfico de barras

El gráfico de barras revela que los encuestados que están más satisfechos con sus trabajos tienden a tener mayores ingresos familiares.

Capítulo 6. Cómo trabajar con la salida

Los resultados de ejecutar un procedimiento estadístico se muestran en el Visor. Los resultados obtenidos pueden ser tablas estadísticas, diagramas, gráficos o texto, dependiendo de las opciones tomadas al ejecutar el procedimiento. Esta sección utiliza los archivos *viewertut.spv* y *demo.sav*. Consulte Capítulo 10, “Archivos de muestra”, en la página 81 para obtener más información.

Uso del Visor

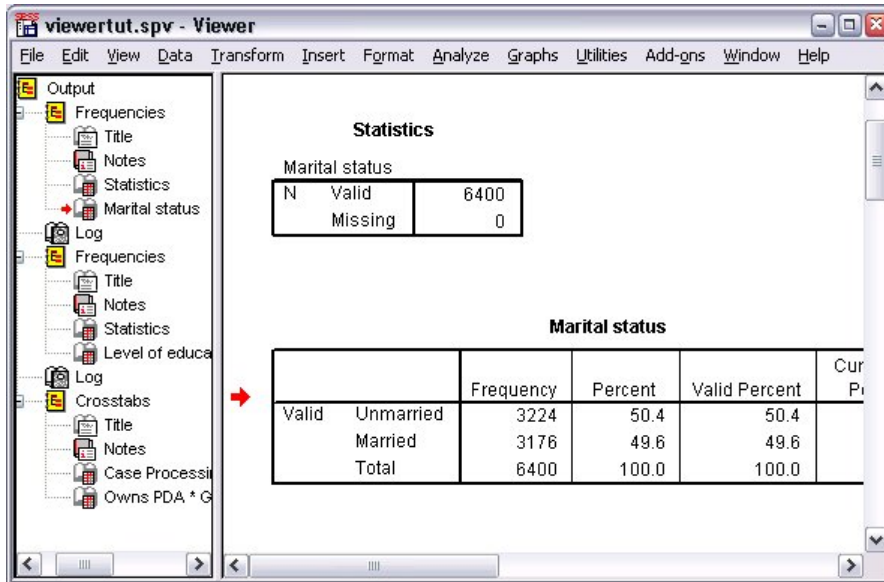


Figura 37. Visor

La ventana Visor se divide en dos paneles. El **panel de titulares** contiene un esquema de titulares con toda la información almacenada en el Visor. El **panel de contenidos** contiene tablas estadísticas, gráficos y resultados de texto.

Utilice las barras de desplazamiento para desplazarse a través de los contenidos de la ventana, tanto vertical como horizontalmente. Para simplificar el acceso, pulse en un elemento en el panel de titulares para visualizarlo en el panel de contenidos.

1. Pulse y arrastre el borde derecho del panel de titulares para modificar su anchura.

Un icono de libro abierto en el panel de esquema indica que actualmente está visible en el visor, aunque puede que no esté actualmente en la parte visible del panel de contenido.

2. Para ocultar una tabla o un gráfico, pulse dos veces en su icono de libro del panel de titulares.

El icono de libro abierto cambia a un icono de libro cerrado, lo que significa que la información asociada a él ahora está oculta.

3. Pulse dos veces en el icono de libro cerrado para volver a mostrar los resultados ocultos.

También puede ocultar toda la salida de un determinado procedimiento estadístico o de toda la salida en el Visor.

4. Para ocultar todos los resultados, pulse la casilla con el signo menos (-) a la izquierda del procedimiento que desea ocultar o en la casilla que está junto al elemento situado más arriba en los titulares.

Los titulares se contraen, indicando visualmente que dichos resultados están ocultos.

También puede cambiar el orden en el que se muestran los resultados.

5. En el panel de titulares, pulse en los elementos que desea mover.
6. Arrastre los elementos seleccionados a una nueva ubicación en los titulares.

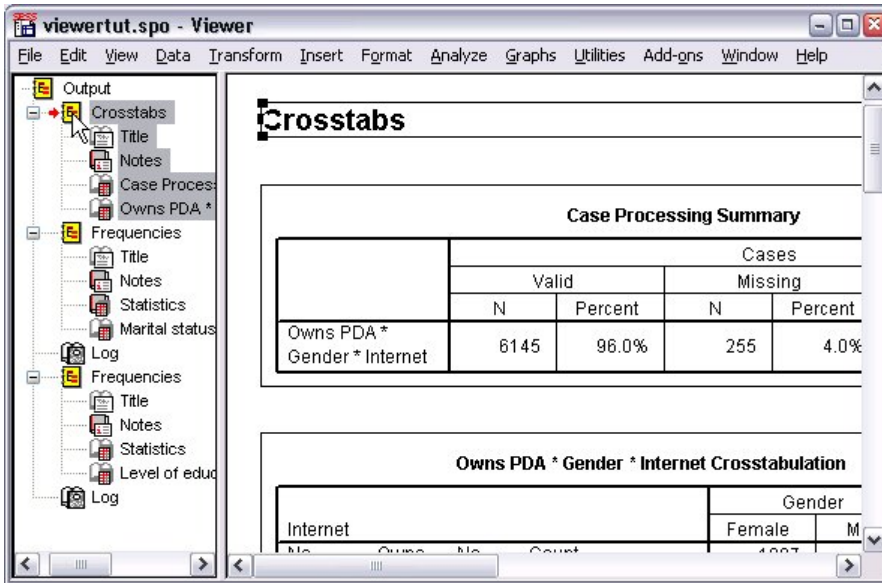


Figura 38. Resultados reordenados en el Visor

También puede mover elementos de la salida pulsando y arrastrándolos en el panel de contenido.

Utilización del editor de tablas dinámicas

Los resultados de la mayoría de los procedimientos estadísticos se muestran en **tablas dinámicas**.

Acceso a las definiciones de resultados

Muchos términos estadísticos se muestran en los resultados. Se puede acceder directamente a las definiciones de dichos términos en el propio Visor.

1. Efectúe una doble pulsación en la tabla *Tabla cruzada Posee PDA * Sexo * Internet*.
2. Haga clic con el botón derecho en *Recuento esperado* y seleccione **¿Qué es esto?** en el menú emergente.

La definición se visualiza en una ventana emergente.

Owns PDA * Gender * Internet Crosstabulation

Internet			Gender		Total		
			Female	Male			
No	Owns PDA	Count	1897	1962	3859		
		Expected Count	1903.4	1955.6	3859.0		
	% within Owns PDA	50.3%	49.7%	100.0%			
	Total	Count	2224	2285	4509		
Yes	Owns PDA	Count	513	512	1025		
		Expected Count	512.5	512.5	1025.0		
	% within Owns PDA	50.0%	50.0%	100.0%			
	Total	Count	305	306	611		
			Yes	Count	305.5	305.5	611.0
			Total	Count	4509	4509	9018

The number of cases that would be expected in the cell if the row and column variables are statistically independent or unrelated to one another.

Figura 39. Definición emergente

Tablas dinámicas

Es posible que las tablas generadas de forma predeterminada no muestren la información con la claridad o nitidez deseadas. Con las tablas dinámicas, se pueden transponer filas y columnas ("voltear" la tabla), ajustar el orden de los datos en una tabla y modificar la tabla de muchas otras maneras. Por ejemplo, puede cambiar una tabla corta y ancha en una larga y estrecha, simplemente transponiendo las filas y las columnas. La modificación del diseño de la tabla no afecta a los resultados. En cambio, es una forma de mostrar la información de una manera distinta o más adecuada.

1. Si todavía no está activado, efectúe una doble pulsación en la *Tabla cruzada Posee PDA * Sexo * Internet* para activarla.
2. Si la ventana *Bandejas dinámicas* no está visible, seleccione en los menús:

Pivote > Bandejas de pivote

Las bandejas de pivote proporcionan una manera de mover datos entre columnas, filas y capas.

Internet			Gender		Total		
			Female	Male			
No	Owns PDA	Count	1897	1962	3859		
		Expected Count	1903.4	1955.6	3859.0		
	% within Owns PDA	49.2%	50.8%	100.0%			
	Count	327	323	650			
Yes	Owns PDA	Count	513	512	1025		
		Expected Count	512.5	512.5	1025.0		
	% within Owns PDA	50.0%	50.0%	100.0%			
	Count	305	306	611			
			Yes	Count	305.5	305.5	611.0
			Total	Count	4509	4509	9018

Pivoting Trays

Gender (COLUMN)

Owns PDA (ROW)

Internet (LAYER)

Figura 40. Bandejas de pivote

3. Arrastre el elemento *Estadísticos* de la dimensión de filas a la dimensión de columnas, debajo de *Género*. La tabla se reconfigurará de forma inmediata para reflejar los cambios.

El orden de los elementos en la bandeja de pivote refleja el orden de los elementos en la tabla.

4. Arrastre y suelte el elemento *Posee PDA* antes del elemento *Internet* en la dimensión de fila para invertir el orden de estas dos filas.

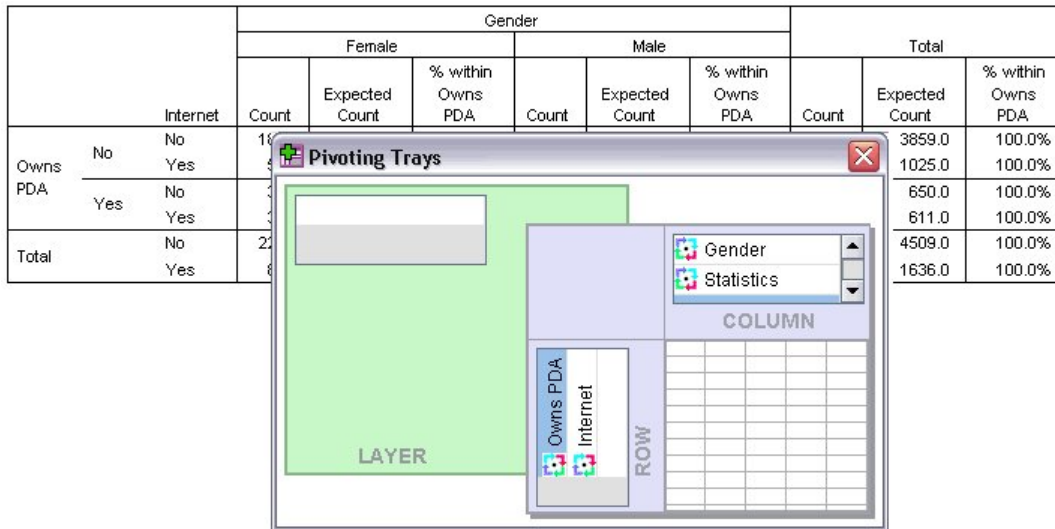


Figura 41. Intercambiar filas

Creación y visualización de capas

Las capas pueden ser de utilidad para tablas grandes con categorías de información anidadas. Al crear capas, se simplifica el aspecto de la tabla, haciendo más sencilla su lectura.

1. Arrastre el elemento *Sexo* de la dimensión Columna a la dimensión Capa.

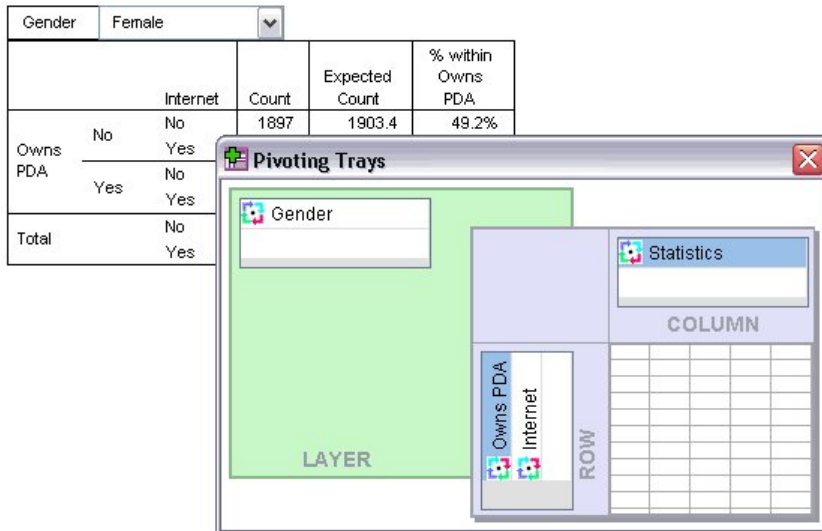


Figura 42. Icono de pivotado Género en la dimensión Capa

Para mostrar una capa distinta, seleccione una categoría de la lista desplegable en la tabla.

Edición de tablas

A menos que se haya tomado la molestia de crear un Aspecto de tabla personalizado, las tablas dinámicas se crearán con el formato estándar. Se puede modificar el formato de cualquier texto dentro de

una tabla. Entre los formatos que se pueden cambiar se incluyen el nombre de la fuente, el tamaño de la fuente, el estilo de la fuente (negrita o cursiva) y el color.

1. Efectúe una doble pulsación en la tabla *Nivel de educación*.
2. Si la barra de herramientas Formato no está visible, seleccione en los menús:

Vista > Barra de herramientas

3. Pulse el texto del título, *Nivel de educación*.
4. En la lista desplegable de tamaños de fuente de la barra de herramientas, seleccione **12**.
5. Para cambiar el color del texto del título, pulse en la herramienta de color de texto y elija un nuevo color.

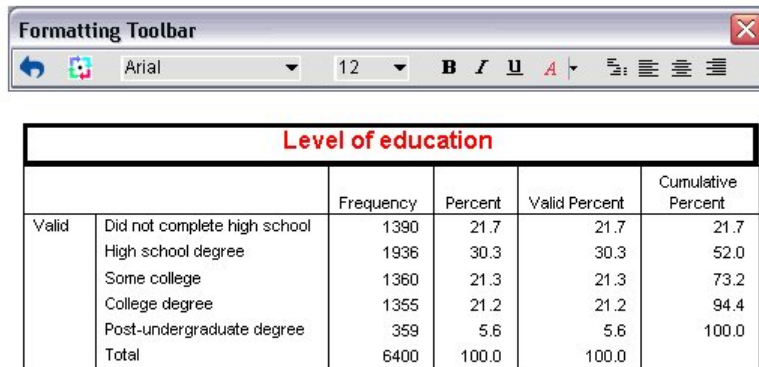


Figura 43. Nuevo formato del Texto del título en la tabla dinámica

También puede editar el contenido de las tablas y etiquetas. Por ejemplo, puede cambiar el título de esta tabla.

6. Efectúe una doble pulsación en el título.
7. Escriba Education Level para la nueva etiqueta.

Note: si se modifican los valores de una tabla, no se volverán a calcular los valores totales ni los demás estadísticos.

Ocultación de filas y columnas

Algunos de los datos que se muestran en una tabla pueden no ser útiles o pueden complicar innecesariamente la tabla. Afortunadamente, puede ocultar filas y columnas enteras sin perder datos.

1. Si todavía no está activada, efectúe una doble pulsación en la tabla *Nivel educativo* para activarla.
2. Pulse la etiqueta de columna *Porcentaje válido* para seleccionarlo.
3. En el menú **Editar** o en el menú emergente que aparece al pulsar el botón derecho del ratón, elija:

Seleccione > Celdas de datos y etiquetas

4. En el menú **Ver**, elija **Ocultar** o en el menú emergente que aparece al pulsar el botón derecho del ratón seleccione **Ocultar categoría**.

La columna ahora está oculta pero no se ha suprimido.

Education Level

		Frequency	Percent	Cumulative Percent
Valid	Did not complete high school	1390	21.7	21.7
	High school degree	1936	30.3	52.0
	Some college	1360	21.3	73.2
	College degree	1355	21.2	94.4
	Post-undergraduate degree	359	5.6	100.0
	Total	6400	100.0	

Figura 44. Columna Porcentaje válido oculta en una tabla

5. Para volver a visualizar la columna, efectúe una doble pulsación en la misma tabla.
6. Pulse **Ver** y seleccione **Mostrar todas las categorías**.

Las filas también se pueden ocultar y volver a visualizar de la misma forma que las columnas.

Cambiar los formatos de visualización de datos

Puede cambiar fácilmente el formato de visualización de los datos en las tablas dinámicas.

1. Si todavía no está activada, efectúe una doble pulsación en la tabla *Nivel educativo* para activarla.
2. Pulse en la etiqueta de columna *Porcentaje* para seleccionarla.
3. En el menú Edición o el menú emergente que aparece al pulsar el botón derecho, seleccione:

Seleccione > Celdas de datos

4. En el menú Edición o el menú emergente que aparece al pulsar el botón derecho, seleccione **Propiedades de casilla**.
5. Pulse en la pestaña **Valor de formato**.
6. Escriba 0 en el campo Decimales para ocultar todas las separaciones decimales de esta columna.

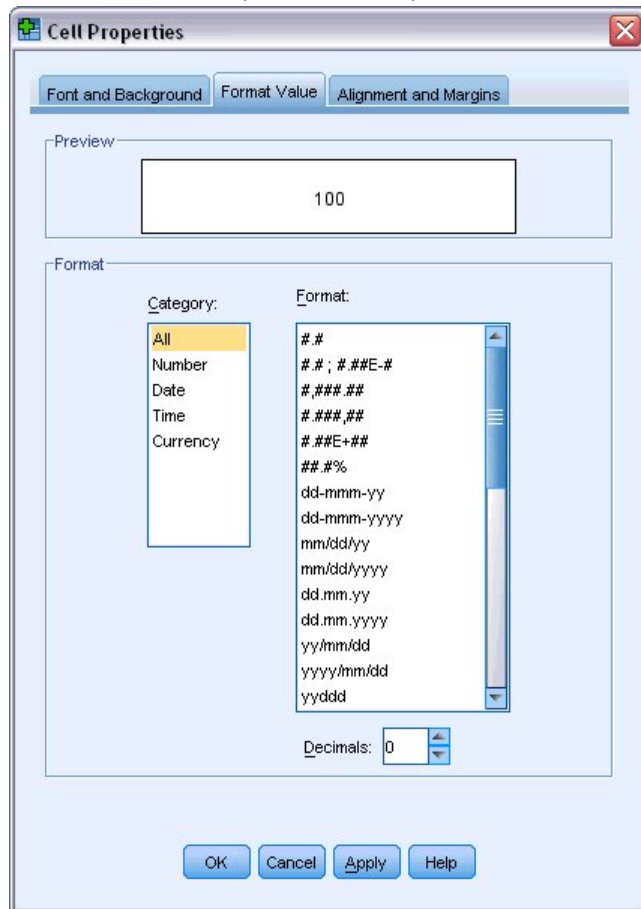


Figura 45. Pestaña Valor de formato del cuadro de diálogo Propiedades de casilla

También puede cambiar el tipo de datos y el formato en este cuadro de diálogo.

7. Seleccione el tipo que desee en la lista Categoría y después seleccione el formato para ese tipo en la lista Formato.
8. Pulse **Aceptar** o **Aplicar** para aplicar los cambios.

Education Level

		Frequency	Percent	Cumulative Percent
Valid	Did not complete high school	1390	22	21.7
	High school degree	1936	30	52.0
	Some college	1360	21	73.2
	College degree	1355	21	94.4
	Post-undergraduate degree	359	6	100.0
Total		6400	100	

Figura 46. Decimales ocultos en la columna *Porcentaje*

Los decimales están ahora ocultos en la columna *Porcentaje*.

TableLooks

El formato de las tablas es una parte muy importante a la hora de proporcionar resultados claros, concisos y con significado. Si la tabla es difícil de leer, es posible que la información contenida en dicha tabla no se entienda fácilmente.

Utilización de formatos predefinidos

1. Efectúe una doble pulsación en la tabla *Estado civil*.
2. En los menús seleccione:

Formato > Aspectos de tabla...

El cuadro de diálogo Aspectos de la tabla recoge una variedad de estilos predefinidos. Seleccione un estilo de la lista para previsualizarlo en la ventana Ejemplo de la derecha.

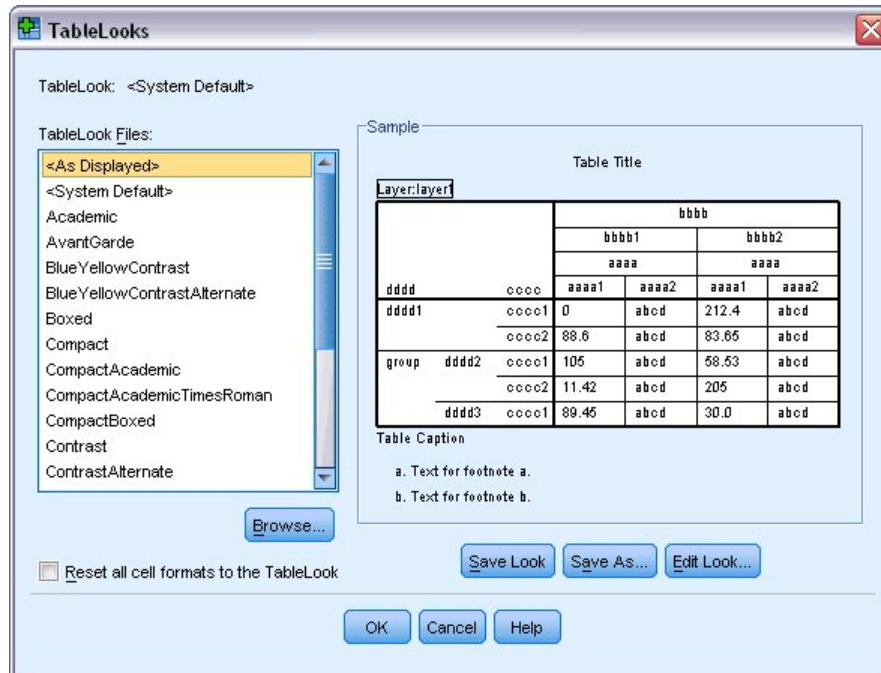


Figura 47. Cuadro de diálogo Aspectos de tabla

Puede utilizar un estilo tal cual, o puede editar un estilo existente para adaptarlo mejor a sus necesidades.

3. Para utilizar un estilo existente, seleccione uno y pulse **Aceptar**.

Personalización de estilos de Aspecto de tabla

Puede personalizar un formato para que se ajuste a sus necesidades específicas. Se pueden personalizar casi todos los aspectos de una tabla, desde el color de fondo hasta los estilos de los bordes.

1. Efectúe una doble pulsación en la tabla *Estado civil*.
2. En los menús seleccione:

Formato > Aspectos de tabla...

3. Seleccione el estilo que más se aproxime al formato que quiere y pulse en **Editar aspecto**.
4. Pulse el separador **Formatos de celda** para ver las opciones de formato.

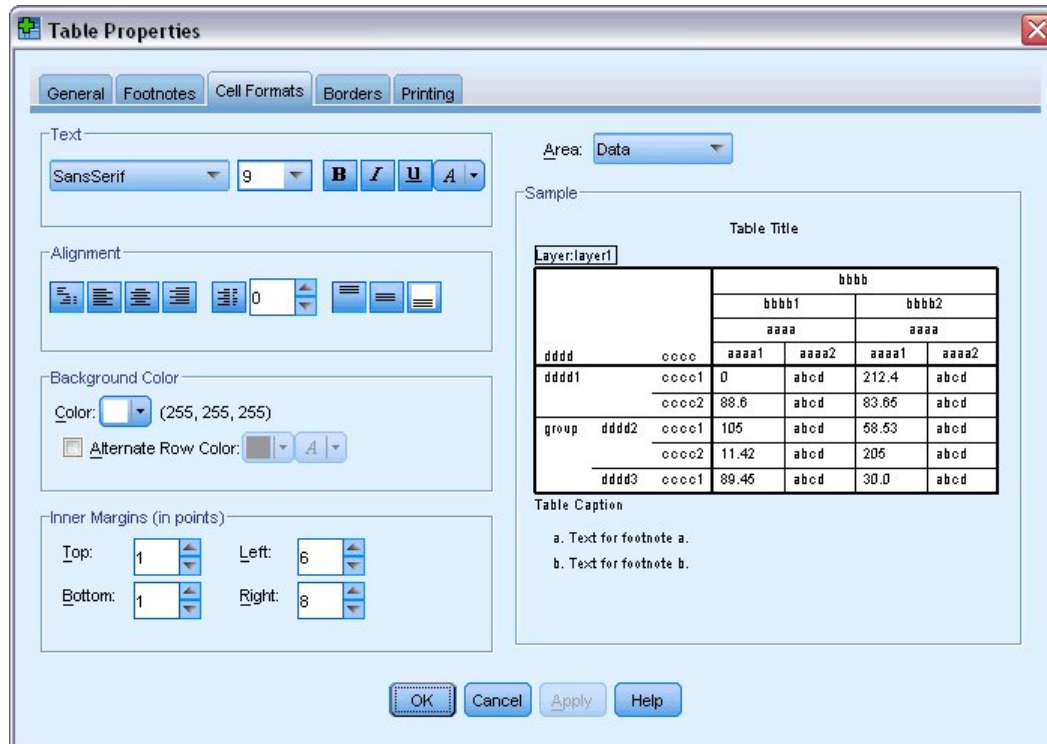


Figura 48. Cuadro de diálogo Propiedades de tabla

Entre las opciones de formato se incluyen el nombre de la fuente, el tamaño de la fuente, el estilo y el color. Las opciones adicionales incluyen alineación, colores de texto y fondo, y tamaños de margen.

La ventana Muestra situada a la derecha proporciona una vista previa de cómo afectan a la tabla los cambios de formato. Cada área de la tabla puede tener un estilo de formato diferente. Por ejemplo, probablemente no querrá que el título tenga el mismo formato que los datos. Para seleccionar un área de la tabla para editar, puede seleccionar el área por el nombre en la lista desplegable Área o pulsar en el área que desee cambiar en la ventana Muestra.

5. Seleccione **Datos** de la lista desplegable Área.
6. Seleccione un nuevo color en la paleta desplegable Fondo.
7. A continuación, seleccione un nuevo color de texto.

La ventana Ejemplo muestra el nuevo estilo.

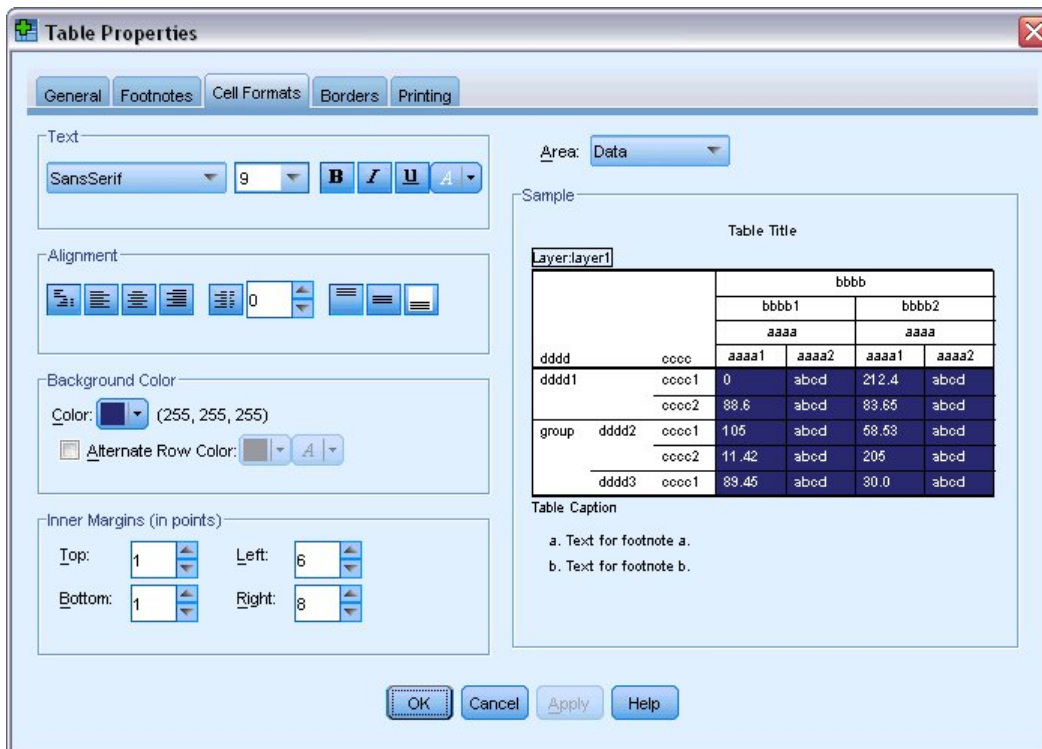


Figura 49. Cambio de los formatos de casilla de la tabla

8. Pulse en **Aceptar** para volver al cuadro de diálogo Aspectos de tabla.

Puede guardar el nuevo estilo para poder aplicarlo fácilmente a tablas futuras.

9. Pulse en **Guardar como**.

10. Acceda al directorio de destino e introduzca un nombre para el nuevo estilo en el cuadro de diálogo Nombre de archivo.

11. Pulse **Guardar**.

12. Pulse **Aceptar** para aplicar los cambios y volver al Visor.

La tabla ahora contiene el formato personalizado que ha especificado.

Figura 50. Aspecto de tabla personalizado

Cambio de los formatos de tablas predeterminadas

Aunque puede cambiar el formato de una tabla después de que se haya creado, puede ser más eficaz cambiar el Aspecto de tabla predeterminado para que no tenga que cambiar el formato cada vez que cree una tabla.

Para cambiar el estilo del Aspecto de tabla predeterminado para las tablas dinámicas, elija en los menús:

Editar > Opciones...

1. Pulse el separador **Tablas dinámicas** del recuadro de diálogo Opciones.

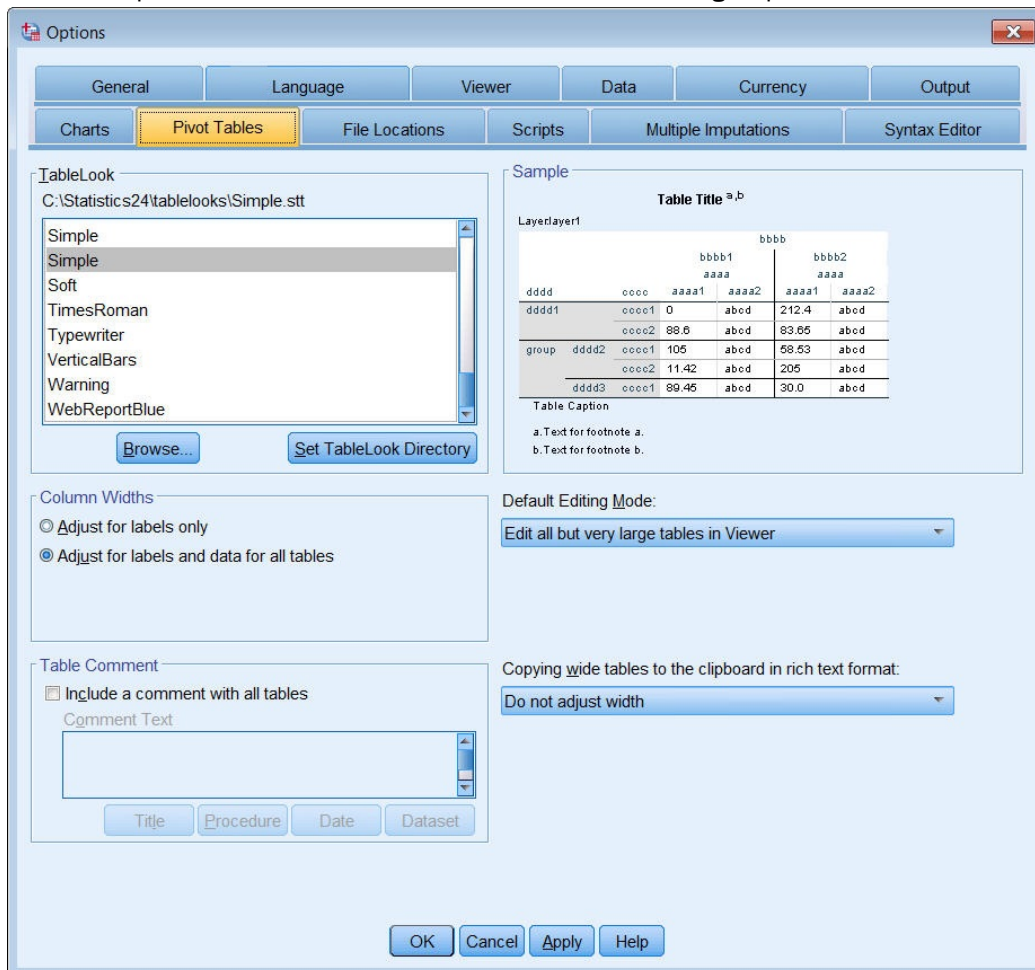


Figura 51. Cuadro de diálogo Opciones

2. Seleccione el estilo de aspecto de tabla que desea utilizar para todas las tablas nuevas.

La ventana Ejemplo de la derecha muestra una vista previa de cada Aspecto de tabla.

3. Pulse **Aceptar** para guardar los valores y cerrar el recuadro de diálogo.

Todas las tablas que cree después de cambiar el Aspecto de tabla predeterminado se ajustan automáticamente a las nuevas reglas de formato.

Personalizar los valores de visualización inicial

Las opciones de presentación inicial incluyen la alineación de objetos en el Visor, si de forma predeterminada se muestran u ocultan los objetos y la anchura de la ventana del Visor. Para cambiar estos valores:

1. En los menús seleccione:

Editar > Opciones...

2. Pulse en la pestaña **Visor**.

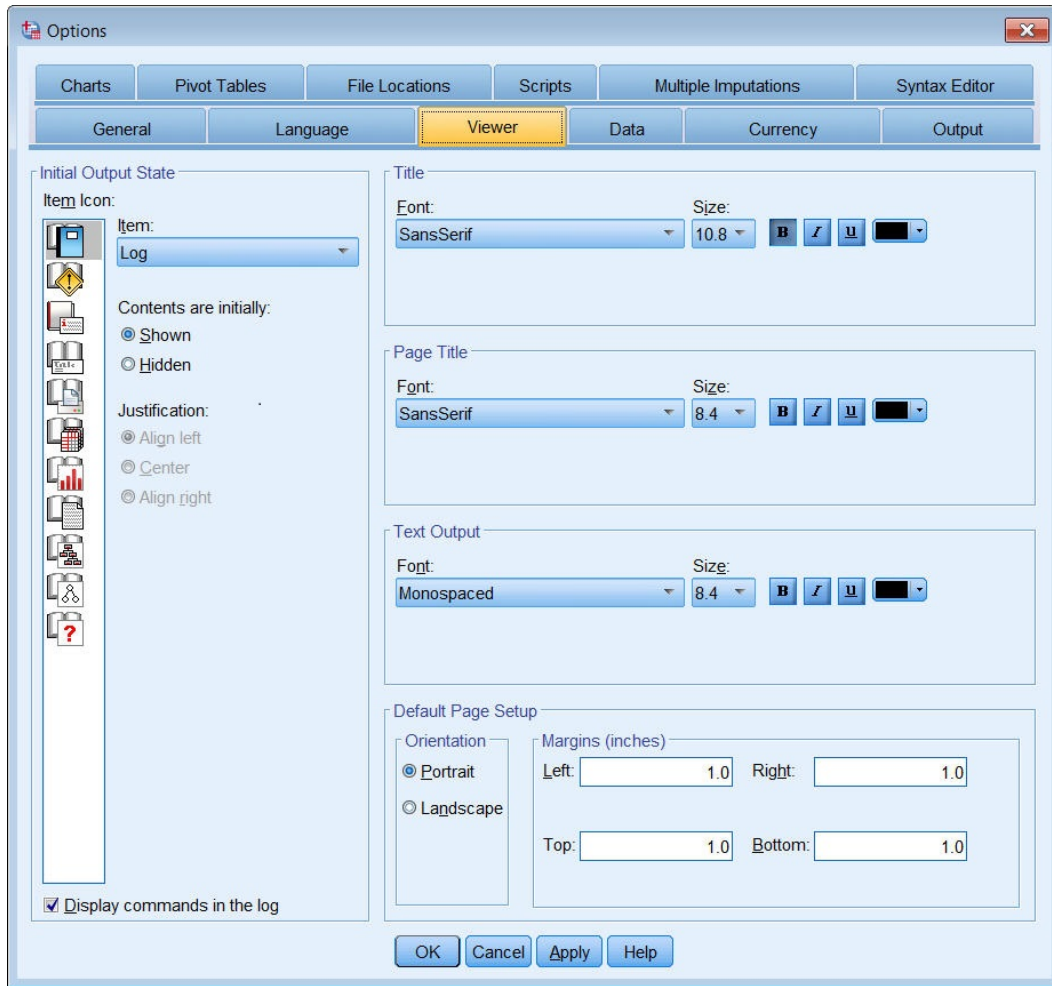


Figura 52. Opciones del Visor

Las opciones se aplican objeto por objeto. Por ejemplo, puede personalizar cómo se muestran los gráficos sin cambiar la forma de presentación de las tablas. Sólo tiene que seleccionar el objeto que desea personalizar y realizar los cambios.

3. Pulse el icono **Título** para visualizar sus valores.

4. Pulse **Centrar** para visualizar todos los títulos en el centro (horizontal) del visor.

También se pueden ocultar elementos, como el registro y los mensajes de advertencia, que normalmente tienden a colapsar los resultados. Al pulsar dos veces en un icono, se cambia de forma automática la propiedad de presentación de dicho objeto.

5. Efectúe una doble pulsación en el icono **Avisos** para ocultar los mensajes de aviso en la salida.

6. Pulse **Aceptar** para guardar los valores y cerrar el recuadro de diálogo.

Visualización de etiquetas de variable y valor

En la mayoría de los casos la presentación de etiquetas de variables y valores es más efectiva que mostrar el nombre de la variable o el valor actual de datos. Sin embargo, pueden darse casos en los que desee mostrar tanto los nombres como las etiquetas.

1. En los menús seleccione:

Editar > Opciones...

2. Pulse en el separador **Etiquetas de salida**.

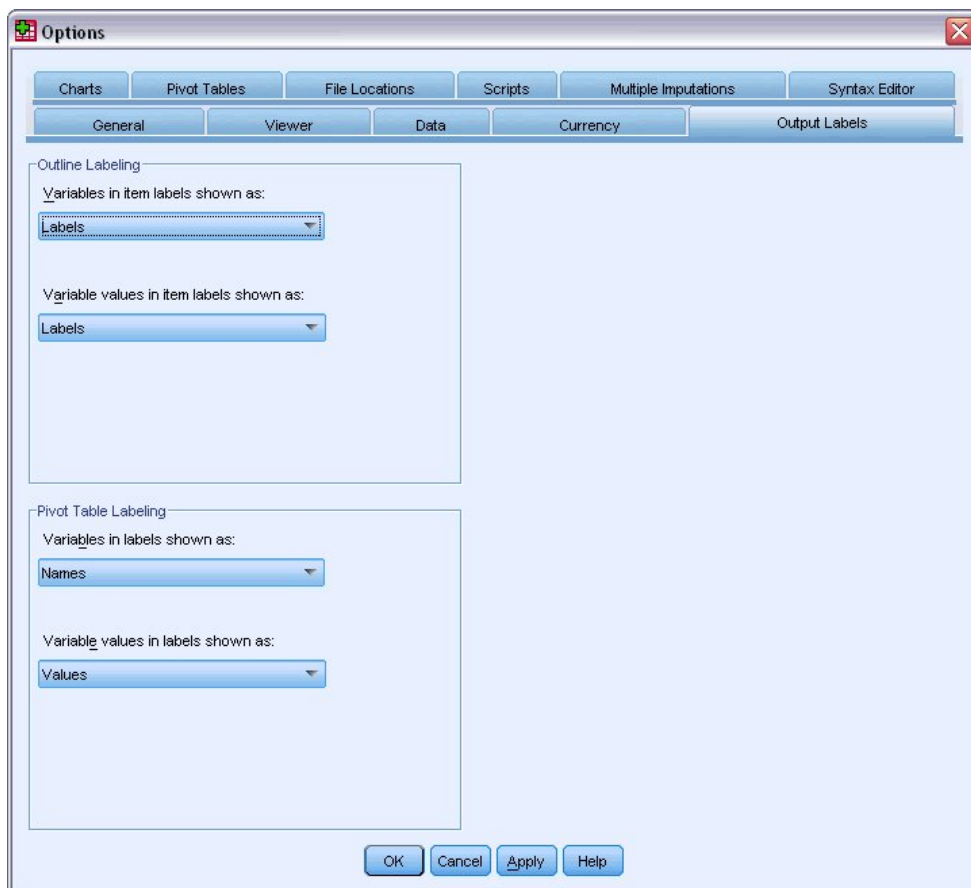


Figura 53. Selecciones de etiquetado de tablas dinámicas

Puede especificar distintas selecciones para los paneles de contenidos y titulares. Por ejemplo, para mostrar las etiquetas en el esquema y los nombres de variable y los valores de datos en el contenido:

3. En el grupo de etiquetado de tabla dinámica, seleccione **Nombres** en la lista desplegable Variables en etiquetas para mostrar nombres de variables en lugar de etiquetas.
4. A continuación, seleccione **Valores** en la lista desplegable Valores de variables en etiquetas para mostrar valores de datos en lugar de etiquetas.

Las tablas subsiguientes producidas en la sesión reflejarán estos cambios.

marital

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	3224	50.4	50.4	50.4
	1	3176	49.6	49.6	100.0
	Total	6400	100.0	100.0	

Figura 54. Nombres y valores de variables mostrados

Uso de resultados en otras aplicaciones

Los resultados se pueden utilizar en muchas aplicaciones. Por ejemplo, es posible que desee incluir una tabla o un gráfico en una presentación o informe.

Los siguientes ejemplos son específicos de Microsoft Word, pero pueden funcionar de forma similar en otras aplicaciones de procesamiento de textos.

Pegar resultados como tablas de Word

Las tablas dinámicas se pueden pegar en Word como si fuesen tablas nativas de Word. Se mantendrán todos los atributos de las tablas, como colores y tamaños de fuentes. Debido a que la tabla se pega en el formato de tabla de Word, puede editarla en Word como cualquier otra tabla.

1. Pulse en una tabla del Visor para activarla.
2. En los menús seleccione:

Editar > Copiar

3. Abra la aplicación de procesamiento de texto.
4. Desde el menú del procesador de texto, elija:

Editar > Pegado especial...

5. Seleccione **Texto con formato (RTF)** en el recuadro de diálogo Pegado especial.
6. Pulse **Aceptar** para pegar los resultados en el documento actual.

La tabla aparecerá en el documento. Puede aplicar formato personalizado, editar los datos y cambiar el tamaño de la tabla para que se ajuste a sus necesidades.

Cómo pegar resultados como texto

Las tablas dinámicas se pueden copiar a otras aplicaciones como texto sin formato. Los estilos de formato no se conservan en este método, pero puede editar los datos de la tabla después de pegarlos en la aplicación de destino.

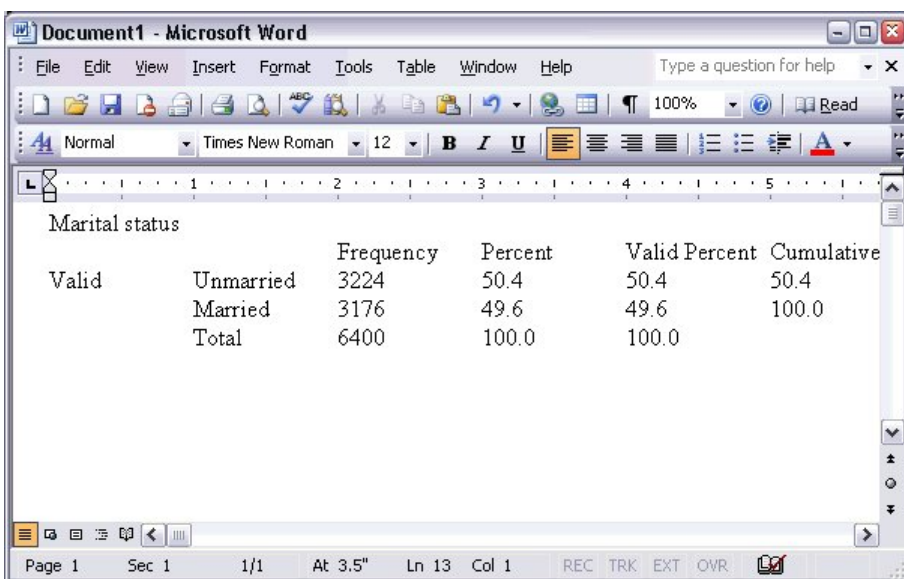
1. Pulse en una tabla del Visor para activarla.
2. En los menús seleccione:

Editar > Copiar

3. Abra la aplicación de procesamiento de texto.
4. Desde el menú del procesador de texto, elija:

Editar > Pegado especial...

5. Seleccione **Texto sin formato** en el recuadro de diálogo Pegado especial.
6. Pulse **Aceptar** para pegar los resultados en el documento actual.



The screenshot shows a Microsoft Word window titled 'Document1 - Microsoft Word'. The menu bar includes File, Edit, View, Insert, Format, Tools, Table, Window, and Help. The toolbar shows various icons for document manipulation. The status bar at the bottom indicates 'Page 1', 'Sec 1', '1/1', 'At 3.5"', 'Ln 13', 'Col 1', and 'REC TRK EXT OVR'. The main content area displays a table with the following data:

Marital status		Frequency	Percent	Valid Percent	Cumulative
Valid	Unmarried	3224	50.4	50.4	50.4
	Married	3176	49.6	49.6	100.0
Total		6400	100.0	100.0	

Figura 55. Tabla dinámica mostrada en Word

Cada columna de la tabla aparece separada por tabulaciones. Puede cambiar los anchos de columna ajustando las detenciones de tabulación en la aplicación de procesamiento de texto.

Exportar resultados a archivos de Microsoft Word, PowerPoint y Excel

Puede exportar resultados a un archivo de Microsoft Word, PowerPoint o Excel. Puede exportar los elementos seleccionados o todos los elementos en el Visor. Esta sección utiliza los archivos *msouuttut.spv* y *demo.sav*. Consulte [Capítulo 10, “Archivos de muestra”](#), en la [página 81](#) para obtener más información.

Nota: la exportación a PowerPoint sólo está disponible en los sistemas operativos Windows y no está disponible con la versión para estudiantes.

En el panel de esquema del Visor, puede seleccionar elementos específicos que desee exportar o bien exportar todos los elementos o todos los elementos visibles.

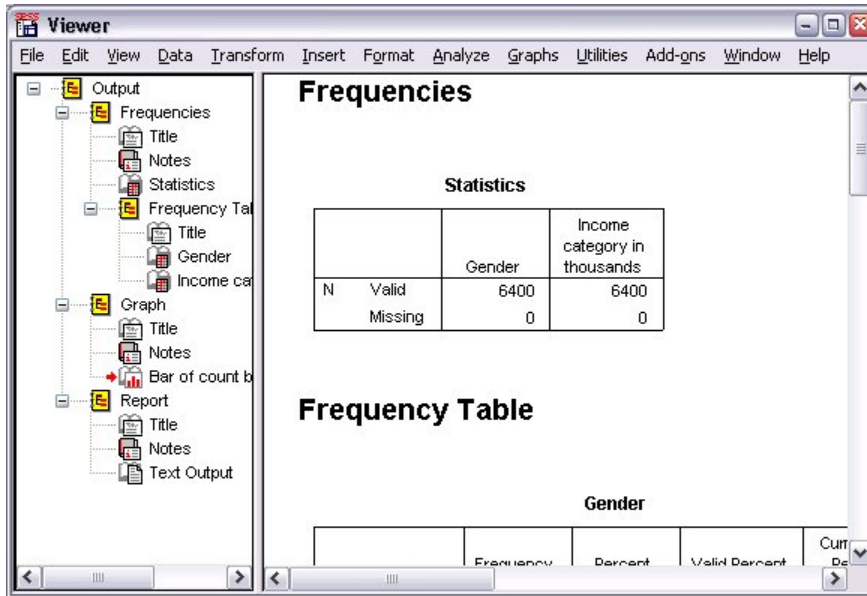


Figura 56. Visor

1. Seleccione en los menús del Visor:

Archivo > Exportar...

En lugar de exportar todos los objetos del Visor, puede optar por exportar únicamente objetos visibles (libros abiertos en el panel de titulares) o sólo los seleccionados en el panel de titulares. Si no ha seleccionado ningún elemento en el panel de esquema, no tiene la opción de exportar objetos seleccionados.

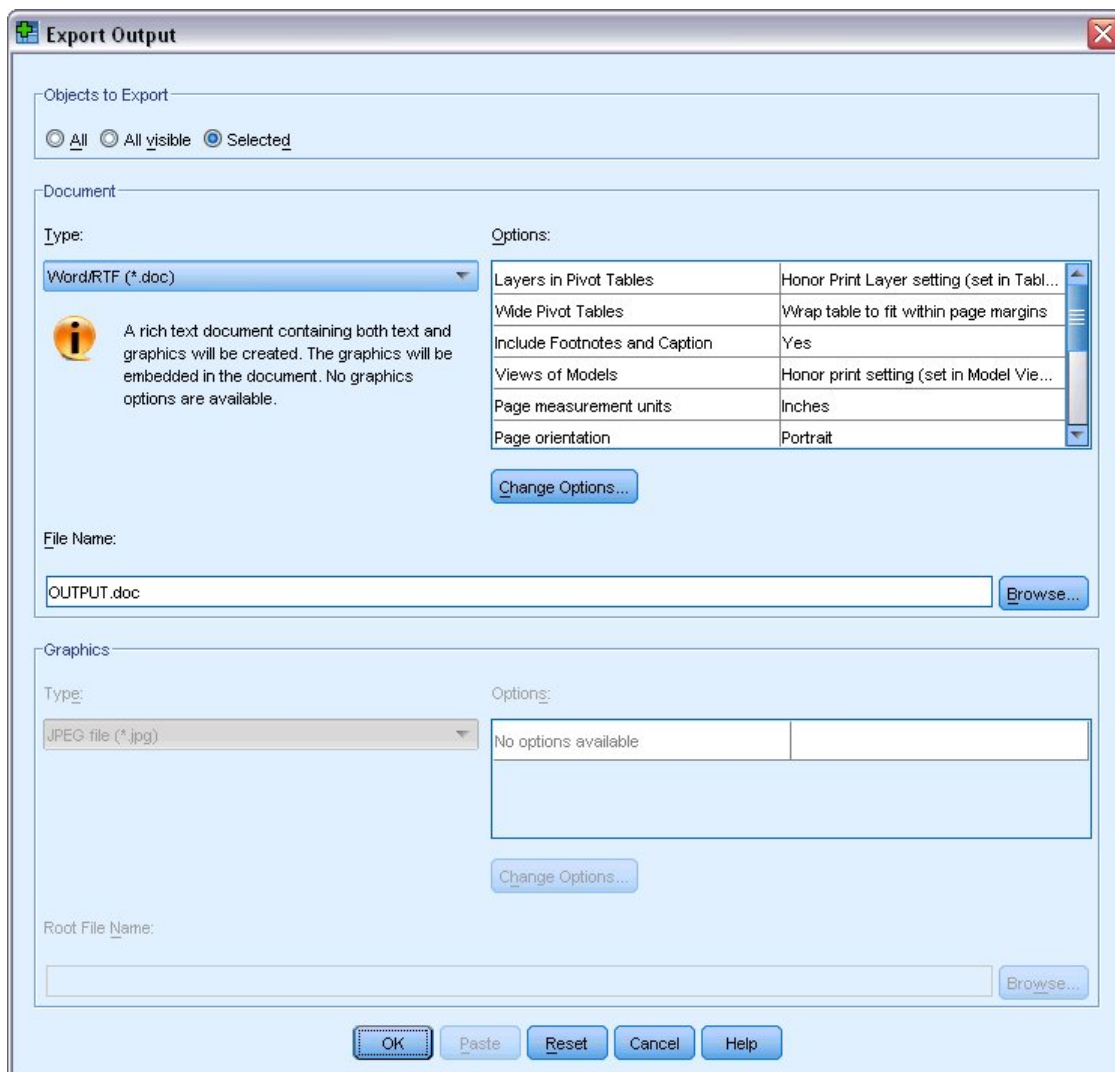


Figura 57. Cuadro de diálogo Exportar resultados

2. En el grupo Objetos a exportar, seleccione **Todos**.
3. En la lista desplegable Tipo, seleccione **Archivo Word/RTF (*.doc)**.
4. Pulse **Aceptar** para generar el archivo Word.

Al abrir el archivo resultante en Word, podrá observar cómo se exportan los archivos. Las notas, que no son objetos visibles, aparecen en Word porque ha elegido exportar todos los objetos.

Las tablas dinámicas se convierten en tablas de Word, en las que se mantiene todo el formato de la tabla dinámica original, incluyendo fuentes, colores, bordes, etc.

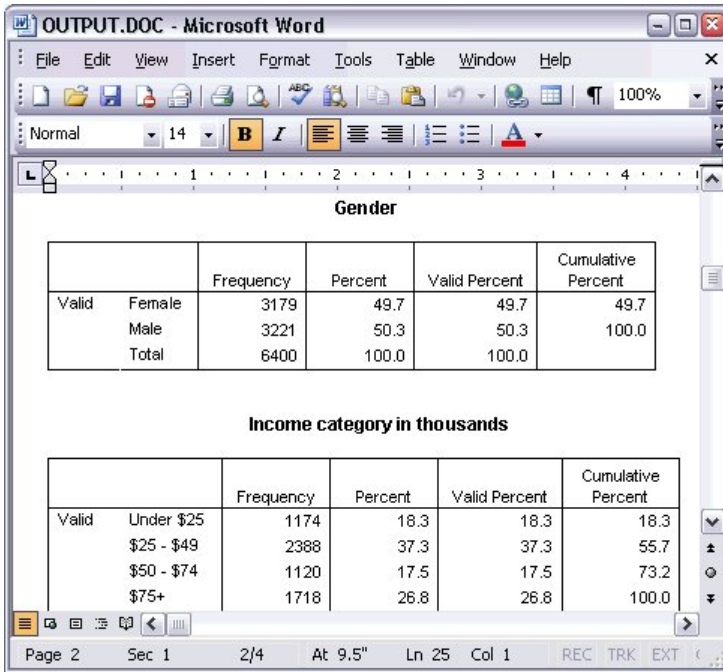


Figura 58. Tablas dinámicas en Word

Los gráficos se incluyen en el documento de Word como imágenes gráficas.

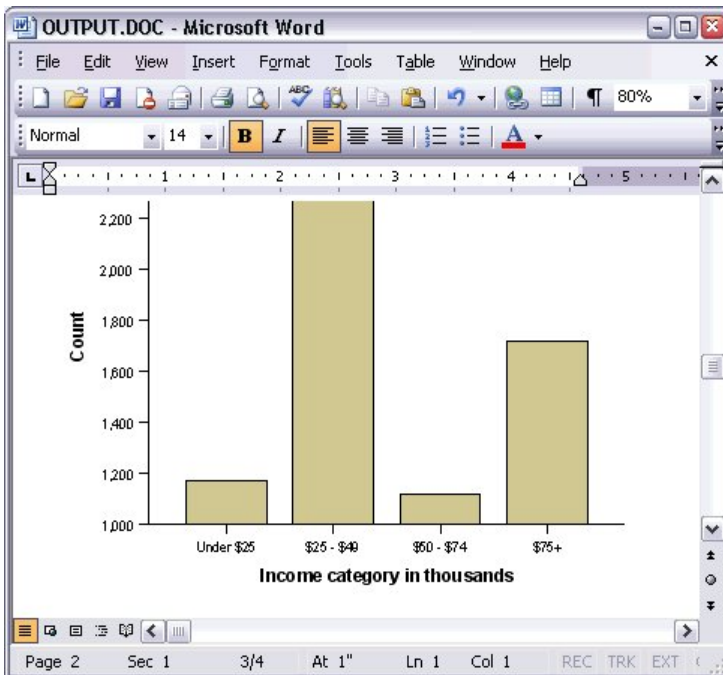


Figura 59. Gráficos en Word

Los resultados de texto aparecen en la misma fuente utilizada para el objeto de texto en el Visor. Para una alineación adecuada, la salida de texto debe utilizar una fuente de paso fijo (monoespaciado).

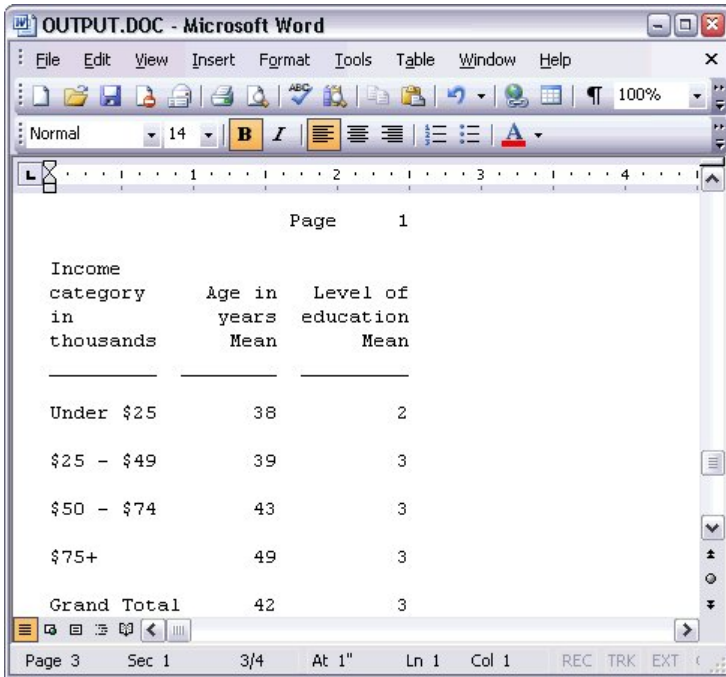


Figura 60. Resultados de texto en Word.

Si exporta a un archivo de PowerPoint, cada elemento exportado se colocará en una diapositiva diferente. Las tablas dinámicas exportadas a PowerPoint se convierten en tablas de Word, con todo el formato de la tabla dinámica original, incluyendo fuentes, colores, bordes, etc.

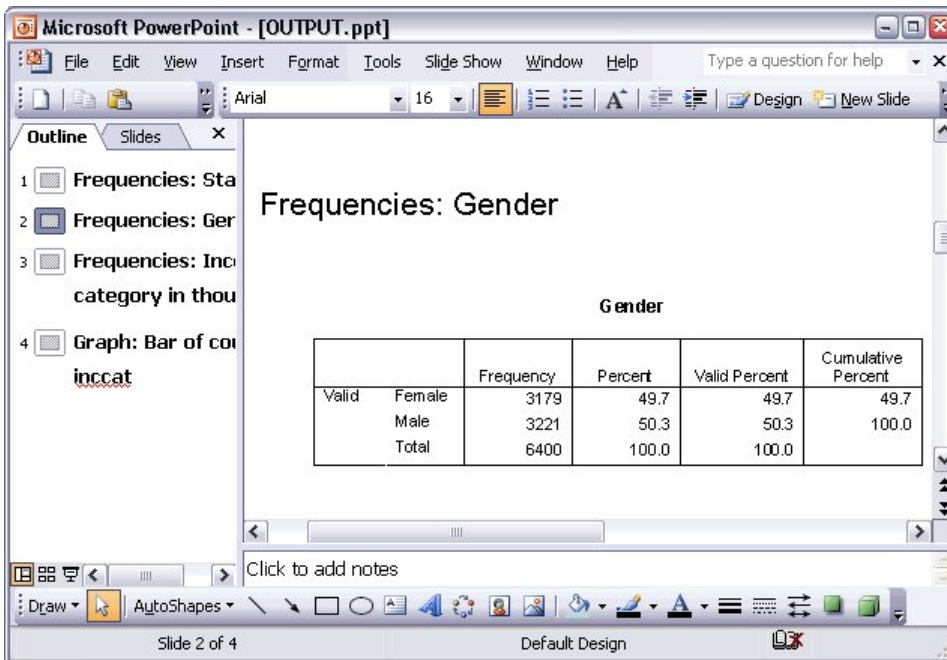


Figura 61. Tablas dinámicas en PowerPoint

Los gráficos seleccionados para la exportación a PowerPoint están incluidos en el archivo de PowerPoint.

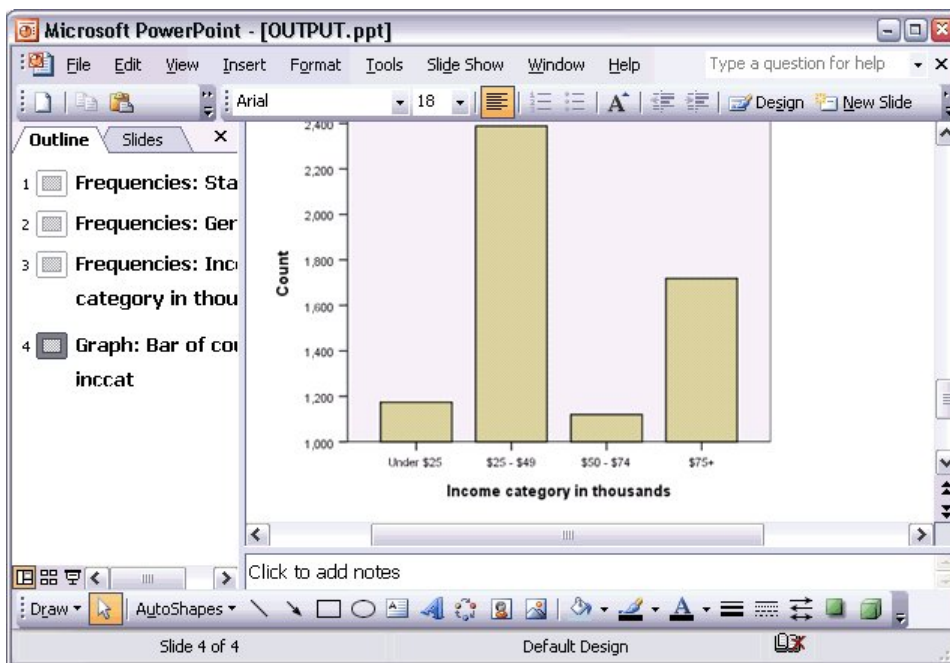


Figura 62. Gráficos en PowerPoint

Nota: la exportación a PowerPoint sólo está disponible en los sistemas operativos Windows y no está disponible con la versión para estudiantes.

Si exporta a un archivo Excel, los resultados se exportan de forma diferente.

The PivotTable data is as follows:

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Female	3,179	49.7	49.7	49.7
	Male	3,221	50.3	50.3	100.0
	Total	6,400	100.0	100.0	
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Under \$25	1,174	18.3	18.3	18.3
	\$25 - \$49	2,388	37.3	37.3	55.7
	\$50 - \$74	1,120	17.5	17.5	73.2
	\$75+	1,718	26.8	26.8	100.0
	Total	6,400	100.0	100.0	

Figura 63. Output.xls en Excel

Las filas, columnas y celdas de la tabla dinámica se convierten en filas, columnas y celdas de Excel.

Cada línea del resultado de texto constituye una fila del archivo de Excel y se incluye todo su contenido en una sola casilla.

The screenshot shows a Microsoft Excel window titled "Microsoft Excel - OUTPUT.xls". The spreadsheet contains the following data:

	A	B	C	D	E
68	Income				
69	category	Age in	Level of		
70	in	years	education		
71	thousands	Mean	Mean		
72					
73					
74	Under \$25	38	2		
75					
76	\$25 - \$49	39	3		
77					
78	\$50 - \$74	43	3		
79					
80	\$75+	49	3		
81					
82	Grand Total	42	3		

Figura 64. Resultados de texto en Excel

Exportar resultados a PDF

Puede exportar los elementos seleccionados o todos los elementos en el Visor a un archivo PDF (formato de documento portátil).

1. Elija en los menús de la ventana Visor que contiene el resultado que desea exportar a PDF:

Archivo > Exportar...

2. En el recuadro de diálogo Exportar salida, seleccione **Portable Document Format** en la lista desplegable Tipo de archivo de formato de exportación.

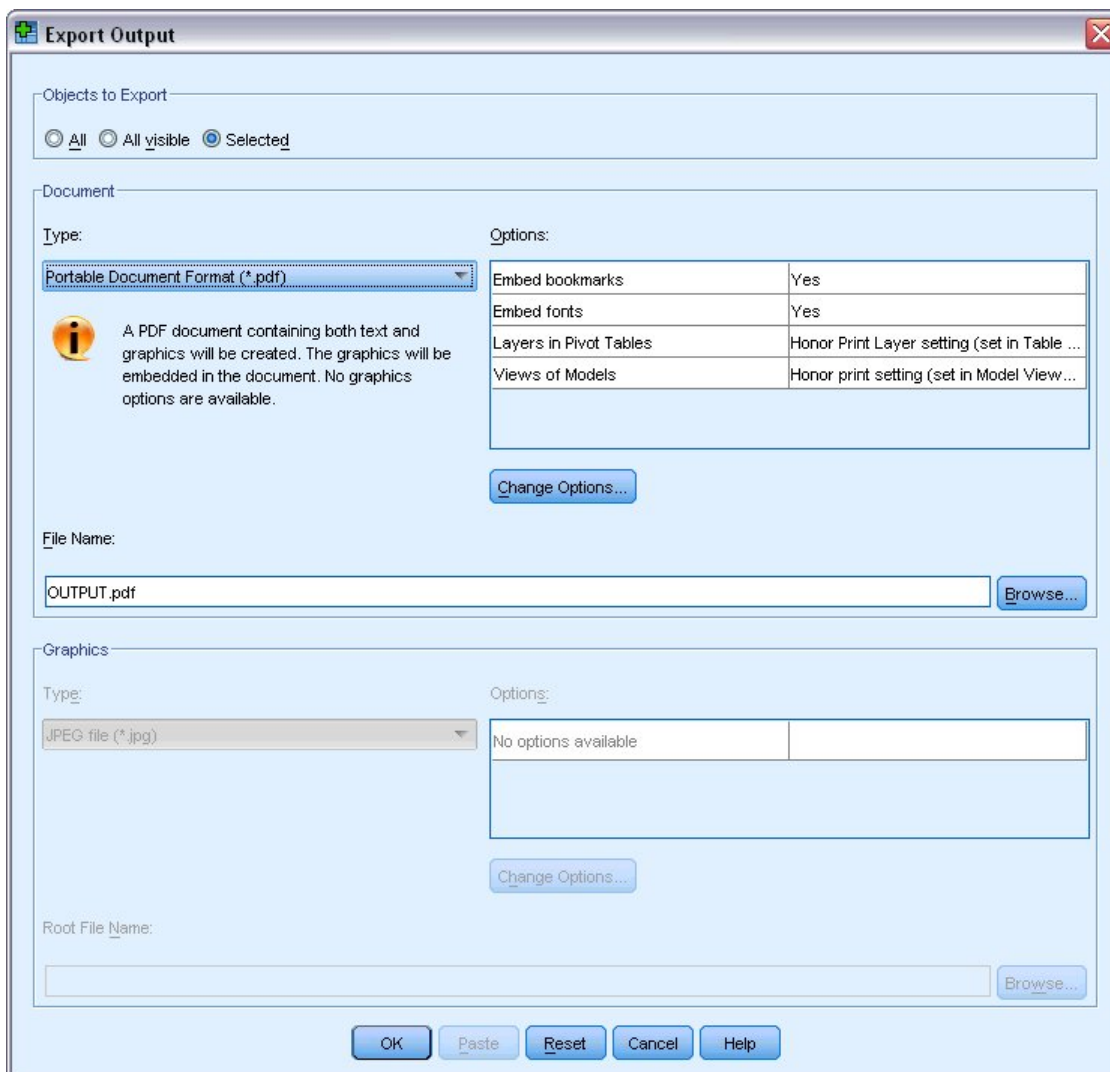


Figura 65. Cuadro de diálogo Exportar resultados

- El panel de esquema del documento del Visor se convierte a marcadores en el archivo PDF para facilitar la navegación.
- El tamaño de página, la orientación, los márgenes, el contenido y la visualización de encabezados y pies de página y el tamaño de gráficos impresos en documentos PDF se controlan mediante opciones de configuración de página (menú Archivo, Configuración de página en la ventana del Visor).
- La resolución (PPP) del documento PDF es la configuración de la resolución actual para la impresora predeterminada 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.

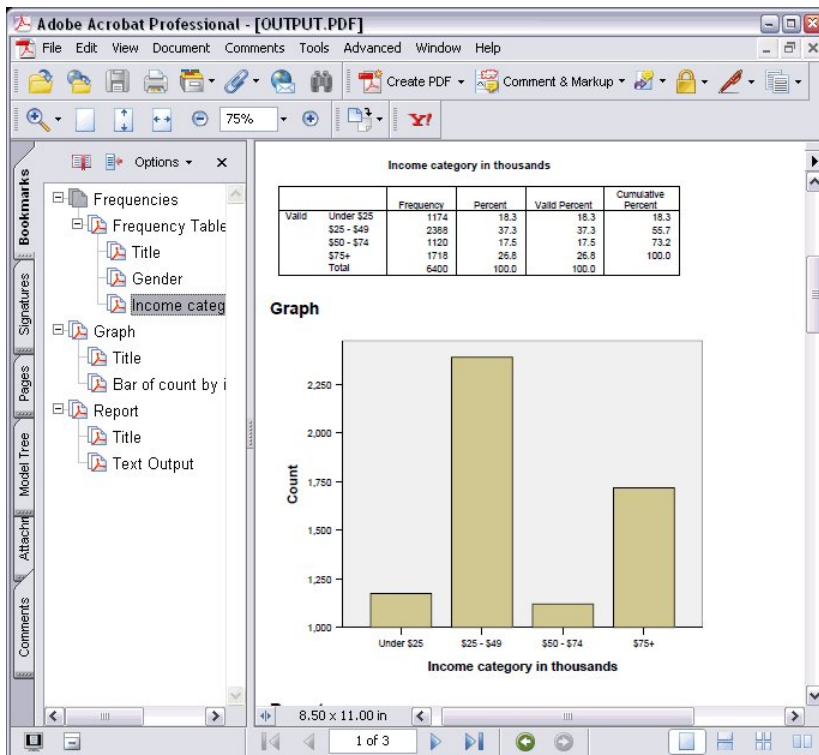


Figura 66. Archivo PDF con marcadores

Exportación de resultados a HTML

Puede exportar también los resultados a HTML (lenguaje de marcas de hipertexto). Al guardar como HTML, toda la salida no gráfica se exporta a un único archivo HTML.

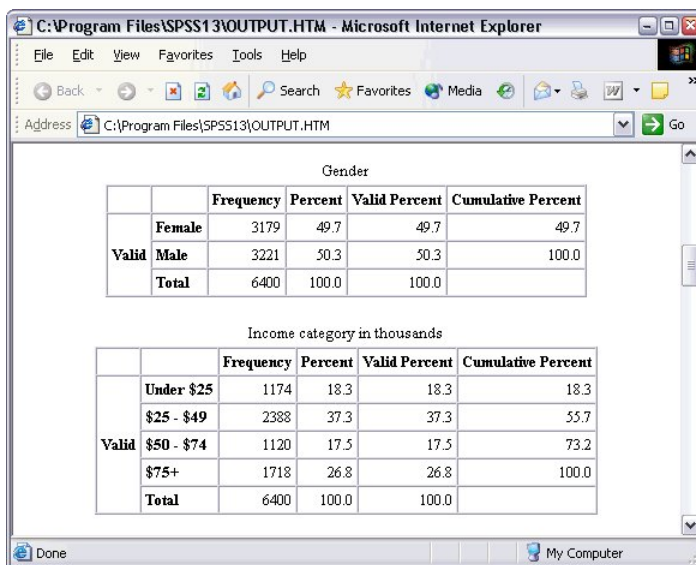


Figura 67. Archivo Output.htm en navegador Web

Al exportar a HTML, se pueden exportar también los gráficos, pero no a un solo archivo.

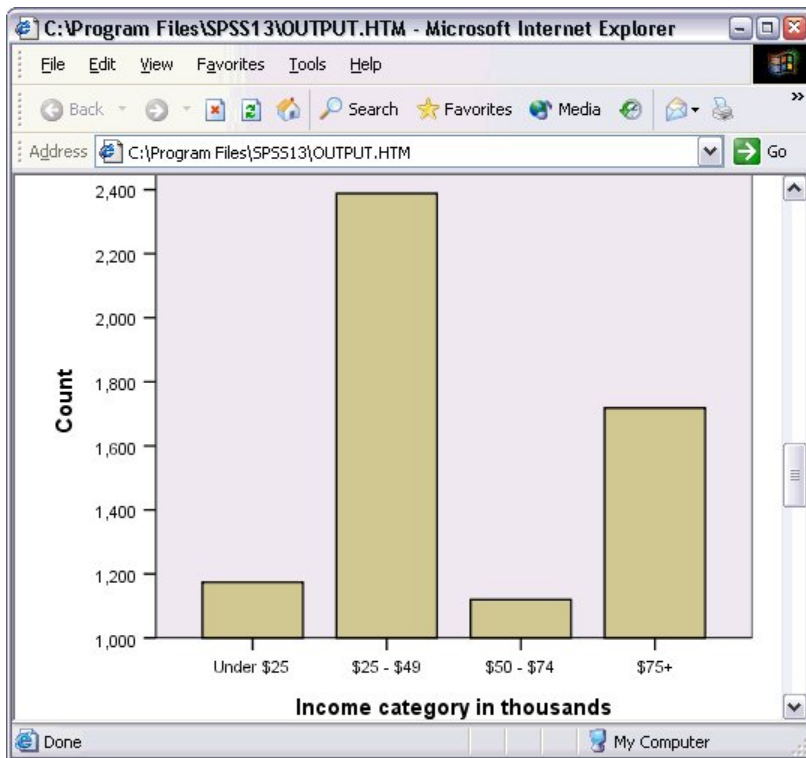


Figura 68. Gráfico en HTML

Cada gráfico se guardará como un archivo en un formato especificado por el usuario y las referencias a dichos archivos gráficos se incluirán en el documento HTML. Existe también una opción para exportar todos los gráficos (o los gráficos seleccionados) a archivos gráficos independientes.

Capítulo 7. Trabajar con sintaxis

Puede ahorrar y automatizar muchas tareas comunes mediante el eficaz lenguaje de comandos. 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 mandatos también le permite guardar los trabajos en un archivo de sintaxis para que pueda repetir el análisis en una fecha posterior.

Un archivo de sintaxis de comandos es simplemente un archivo de texto que contiene comandos de IBM SPSS Estadística. Puede abrir una ventana de sintaxis y escribir mandatos directamente, pero a menudo es más fácil dejar que los recuadros de diálogo hagan automáticamente todo o parte del trabajo.

En los ejemplos de este capítulo se utiliza el archivo de datos *demo.sav*. Consulte el tema [Capítulo 10, “Archivos de muestra”](#), en la [página 81](#) para obtener más información.

Nota: en la versión para estudiantes no está disponible la sintaxis de comandos.

Pegar sintaxis

La forma más fácil de crear sintaxis es utilizar el botón Pegar que se encuentra en la mayoría de los recuadros de diálogo.

1. Abra el archivo de datos *demo.sav*. Consulte el tema [Capítulo 10, “Archivos de muestra”](#), en la [página 81](#) para obtener más información.
2. En los menús seleccione:
Analizar > Estadísticos descriptivos > Frecuencias...
3. Seleccione *Estado civil [marital]* y muévelo a la lista Variable(s).
4. Pulse en **Gráficos**.
5. En el recuadro de diálogo Gráficos, seleccione **Gráficos de barras**.
6. En el grupo Valores del gráfico, seleccione **Porcentajes**.
7. Pulse **Continuar**. Pulse **Pegar** para copiar la sintaxis creada como resultado de las selecciones del cuadro de diálogo en el Editor de sintaxis.

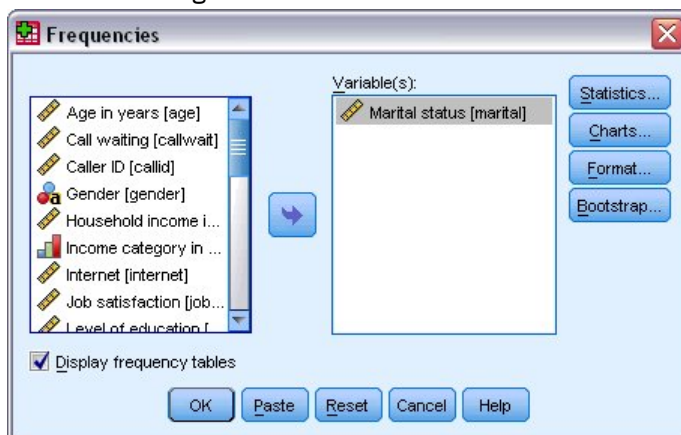


Figura 69. Cuadro de diálogo Frecuencias

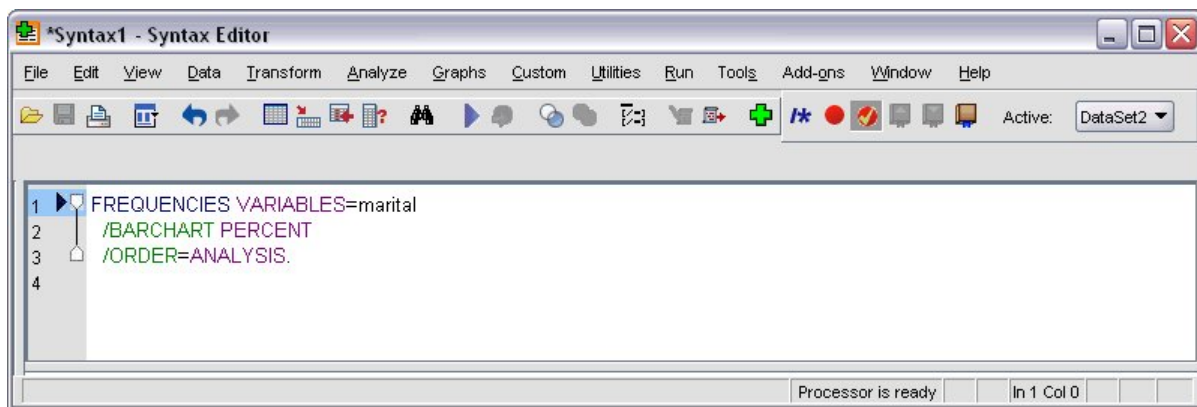


Figura 70. Sintaxis de frecuencias

8. Para ejecutar la sintaxis que se muestra actualmente, elija en los menús:

Ejecutar > Selección

Editar la sintaxis

En la ventana de sintaxis, puede editar la sintaxis. Por ejemplo, puede modificar el subcomando / BARCHART para que muestre frecuencias en lugar de porcentajes. (Cada subcomando se indica con una barra inclinada.) Si conoce la palabra clave para mostrar frecuencias, puede introducirlas directamente. Si no conoce la palabra clave, puede obtener una lista de palabras claves disponibles para el subcomando situando el cursor en cualquier parte siguiendo el nombre del subcomando y pulsando Ctrl+barra espaciadora. Muestra el control de terminación automática para el submandato.

Eliminar la palabra clave PERCENT del subcomando BARCHART.

Pulse Ctrl-Barra de espacio.

1. Pulse en el elemento etiquetado **FREC** para mostrar las frecuencias. Al pulsar en un elemento en el control autocompletar lo insertará en el lugar donde se encuentre el cursor actualmente.

De forma predeterminada, el control autocompletar mostrará una lista de términos disponibles conforme escribe. Por ejemplo, le gustaría incluir un gráfico circular en el gráfico de barras. El diagrama circular se especifica con un submandato aparte.

2. Pulse Intro después de la palabra clave **FREQ** y escriba una barra inclinada para indicar el inicio de un submandato.

El Editor de sintaxis le solicita la lista de submandatos para el mandato actual.

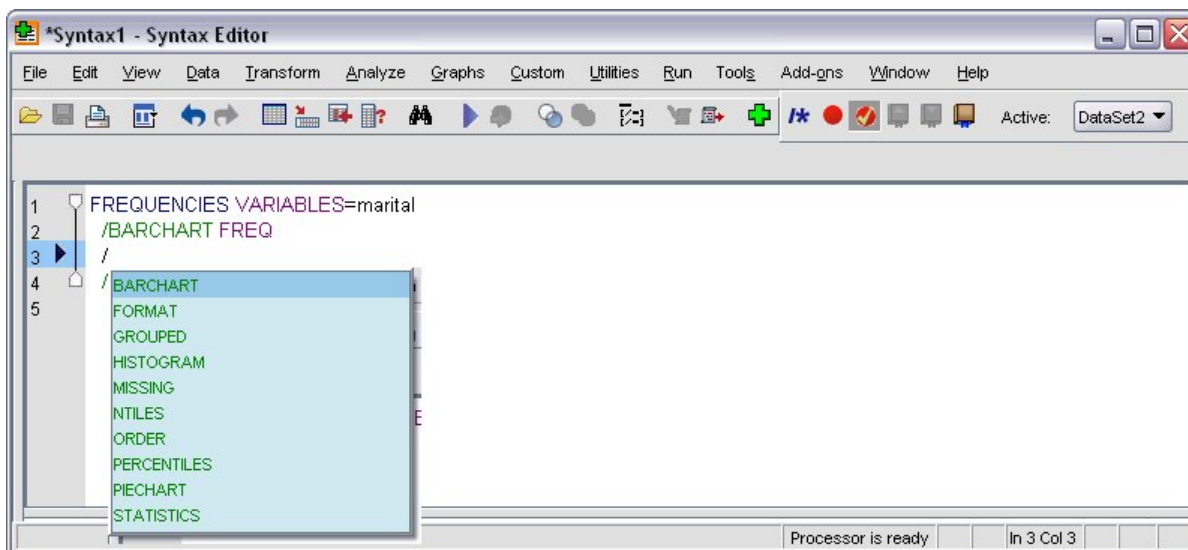


Figura 71. El control autocompletar muestra los subcomandos.

Para obtener más ayuda detallada acerca del comando actual, pulse la tecla F1. Esto le lleva directamente a la información de referencia de sintaxis del mandato para el mandato actual.

Puede que haya notado que el texto que aparece en la ventana de sintaxis tiene color. La codificación de colores le permite identificar rápidamente los términos no reconocidos, si sólo los términos reconocidos tienen color. Por ejemplo, escribes mal el subcomando FORMAT como FRMAT. Los subcomandos aparecen coloreados en verde por defecto, pero el texto FRMAT aparecerá sin colorear ya que no se reconoce.

Abrir y ejecutar un archivo de sintaxis

1. Para abrir un archivo de sintaxis almacenado, elija en los menús:

Archivo > Abrir > Sintaxis...

Se muestra un recuadro de diálogo estándar para abrir archivos.

2. Seleccione un archivo de sintaxis. Si no se visualizan archivos de sintaxis, asegúrese de que **Sintaxis (*.sps)** esté seleccionado como el tipo de archivo que desea ver.

3. Pulse en **Abrir**.

4. Utilice el menú Ejecutar en la ventana de sintaxis para ejecutar los mandatos.

Si los comandos se aplican a un archivo de datos específico, deberá abrir el archivo de datos antes de ejecutar los comandos o bien deberá incluir un comando que lo abra. Puede pegar este tipo de mandato desde los recuadros de diálogo que abren archivos de datos.

Utilizar puntos de interrupción

Los puntos de corte le permiten detener la ejecución de la sintaxis de comando en puntos específicos de la sintaxis y continuar con la ejecución cuando esté listo. Esto le permite visualizar los resultados o datos en el punto intermedio durante un trabajo de sintaxis o ejecutar la sintaxis de comandos que muestra la información sobre el estado actual de los datos, como FREQUENCIES. Los puntos de interrupción solo se pueden establecer en el nivel de un mandato, no en líneas específicas dentro de un mandato.

Para insertar un punto de interrupción en un mandato:

1. Pulse en cualquier lugar de la región que aparece a la izquierda del texto asociado al comando.

El punto de interrupción se representa como un círculo rojo en la región a la izquierda del texto del mandato y en la misma línea que el nombre del mandato, independientemente de donde haya pulsado.

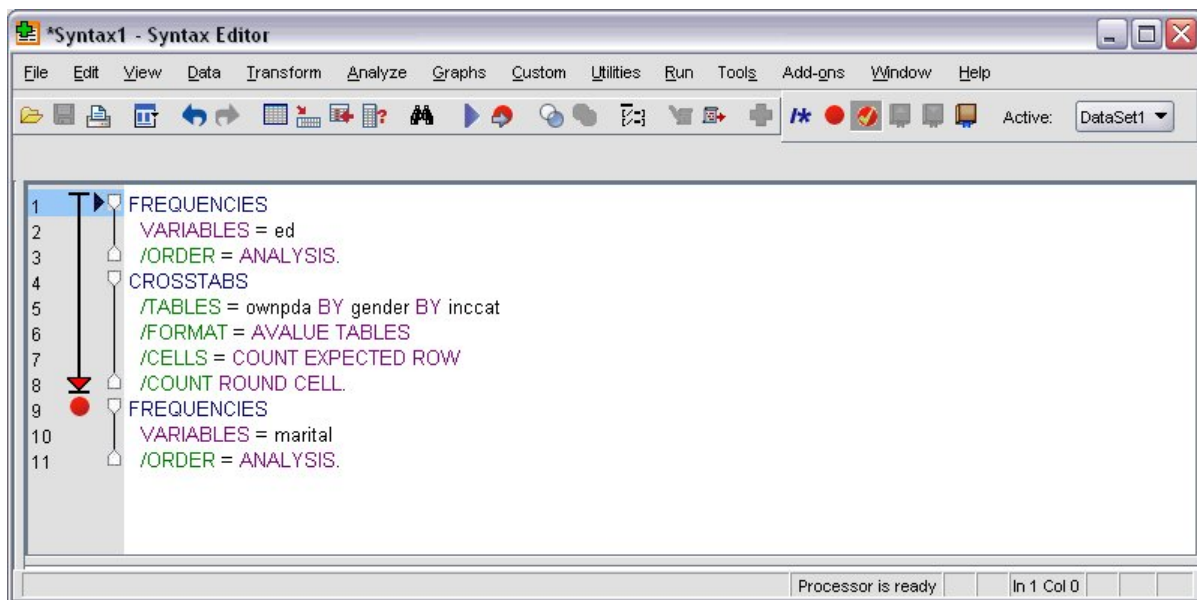


Figura 72. La ejecución se detiene en un punto de corte.

Cuando ejecuta una sintaxis de mandato que contenga puntos de interrupción, la ejecución se detiene antes de cada mandato que contenga un punto de interrupción.

La flecha que apunta hacia abajo, a la izquierda del texto de comando muestra el progreso de la ejecución de la sintaxis. Abarca la región desde el primer mandato que se ejecuta hasta la última ejecución de mandato y resulta especialmente útil cuando se ejecuta una sintaxis de mandato que contenga puntos de interrupción.

Para reanudar la ejecución después de un punto de interrupción:

2. En los menús del Editor de sintaxis, elija:

Ejecutar > Continuar

Capítulo 8. Modificar los valores de los datos

Puede que los datos con los que comience a trabajar no siempre estén organizados de la forma más conveniente para sus necesidades de análisis y de creación de informes. Por ejemplo, es posible que desee:

- Crear una variable categórica a partir de una variable de escala.
- Combinar varias categorías de respuesta en una sola categoría.
- Crear una nueva variable que sea la diferencia calculada entre dos variables existentes.
- Calcular el tiempo entre dos fechas.

En este capítulo se utiliza el archivo de datos *demo.sav*. Consulte el tema [Capítulo 10, “Archivos de muestra”](#), en la página 81 para obtener más información.

Crear una variable categórica a partir de una variable de escala

Muchas de las variables categóricas del archivo de datos *demo.sav* se derivan, de hecho, de variables de escala de ese archivo de datos. Por ejemplo, la variable *cat1ng* consiste simplemente en la variable *ingres* agrupada en cuatro categorías. Esta variable categórica emplea los valores enteros 1–4 para representar las categorías de ingresos siguientes: menos que 25\$, entre 25\$ y 49\$, entre 50\$ y 74\$ y, 75\$ o superior.

Para crear la variable categórica *inccat*:

1. Seleccione en los menús de la ventana Editor de datos:

Transformar > Agrupación visual...

En el cuadro de diálogo inicial Agrupación visual, puede seleccionar las variables de escala y ordinales para las cuales desee crear nuevas variables agrupadas. **Agrupación** significa tomar dos o más valores contiguos y agruparlos en la misma categoría.

Debido a que la agrupación visual se basa en los valores reales del archivo de datos para ayudarle a realizar buenas elecciones de agrupación, antes necesita leer el archivo de datos. Puesto que dicha operación puede llevar algún tiempo si su archivo de datos contiene un gran número de casos, este cuadro de diálogo inicial le permite también limitar el número de casos que se debe leer (que se exploran). Este proceso no es preciso con nuestro archivo de datos de muestra. A pesar de que contiene más de 6.000 casos, explorar ese número de casos no requiere mucho tiempo.

2. Arrastre y suelte *Ingresos del hogar en miles [income]* de la lista Variables a la lista Variables a agrupar y, a continuación, pulse **Continuar**.

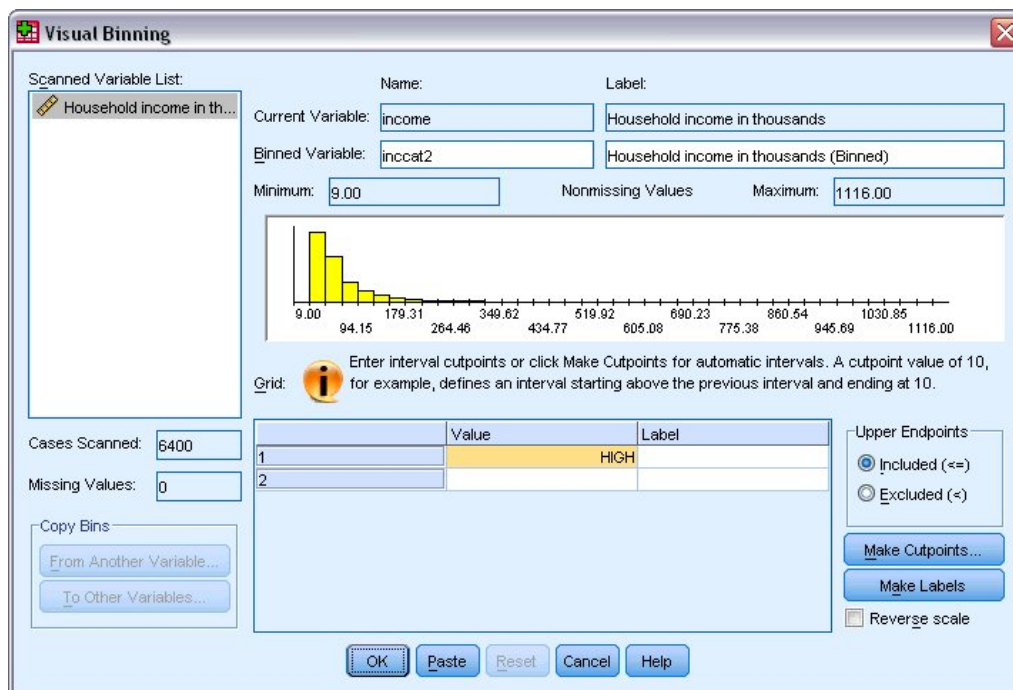


Figura 73. Cuadro de diálogo principal Agrupación visual

3. En el cuadro de diálogo principal Agrupación visual, seleccione *Ingresos del hogar en miles [ingres]* en la Lista de variables exploradas.

Un histograma muestra la distribución de la variable seleccionada (que en este caso tiene un sesgo alto).

4. Introduzca `inccat2` para el nombre de la nueva variable binned y `Income category [in thousands]` para la etiqueta de la variable.

5. Pulse **Crear puntos de corte**.

6. Seleccione **Intervalos de igual amplitud**.

7. Introduzca 25 para el primer punto de corte, 3 para el número de puntos de corte y 25 para la amplitud.

El número de categorías agrupadas es superior al número de puntos de corte. Así, en este ejemplo, la nueva variable agrupada tendrá cuatro categorías, donde cada una de las tres primeras categorías contendrá rangos de 25 (mil) y la última contendrá todos los valores por encima del valor de punto de corte más alto de 75 (mil).

8. Haga clic en **Aplicar**.

Los valores que se muestran ahora en la cuadrícula representan los puntos de corte definidos, que son los límites superiores de cada categoría. Las líneas verticales en el histograma también indican las ubicaciones de los puntos de corte.

De forma predeterminada, los valores de dichos puntos de corte se incluyen en las categorías correspondientes. Por ejemplo, el primer valor 25 incluiría entonces todos los valores menores o iguales a 25. Pero, en este ejemplo, queremos contar con categorías que correspondan a cantidades menores que 25, 25–49, 50–74 y 75 o superiores.

9. En el grupo Puntos finales superiores, seleccione **Excluidos (<)**.
10. A continuación, pulse **Crear etiquetas**.

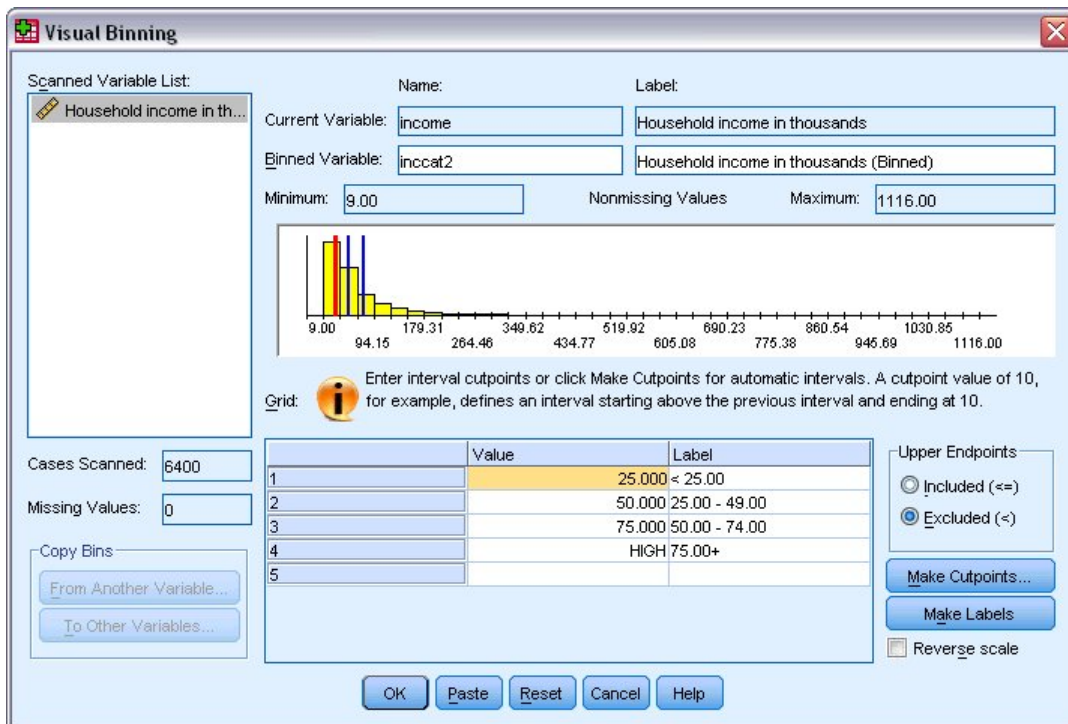


Figura 74. Etiquetas de valor generadas automáticamente.

Con ello se generan automáticamente etiquetas de valor descriptivas para cada categoría. Puesto que los valores reales asignados a la nueva variable agrupada son simplemente enteros secuenciales a partir de 1, las etiquetas de valor pueden ser muy útiles.

También puede especificar o cambiar manualmente los puntos de corte y las etiquetas de la cuadrícula, cambiar las ubicaciones de los puntos de corte arrastrando y soltando las líneas de punto de corte en el histograma y suprimir los puntos de corte arrastrando las líneas de punto de corte del histograma.

11. Pulse **Aceptar** para crear la nueva variable agrupada.

La nueva variable aparecerá en el Editor de datos. Puesto que la variable se añade al final del archivo, se visualiza en la columna del extremo derecho de la vista de datos y en la última fila de la vista de variables.

Cálculo de nuevas variables

Es posible calcular nuevas variables basadas en ecuaciones muy complejas utilizando una amplia variedad de funciones matemáticas. En este ejemplo, sin embargo, simplemente calcularemos una nueva variable que es la diferencia entre los valores de dos variables existentes.

El archivo de datos *demo.sav* contiene una variable para la edad actual de los encuestados y otra variable para el número de años transcurridos en el trabajo actual. Sin embargo, no contiene una variable para la edad del encuestado en el momento que comenzó ese trabajo. Podemos crear una nueva variable que es la diferencia calculada entre la edad actual y el número de años en el trabajo actual, que debe ser la edad aproximada en la que el encuestado inició ese trabajo.

1. Seleccione en los menús de la ventana Editor de datos:

Transformación > Calcular variable...

2. Para la Variable objetivo introduzca `jobstart`.
3. Seleccione *Edad en años [edad]* en la lista de variables de origen y pulse el botón de flecha para copiarlo en el recuadro de texto Expresión numérica.
4. Pulse en el botón del símbolo menos (-) de la calculadora en el cuadro de diálogo (o la tecla menos del teclado).

5. Seleccione *Años con el empleador actual [empleo]* y pulse el botón de flecha para copiarlo en la expresión.

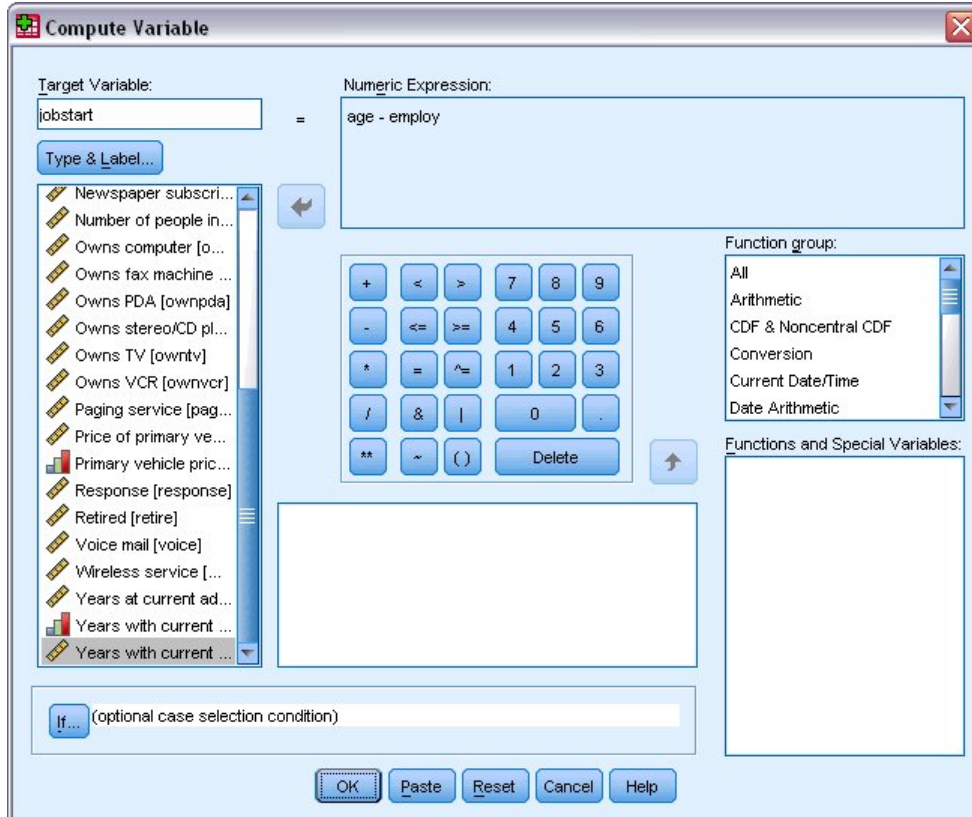


Figura 75. Cuadro de diálogo Calcular variable

Nota: esté atento al seleccionar la variable de empleo adecuada. Existe también una versión categórica recodificada de la variable, que *no* es la que usted desea. La expresión numérica debe ser *edad-empleo*, no *edad-empcat*.

6. Pulse **Aceptar** para calcular la nueva variable.

La nueva variable aparecerá en el Editor de datos. Puesto que la variable se añade al final del archivo, se visualiza en la columna del extremo derecho de la vista de datos y en la última fila de la vista de variables.

Utilizar funciones en expresiones

Puede utilizar también funciones predefinidas en las expresiones. Hay disponibles más de 70 funciones incorporadas, entre las que se incluyen:

- Funciones aritméticas
- Funciones estadísticas
- Funciones de distribución
- Funciones lógicas
- Funciones de agregación y extracción de fecha y hora
- Funciones de valor perdido
- Funciones transversales
- Funciones de cadena

Las funciones se organizan en grupos lógicos distintos, por ejemplo, un grupo para operaciones aritméticas y otro para calcular métricas estadísticas. Para mayor comodidad, también se incluyen diversas variables del sistema utilizadas habitualmente, como por ejemplo *\$TIME* (fecha y hora actual), en grupos de funciones adecuados.

Pegar una función en una expresión

Para pegar una función en una expresión:

1. Coloque el cursor en la expresión en el punto en el que desea que aparezca la función.
2. Seleccione el grupo adecuado de la lista Grupo de funciones. El grupo etiquetado como **Todos** proporciona un listado de todas las funciones y variables del sistema disponibles.
3. Efectúe una doble pulsación en la función de la lista Funciones y Variables especiales (o seleccione la función y pulse la flecha adyacente a la lista de grupos de funciones).

Se insertará la función en la expresión. Si resalta parte de la expresión y luego inserta la función, la parte resaltada de la expresión se utiliza como primer argumento en la función.

Editar una función en una expresión

La función no estará completa mientras que no se introduzcan los argumentos, que aparecen representados por interrogaciones en la función que se acaba de pegar. El número de signos de interrogación indica el número mínimo de argumentos necesarios para completar la función.

1. Resalte el signo de interrogación en la función pegada.
2. Introduzca los argumentos. Si los argumentos son nombres de variable, puede pegarlos desde la lista de variables.

Utilizar expresiones condicionales

Se pueden utilizar expresiones condicionales (también denominadas expresiones lógicas) para aplicar transformaciones a determinados subconjuntos de casos. Una expresión condicional devuelve el valor verdadero, falso o perdido para cada caso. Si el resultado de una expresión condicional es verdadero, la transformación se aplicará a ese caso. Si el resultado es falso o perdido, la transformación no se aplica al caso.

Para especificar una expresión condicional:

1. Pulse **Si** en el cuadro de diálogo Calcular variable. Esto abre el recuadro de diálogo Si los casos.
2. Seleccione **Incluir si el caso satisface la condición**.
3. Introduzca la expresión condicional.

La mayoría de las expresiones condicionales contienen al menos un operador relacional, como en:

`age>=21`

o

`income*3<100`

En el primer ejemplo, sólo aquellos casos que tengan un valor de 21 o mayor para la *Edad [edad]* seleccionada. En el segundo ejemplo, *Ingresos del hogar en miles [income]* multiplicado por 3 debe ser inferior a 100 para que se seleccione un caso.

También puede enlazar dos o más expresiones condicionales utilizando operadores lógicos, como en:

`age>=21 | ed>=4`

o

`income*3<100 & ed=5`

En el primer ejemplo, se seleccionarán los casos que cumplan la condición de *Edad [edad]* o la condición de *Nivel educativo [educ]*. En el segundo ejemplo, para que se seleccione un caso se deben cumplir ambas condiciones, *Ingresos del hogar en miles [income]* y *Nivel de educación [ed]*.

Trabajar con fechas y horas

Hay determinadas tareas que generalmente se llevan a cabo con fechas y horas que se pueden realizar fácilmente usando el Asistente para fecha y hora. Mediante este asistente, puede:

- Crear una variable de fecha/hora a partir de una variable de cadena que contenga una fecha u hora.
- Crear una variable de fecha/hora fusionando variables que contengan distintas partes de la fecha o la hora.
- Añadir o restar valores de variables de fecha/hora, incluyendo sumar o restar dos variables de fecha/hora.
- Extraer una parte de una variable de fecha u hora; por ejemplo, el día del mes de una variable de fecha/hora que tiene el formato mm/dd/aaaa.

En los ejemplos de este apartado se utiliza el archivo de datos *upgrade.sav*. Consulte el tema [Capítulo 10, “Archivos de muestra”](#), en la página 81 para obtener más información.

Para utilizar el asistente de fecha y hora:

1. En los menús seleccione:

Transformar > Asistente de fecha y hora...



Figura 76. Pantalla principal del Asistente para fecha y hora

En la pantalla principal del Asistente para fecha y hora se presenta una serie de tareas generales. Las tareas no aplicables para los datos actuales se desactivan. Por ejemplo, el archivo de datos *upgrade.sav* no contiene ninguna variable de cadena, por lo que la tarea de crear una variable de fecha a partir de una serie está inhabilitada.

Si no está familiarizado con las fechas y las horas en IBM SPSS Estadística, puede seleccionar **Aprender cómo se representan las fechas y las horas** y pulsar en **Siguiente**. Esto conduce a una pantalla que proporciona una breve visión general de las variables de fecha/hora y un enlace, a través del botón Ayuda, a información más detallada.

Cálculo de la longitud de tiempo entre dos fechas

Una de las tareas más comunes relacionadas con las fechas es el cálculo de la distancia temporal entre dos fechas. Como ejemplo, imagine que una compañía de software está interesada en analizar las compras de licencias actualizadas determinando el número de años que transcurren desde que cada cliente compró su última actualización. El archivo de datos *upgrade.sav* contiene una variable para la fecha en que cada cliente compró una actualización, pero no el número de años transcurridos desde entonces. Una nueva variable que es la longitud de tiempo en años entre la fecha de la última actualización y la fecha del siguiente release del producto proporcionará una medida de esta cantidad.

Para calcular la longitud de tiempo entre dos fechas:

1. Seleccione **Calcular con fechas y horas** en la pantalla de introducción del asistente de fecha y hora y pulse **Siguiente**.
2. Seleccione **Calcular el número de unidades de tiempo entre dos fechas** y pulse **Siguiente**.

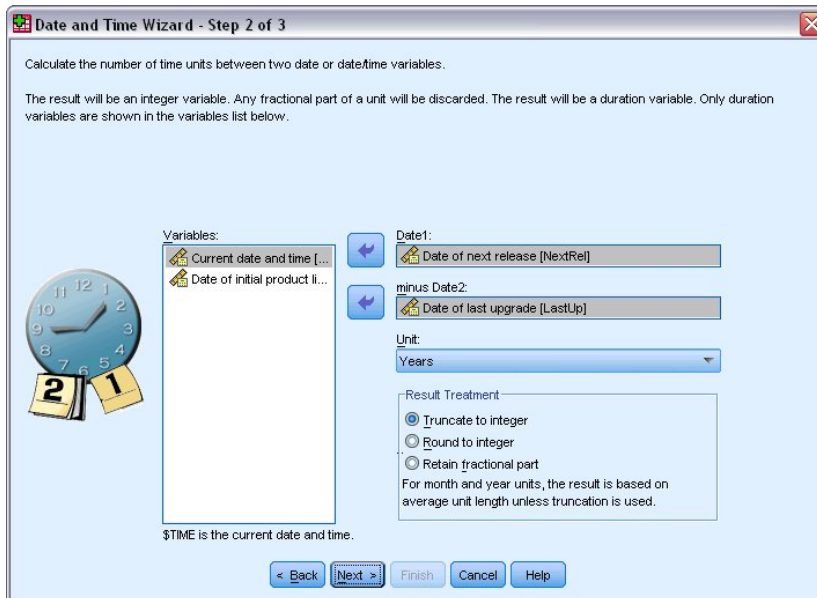


Figura 77. Cálculo de la distancia temporal entre dos fechas: paso 2

3. En el paso 2, seleccione *Fecha de la siguiente versión* en Fecha1.
4. Seleccione *Fecha de la última actualización* en Fecha2.
5. Seleccione **Años** para Unidad y **Truncar a entero** para Tratamiento resultante. (Estas son las selecciones predeterminadas).
6. Pulse **Siguiente**.
7. En paso 3, escriba *YearsLastUp* como nombre de la variable resultante. Las variables de resultado no pueden tener el mismo nombre que una variable existente.
8. Escriba *Años desde última actualización* como etiqueta de la variable resultante. Las etiquetas de variables para las variables de resultado son opcionales.
9. Deje la selección predeterminada en **Crear la variable ahora** y pulse **Finalizar** para crear la nueva variable.

La nueva variable, *AñosÚltimAct*, que se muestra en el Editor de datos, es el número entero de años transcurridos entre las dos fechas. Se han truncado partes fraccionarias de un año.

Añadir una duración a una fecha

Si lo desea, puede añadir o sustraer duraciones, como 10 días o 12 meses, a una fecha. Siguiendo con el ejemplo de la compañía de software del apartado anterior, imagine que se desea determinar la fecha en la que finaliza el contrato inicial de asistencia técnica de cada cliente. El archivo de datos *upgrade.sav* contiene una variable para el número de años de servicio de asistencia técnica contratado y otra para la fecha inicial de la compra. A continuación, puede determinar la fecha de finalización del soporte inicial añadiendo años de soporte a la fecha de compra.

Para añadir una duración a una fecha:

1. Seleccione **Calcular con fechas y horas** en la pantalla de introducción del asistente de fecha y hora y pulse **Siguiente**.
2. En el paso 1, seleccionar **Añadir o sustraer una duración a una fecha** y pulse en **Siguiente**.

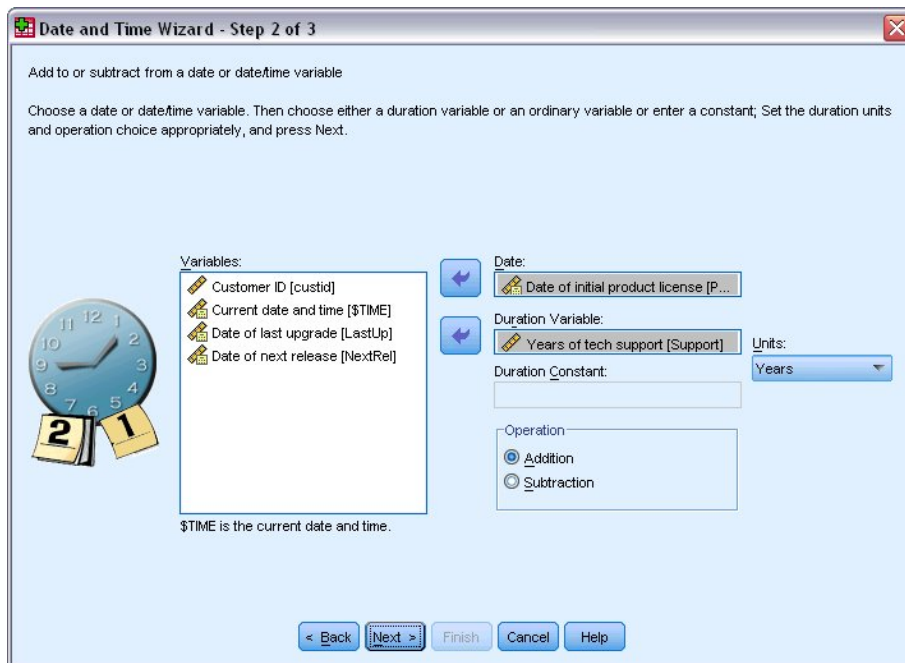


Figura 78. Adición de una duración a una fecha: paso 2

3. Seleccione *Fecha de licencia inicial del producto* para la fecha.
4. En paso 2, seleccione *Años de asistencia técnica* como la variable de duración.

Puesto que *Años de soporte técnico* es simplemente una variable numérica, debe indicar las unidades que se deben utilizar al añadir esta variable como una duración.

5. Seleccione **Años** en la lista desplegable Unidades.
6. Pulse **Siguiente**.
7. En paso 3, escriba *SupEndDate* como nombre de la variable resultante. Las variables de resultado no pueden tener el mismo nombre que una variable existente.
8. Escriba *Fecha de finalización de asistencia técnica* como etiqueta para la variable resultante. Las etiquetas de variables para las variables de resultado son opcionales.
9. Pulse **Finalizar** para crear la nueva variable.

La nueva variable aparecerá en el Editor de datos.

Capítulo 9. Ordenación y selección de datos

Los archivos de datos no siempre están organizados de la forma ideal para las necesidades específicas del usuario. Para preparar los datos para el análisis, puede seleccionar entre una amplia gama de transformaciones de archivos, incluida la posibilidad de:

- **Ordenar datos.** Puede ordenar los casos en función del valor de una o más variables.
- **Seleccionar subconjuntos de casos.** Puede restringir el análisis a un subconjunto de casos o efectuar análisis simultáneos de subconjuntos diferentes.

En los ejemplos de este capítulo se utiliza el archivo de datos *demo.sav*. Consulte el tema [Capítulo 10](#), “Archivos de muestra”, en la página 81 para obtener más información.

Ordenación de datos

La ordenación de casos (ordenación de filas del archivo de datos) es a menudo útil y a veces necesaria para determinados tipos de análisis.

Para reordenar la secuencia de casos en el archivo de datos basándose en el valor de una o más variables de ordenación:

1. En los menús seleccione:

Datos > Ordenar casos...

Se muestra el cuadro de diálogo Ordenar casos.

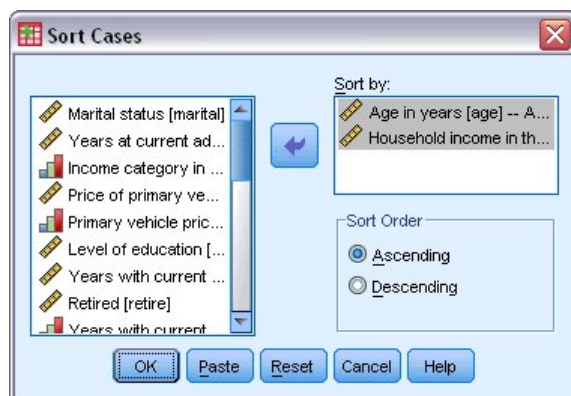


Figura 79. Cuadro de diálogo Ordenar casos

2. Añada las variables *Edad en años [edad]* e *Ingresos del hogar en miles [ingres]* a la lista Ordenar por.

Si selecciona varias variables de ordenación, los casos se ordenarán de acuerdo con el orden en que estas variables aparezcan en la lista Ordenar por. En este ejemplo, basado en las entradas de la lista Ordenar por, los casos se ordenarán por el valor de *Ingresos del hogar en miles [ingres]* dentro de las categorías de *Edad en años [edad]*. Para las variables de cadena, las letras mayúsculas preceden a las minúsculas correspondientes en orden (por ejemplo, el valor de cadena *Sí* es anterior *sí* en el orden de clasificación).

procesamiento de archivos segmentados

Para dividir el archivo de datos en grupos separados para el análisis:

1. En los menús seleccione:

Datos > Segmentar archivo...

Se muestra el cuadro de diálogo Segmentar archivo.

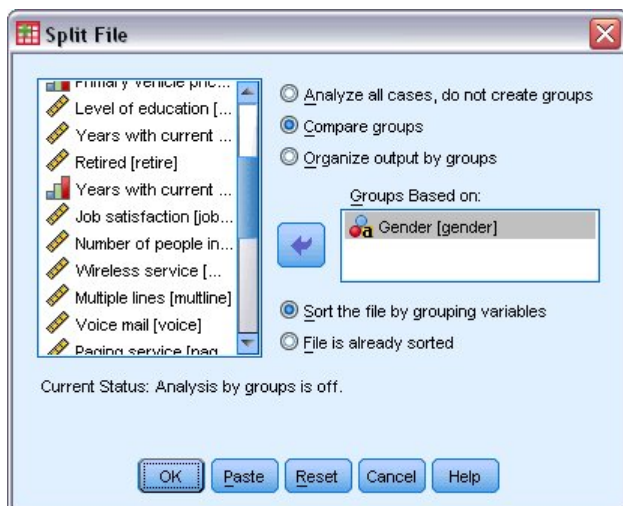


Figura 80. Cuadro de diálogo Segmentar archivo

2. Seleccione **Comparar los grupos** u **Organizar los resultados por grupos**. (Los ejemplos que siguen a estos pasos muestran las diferencias entre estas dos opciones.)
3. Seleccione *Sexo [sexo]* para dividir el archivo en grupos independientes para estas variables.

Puede utilizar variables numéricas, de cadena corta y de cadena larga como variables de agrupación. Se realizará un análisis distinto para cada uno de los subgrupos definidos por las variables de agrupación. Si selecciona varias variables de agrupación, el orden en el que aparecen en la lista Grupos basados en determina la forma en que se agrupan los casos.

Si selecciona **Comparar los grupos**, los resultados de todos los grupos de archivos segmentados se incluirán en las mismas tablas, como se muestra en la siguiente tabla de estadísticos de resumen generada por el procedimiento Frecuencias.

Statistics

Household income in thousands

Female	N	Valid	3179
		Missing	0
	Mean		68.7798
	Median		44.0000
	Std. Deviation		75.73510
Male	N	Valid	3221
		Missing	0
	Mean		70.1608
	Median		45.0000
	Std. Deviation		81.56216

Figura 81. Salida del archivo segmentado con una sola tabla dinámica

Si selecciona **Organizar los resultados por grupos** y ejecuta el procedimiento Frecuencias, se crean dos tablas dinámicas: una tabla para mujeres y una tabla para hombres. :

Statistics^a

Household income in thousands

N	Valid	3179
	Missing	0
Mean		68.7798
Median		44.0000
Std. Deviation		75.73510

a. Gender = Female

Figura 82. Salida del archivo segmentado con tabla dinámica para mujeres

Statistics^a

Household income in thousands

N	Valid	3221
	Missing	0
Mean		70.1608
Median		45.0000
Std. Deviation		81.56216

a. Gender = Male

Figura 83. Salida del archivo segmentado con tabla dinámica para hombres

Ordenación de casos para el procesamiento de archivos segmentados

El procedimiento Segmentar archivo crea un nuevo subgrupo cada vez que se encuentra un valor diferente para una de las variables de agrupación. Por lo tanto, es importante ordenar los casos basándose en los valores de las variables de agrupación antes de invocar el proceso de archivos segmentados.

De forma predeterminada, el procedimiento Segmentar archivo ordena automáticamente el archivo de datos de acuerdo con los valores de las variables de agrupación. Si el archivo ya está ordenado con el orden correcto, puede ahorrar tiempo de procesamiento si selecciona **El archivo ya está ordenado**.

Activación y desactivación del procesamiento de archivos segmentados

Después de invocar el proceso de archivos segmentados, permanece en vigor durante el resto de la sesión a menos que lo desactive.

- **Analizar todos los casos.** Esta opción desactiva el proceso de archivos segmentados.
- **Comparar grupos y Organizar los resultados por grupos.** Esta opción activa el proceso de archivos segmentados.

Mientras permanezca activado el procesamiento de archivos segmentados, aparecerá el mensaje **Segmentar archivo activado** en la barra de estado situada en la parte inferior de la ventana de la aplicación.

Selección de subconjuntos de casos

Puede limitar el análisis a un determinado subconjunto de acuerdo con criterios que pueden incluir 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

Para seleccionar un subconjunto de casos para el análisis:

1. En los menús seleccione:

Datos > Seleccionar casos...

Se abrirá el cuadro de diálogo Seleccionar casos.

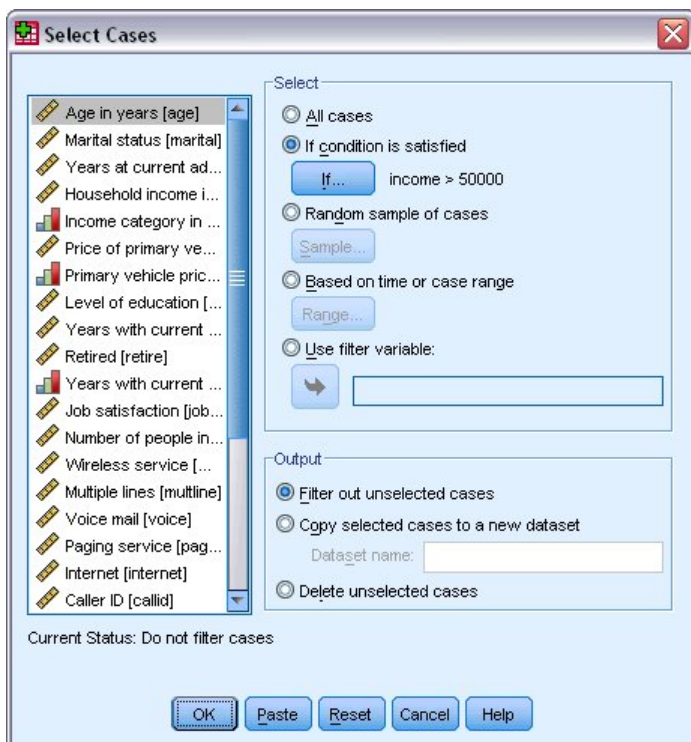


Figura 84. Cuadro de diálogo Seleccionar casos

Para seleccionar subconjuntos de casos basados en una expresión condicional

Para seleccionar casos basados en una expresión condicional:

1. Seleccione **Si se satisface la condición** y pulse en **Si** en el cuadro de diálogo Seleccionar casos. Se abrirá el cuadro de diálogo Seleccionar casos si.

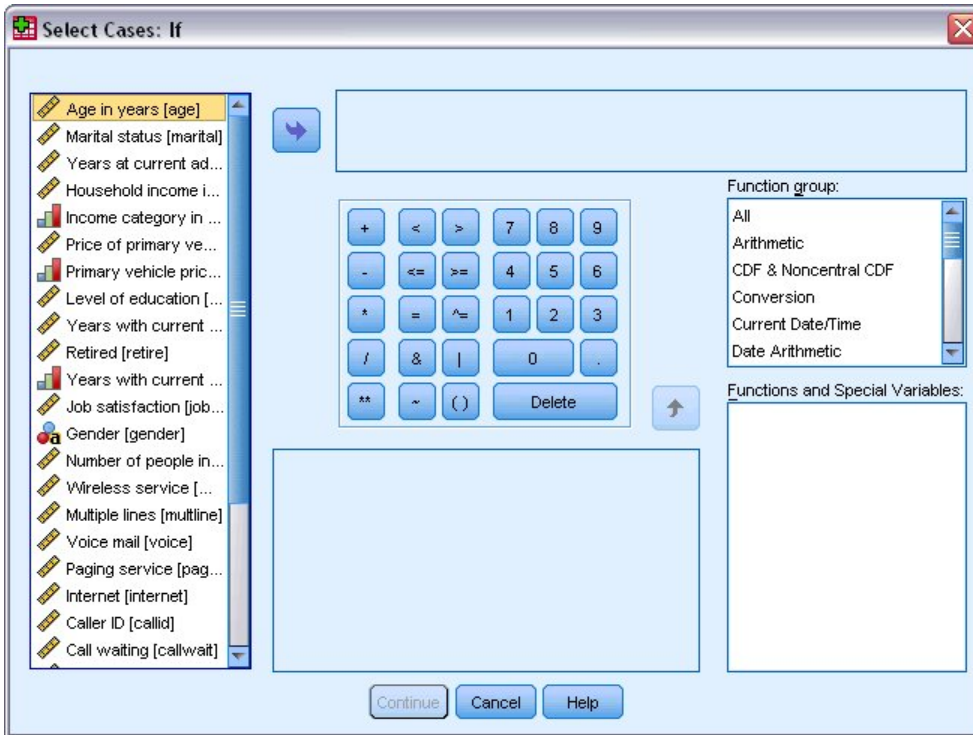


Figura 85. Cuadro de diálogo Seleccionar casos: Si

En la expresión condicional se pueden utilizar nombres de variables existentes, constantes, operadores aritméticos, operadores lógicos, operadores de relación y funciones. Puede escribir y editar la expresión en el cuadro de texto de la misma manera que cualquier otro texto de una ventana de resultados. También puede utilizar el teclado de calculadora, la lista de variables y la lista de funciones para pegar elementos en la expresión. Consulte el tema “Utilizar expresiones condicionales” en la página 69 para obtener más información.

Para seleccionar una muestra aleatoria de casos

Para obtener una muestra aleatoria:

1. Seleccione **Muestra aleatoria de casos** en el cuadro de diálogo Seleccionar casos.
2. Haga clic en **Muestra**.

Se abrirá el cuadro de diálogo Seleccionar casos de muestra aleatoria.

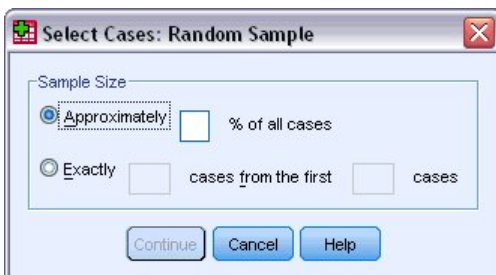


Figura 86. Cuadro de diálogo Seleccionar casos: Muestra aleatoria

Puede seleccionar una de las siguientes alternativas para el tamaño de la muestra:

- **Aproximadamente.** Un porcentaje especificado por el usuario. Esta opción genera una muestra aleatoria de aproximadamente el porcentaje de casos 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.

Selección de un rango de tiempo o un rango de casos

Para seleccionar un rango de casos basados en fechas, horas o números de observación (fila):

1. Seleccione **Basándose en el rango del tiempo o de los casos** y pulse en **Rango** en el cuadro de diálogo Seleccionar casos.

Se abrirá el cuadro de diálogo Seleccionar rango de casos, en el que puede seleccionar un rango de números de observación (fila).

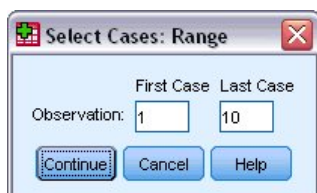


Figura 87. Cuadro de diálogo Seleccionar casos: Rango

- **Primer caso.** Introduzca los valores de la fecha y hora iniciales del rango. Si no se ha definido ninguna variable de fecha, introduzca el número de la observación inicial (número de fila del Editor de datos, a no ser que se haya activado Segmentar archivo). Si no especifica un valor de Último caso, se seleccionan todos los casos desde la fecha/hora de inicio hasta el final de la serie temporal.
- **Último caso.** Introduzca los valores de la fecha y hora finales del rango. Si no se ha definido ninguna variable de fecha, introduzca el número de la observación final (número de fila del Editor de datos, a no ser que se haya activado Segmentar archivo). Si no especifica un valor de primer caso, se seleccionan todos los casos desde el principio de la serie temporal hasta la fecha/hora final.

Para datos de serie temporal en los que se hayan definido variables de fecha, puede seleccionar un rango de fechas u horas de acuerdo con las variables de fecha definidas. Cada caso representa observaciones en un momento diferente, y el archivo se ordena por orden cronológico.



Figura 88. Cuadro de diálogo Seleccionar casos: Rango (series temporales)

Para generar variables de fecha para datos de serie temporal:

2. En los menús seleccione:

Datos > Definir fechas...

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

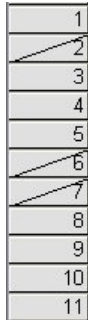
seleccionados no se incluirán en el nuevo conjunto de datos y se mantendrán en su estado original en el conjunto de datos original.

- **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.

Estado de selección de casos

Si ha seleccionado un subconjunto de casos pero no ha descartado casos no seleccionados, los casos no seleccionados se marcan en el Editor de datos con una línea diagonal a través del número de fila.



1
2
3
4
5
6
7
8
9
10
11

Figura 89. Estado de selección de casos

Capítulo 10. Archivos de muestra

Los archivos muestrales instalados con el producto se encuentran en el subdirectorio *Samples* del directorio de instalación. En el subdirectorio *Samples* hay una carpeta distinta para cada uno de los idiomas siguientes: Inglés, francés, alemán, italiano, japonés, coreano, polaco, chino simplificado, español y chino tradicional.

No todos los archivos muestrales están disponibles en todos los idiomas. Si un archivo muestral no está disponible en un idioma, esa carpeta de idioma contendrá una versión en inglés del archivo muestral.

Descripciones

A continuación, se describen brevemente los archivos muestrales usados en varios ejemplos que aparecen a lo largo de la documentación.

- **accidents.sav.** Archivo de datos hipotéticos sobre una compañía de seguros que estudia los factores de riesgo de edad y género que influyen en los accidentes de automóviles de una región determinada. Cada caso corresponde a una clasificación cruzada de categoría de edad y género.
- **adl.sav.** Archivo de datos hipotéticos relativo a los esfuerzos para determinar las ventajas de un tipo propuesto de tratamiento para pacientes que han sufrido un derrame cerebral. Los médicos dividieron de manera aleatoria a pacientes (mujeres) que habían sufrido un derrame cerebral en dos grupos. El primer grupo recibió el tratamiento físico estándar y el segundo recibió un tratamiento emocional adicional. Tres meses después de los tratamientos, se puntuaron las capacidades de cada paciente para realizar actividades cotidianas como variables ordinales.
- **advert.sav.** Archivo de datos hipotéticos sobre las iniciativas de un minorista para examinar la relación entre el dinero invertido en publicidad y las ventas resultantes. Para ello, se han recopilado cifras de ventas anteriores y los costes de publicidad asociados.
- **aflatoxin.sav.** Archivo de datos hipotéticos sobre las pruebas realizadas en las cosechas de maíz con relación a la aflatoxina, un veneno cuya concentración varía ampliamente en los rendimientos de cultivo y entre los mismos. Un procesador de grano ha recibido 16 muestras de cada uno de los 8 rendimientos de cultivo y ha medido los niveles de aflatoxinas en partes por millón (PPM).
- **anorectic.sav.** Mientras se trabaja hacia una sintomatología estandarizada del comportamiento anoréxico/bulímico, los investigadores¹hicieron un estudio de 55 adolescentes con trastornos alimenticios conocidos. Cada paciente fue examinado cuatro veces durante cuatro años, lo que representa un total de 220 observaciones. En cada observación, se puntuó a los pacientes por cada uno de los 16 síntomas. Faltan las puntuaciones de los síntomas para el paciente 71 en el tiempo 2, el paciente 76 en el tiempo 2 y el paciente 47 en el tiempo 3, lo que nos deja 217 observaciones válidas.
- **bankloan.sav.** Archivo de datos hipotéticos sobre las iniciativas de un banco para reducir la tasa de moras de créditos. El archivo contiene información financiera y demográfica de 850 clientes anteriores y posibles clientes. Los primeros 700 casos son clientes a los que anteriormente se les ha concedido un préstamo. Al menos 150 casos son posibles clientes cuyos riesgos de crédito el banco necesita clasificar como positivos o negativos.
- **bankloan_binning.sav.** Archivo de datos hipotéticos que contiene información financiera y demográfica sobre 5.000 clientes anteriores.
- **behavior.sav.** En un ejemplo clásico², se pidió a 52 estudiantes que calificaran las combinaciones de 15 situaciones y 15 comportamientos en una escala de 10 puntos que va desde 0 ="extremadamente apropiado" a 9 ="extremadamente inapropiado". Los valores promediados respecto a los individuos se toman como disimilaridades.

¹ Van der Ham, T., J. J. Meulman, D. C. Van Strien y H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry*, 170, 363-368.

² Price, R. H., y D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology*, 30, 579-586.

- **behavior_ini.sav.** Este archivo de datos contiene una configuración inicial para una solución bidimensional de *behavior.sav*.
- **brakes.sav.** Archivo de datos hipotéticos sobre el control de calidad de una fábrica que produce frenos de disco para automóviles de alto rendimiento. El archivo de datos contiene las mediciones del diámetro de 16 discos de cada una de las 8 máquinas de producción. El diámetro objetivo para los frenos es de 322 milímetros.
- **breakfast.sav.** En un estudio clásico³, a 21 estudiantes de ADE de la Wharton School y sus cónyuges se les pidió que clasificaran 15 artículos del desayuno en orden de preferencia con 1 ="más preferido" a 15 ="menos preferido". Sus preferencias se registraron en seis escenarios distintos, de "Preferencia global" a "Aperitivo, con bebida sólo".
- **breakfast-overall.sav.** Este archivo de datos sólo contiene las preferencias de elementos de desayuno para el primer escenario, "Preferencia global".
- **broadband_1.sav.** Archivo de datos hipotéticos que contiene el número de suscriptores, por región, a un servicio de banda ancha nacional. El archivo de datos contiene números de suscriptores mensuales para 85 regiones durante un período de cuatro años.
- **broadband_2.sav.** Este archivo de datos es idéntico a *broadband_1.sav* pero contiene datos para tres meses adicionales.
- **car_insurance_claims.sav.** Un conjunto de datos presentado y analizado en otro lugar⁴ está preocupado por las reclamaciones por daños a automóviles. La cantidad de reclamaciones media se puede modelar como si tuviera una distribución Gamma, mediante una función de enlace inversa para relacionar la media de la variable dependiente con una combinación lineal de la edad del asegurado, el tipo de vehículo y la antigüedad del vehículo. El número de reclamaciones presentadas se puede utilizar como ponderación de escala.
- **car_sales.sav.** Este archivo de datos contiene estimaciones de ventas, precios de lista y especificaciones físicas hipotéticas de varias marcas y modelos de vehículos. Los precios de lista y las especificaciones físicas se han obtenido de *edmunds.com* y de sitios de fabricantes.
- **car_sales_uprepared.sav.** Ésta es una versión modificada de *car_sales.sav* que no incluye ninguna versión transformada de los campos.
- **carpet.sav.** En un ejemplo popular⁵, una empresa interesada en la comercialización de un nuevo limpiador de alfombras quiere examinar la influencia de cinco factores en la preferencia del consumidor: diseño del paquete, nombre de la marca, precio, el sello de calidad de *Good Housekeeping* y una garantía de devolución de dinero. Hay tres niveles de factores para el diseño del producto, cada uno con una diferente colocación del cepillo del aplicador; tres nombres comerciales (*K2R*, *Glory* y *Bissell*); tres niveles de precios; y dos niveles (no o sí) para los dos últimos factores. Diez consumidores clasificaron 22 perfiles definidos por estos factores. La variable *Preferencia* contiene el rango de las clasificaciones medias de cada perfil. Las clasificaciones inferiores corresponden a preferencias elevadas. Esta variable refleja una medida global de la preferencia de cada perfil.
- **carpet_prefs.sav.** Este archivo de datos se basa en el mismo ejemplo que el descrito para *carpet.sav*, pero contiene las clasificaciones reales recogidas de cada uno de los 10 consumidores. Se pidió a los consumidores que clasificaran los 22 perfiles de los productos empezando por el menos preferido. Las variables desde *PREF1* hasta *PREF22* contienen los ID de los perfiles asociados, como se definen en *carpet_plan.sav*.
- **catalog.sav.** Este archivo de datos contiene cifras de ventas mensuales hipotéticas de tres productos vendidos por una compañía de venta por catálogo. También se incluyen datos para cinco variables predictoras posibles.
- **catalog_seasfac.sav.** Este archivo de datos es igual que *catalog.sav*, con la excepción de que incluye un conjunto de factores estacionales calculados a partir del procedimiento Descomposición estacional junto con las variables de fecha que lo acompañan.

³ Green, P. E., y V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.

⁴ McCullagh, P., y J. A. Nelder. 1989. *Modelos lineales generalizados*, 2.^a ed. Londres: Chapman & Hall.

⁵ Green, P. E., e Y. Viento. 1973. *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.

- **cellular.sav.** Archivo de datos hipotéticos sobre las iniciativas de una compañía de telefonía móvil para reducir el abandono de clientes. Las puntuaciones de propensión al abandono se aplican a las cuentas, que van de 0 a 100. Los usuarios de cuentas con una puntuación de 50 o más puede que ya estén buscando otro proveedor.
- **ceramics.sav.** Archivo de datos hipotéticos sobre las iniciativas de un fabricante para determinar si una nueva aleación de calidad tiene una mayor resistencia al calor que una aleación estándar. Cada caso representa una prueba independiente de una de las aleaciones; la temperatura a la que registró el fallo del rodamiento.
- **cereal.sav.** Archivo de datos hipotéticos sobre una encuesta realizada a 880 personas sobre sus preferencias en el desayuno, teniendo también en cuenta su edad, sexo, estado civil y si tienen un estilo de vida activo o no (en función de si practican ejercicio al menos dos veces a la semana). Cada caso representa un encuestado diferente.
- **clothing_defects.sav.** Archivo de datos hipotéticos sobre el proceso de control de calidad en una fábrica de prendas. Los inspectores toman una muestra de prendas de cada lote producido en la fábrica, y cuentan el número de prendas que no son aceptables.
- **coffee.sav.** Este archivo de datos corresponde a las imágenes percibidas de seis marcas de café helado⁶. Para cada uno de los 23 atributos de la imagen de café helado, la gente seleccionó todas las marcas descritas por el atributo. Las seis marcas se denotan AA, BB, CC, DD, EE y FF para mantener la confidencialidad.
- **contacts.sav.** Archivo de datos hipotéticos sobre las listas de contactos de un grupo de representantes de ventas de ordenadores de empresa. Cada uno de los contactos está categorizado por el departamento de la compañía en el que trabaja y su categoría en la compañía. Además, también se registran los importes de la última venta realizada, el tiempo transcurrido desde la última venta y el tamaño de la compañía del contacto.
- **creditpromo.sav.** Archivo de datos hipotéticos sobre las iniciativas de unos almacenes para evaluar la eficacia de una promoción de tarjetas de crédito reciente. Para este fin, se seleccionaron aleatoriamente 500 titulares. La mitad recibieron un anuncio promocionando una tasa de interés reducida sobre las ventas realizadas en los siguientes tres meses. La otra mitad recibió un anuncio estacional estándar.
- **customer_dbase.sav.** Archivo de datos hipotéticos sobre las iniciativas de una compañía para usar la información de su almacén de datos para realizar ofertas especiales a los clientes con más probabilidades de responder. Se seleccionó un subconjunto de la base de clientes aleatoriamente a quienes se ofrecieron las ofertas especiales y sus respuestas se registraron.
- **customer_information.sav.** Archivo de datos hipotéticos que contiene la información de correo del cliente, como el nombre y la dirección.
- **customer_subset.sav.** Un subconjunto de 80 casos de *customer_dbase.sav*.
- **debate.sav.** Archivos de datos hipotéticos sobre las respuestas emparejadas de una encuesta realizada a los asistentes a un debate político antes y después del debate. Cada caso corresponde a un encuestado diferente.
- **debate_aggregate.sav.** Archivo de datos hipotéticos que agrega las respuestas de *debate.sav*. Cada caso corresponde a una clasificación cruzada de preferencias antes y después del debate.
- **demo.sav.** Archivos de datos hipotéticos sobre una base de datos de clientes adquirida con el fin de enviar por correo ofertas mensuales. Se registra si el cliente respondió a la oferta, junto con información demográfica diversa.
- **demo_cs_1.sav.** Archivo de datos hipotéticos sobre el primer paso de las iniciativas de una compañía para recopilar una base de datos de información de encuestas. Cada caso corresponde a una ciudad diferente, y se registra la identificación de la ciudad, la región, la provincia y el distrito.
- **demo_cs_2.sav.** Archivo de datos hipotéticos sobre el segundo paso de las iniciativas de una compañía para recopilar una base de datos de información de encuestas. Cada caso corresponde a una unidad familiar diferente de las ciudades seleccionadas en el primer paso, y se registra la identificación de la

⁶ Kennedy, R., C. Riquier y B. Afilado. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement, and Analysis for Marketing*, 5, 56-70.

unidad, la subdivisión, la ciudad, el distrito, la provincia y la región. También se incluye la información de muestreo de las primeras dos etapas del diseño.

- **demo_cs.sav.** Archivo de datos hipotéticos que contiene información de encuestas recopilada mediante un diseño del muestreo complejo. Cada caso corresponde a una unidad familiar distinta, y se recopila información demográfica y de muestreo diversa.
- **diabetes_costs.sav.** Es un archivo de datos hipotético que contiene información mantenida por una compañía de seguros acerca de los titulares de póliza que tienen diabetes. Cada caso corresponde a un titular de política diferente.
- **dietstudy.sav.** Este archivo de datos hipotéticos contiene los resultados de un estudio de la "dieta Stillman"⁷. Cada caso corresponde a un sujeto separado y registra su peso en libras antes y después de la dieta y los niveles de triglicéridos en mg/100 ml.
- **dmdata.sav.** Es un archivo de datos hipotético que contiene información demográfica y de compra para una empresa de marketing directo. *dmdata2.sav* contiene información para un subconjunto de contactos que recibieron un envío de correos de prueba y *dmdata3.sav* contiene información sobre el resto de contactos que no recibieron el envío de correos de prueba.
- **dvdplayer.sav.** Archivo de datos hipotéticos sobre el desarrollo de un nuevo reproductor de DVD. El equipo de marketing ha recopilado datos de grupo de enfoque mediante un prototipo. Cada caso corresponde a un usuario encuestado diferente y registra información demográfica sobre los encuestados y sus respuestas a preguntas acerca del prototipo.
- **german_credit.sav.** Este archivo de datos se toma del conjunto de datos "crédito alemán" en el repositorio de bases de datos de aprendizaje automático⁸ en la Universidad de California, Irvine.
- **grocery_1month.sav.** Este archivo de datos hipotéticos es el archivo de datos *grocery_coupons.sav* con las compras semanales "acumuladas" para que cada caso corresponda a un cliente diferente. Algunas de las variables que cambiaban semanalmente desaparecen de los resultados, y la cantidad gastada registrada se convierte ahora en la suma de las cantidades gastadas durante las cuatro semanas del estudio.
- **grocery_coupons.sav.** Archivo de datos hipotéticos que contiene datos de encuestas recopilados por una cadena de tiendas de alimentación interesada en los hábitos de compra de sus clientes. Se sigue a cada cliente durante cuatro semanas, y cada caso corresponde a un cliente-semana distinto y registra información sobre dónde y cómo compran los clientes, incluida la cantidad que invierten en comestibles durante esa semana.
- **guttman.sav.** Campana⁹ se presenta en una tabla para ilustrar posibles grupos sociales. Guttman¹⁰ utilizó una parte de esta tabla, en la que cinco variables que describen aspectos como la interacción social, los sentimientos de pertenencia a un grupo, la proximidad física de los miembros y la formalidad de la relación se cruzaron con siete grupos sociales teóricos, incluyendo multitudes (por ejemplo, personas en un partido de fútbol), audiencias (por ejemplo, personas en un teatro o conferencia en un aula), público (por ejemplo, audiencias de periódicos o de televisión), multitudes (pero con una interacción mucho más intensa), grupos primarios (íntimos), grupos secundarios (voluntarios) y la comunidad moderna (confederación más amplia resultado de una proximidad física cercana y de la necesidad de servicios especializados).
- **health_funding.sav.** Archivo de datos hipotéticos que contiene datos sobre inversión en sanidad (cantidad por 100 personas), tasas de enfermedad (índice por 10.000 personas) y visitas a centros de salud (índice por 10.000 personas). Cada caso representa una ciudad diferente.
- **hivassay.sav.** Archivo de datos hipotéticos sobre las iniciativas de un laboratorio farmacéutico para desarrollar un ensayo rápido para detectar la infección por VIH. Los resultados del ensayo son ocho

⁷ Rickman, R., N. Mitchell, J. Dingman y J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association*, 228:, 54-58.

⁸ Blake, C. L., y C. J. Merz. 1998. "UCI Repository of machine learning databases." Available at <http://www.ics.uci.edu/~mllearn/MLRepository.html>.

⁹ Bell, E. H. 1961. *Social foundations of human behavior: Introduction to the study of sociology*. Nueva York: Harper & Row.

¹⁰ Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika*, 33, 469-506.

tonos de rojo con diferentes intensidades, donde los tonos más oscuros indican una mayor probabilidad de infección. Se llevó a cabo una prueba de laboratorio de 2.000 muestras de sangre, de las cuales una mitad estaba infectada con el VIH y la otra estaba limpia.

- **hourlywagedata.sav.** Archivo de datos hipotéticos sobre los salarios por horas de enfermeras de puestos de oficina y hospitales y con niveles distintos de experiencia.
- **insurance_claims.sav.** Este es un archivo de datos hipotéticos que se refiere a una compañía de seguros que quiere construir un modelo para marcar reclamaciones sospechosas, potencialmente fraudulentas. Cada caso representa una reclamación diferente.
- **insure.sav.** Archivo de datos hipotéticos sobre una compañía de seguros que estudia los factores de riesgo que indican si un cliente tendrá que hacer una reclamación a lo largo de un contrato de seguro de vida de 10 años. Cada caso del archivo de datos representa un par de contratos (de los que uno registró una reclamación y el otro no), agrupados por edad y sexo.
- **judges.sav.** Archivo de datos hipotéticos sobre las puntuaciones concedidas por jueces cualificados (y un aficionado) a 300 actuaciones gimnásticas. Cada fila representa una actuación diferente; los jueces vieron las mismas actuaciones.
- **kinship_dat.sav.** Rosenberg y Kim¹¹ se propuso analizar 15 términos de parentesco (tía, hermano, primo, hija, padre, nieta, abuelo, abuela, nieto, madre, sobrino, sobrina, hermana, hijo, tío). Le pidieron a cuatro grupos de estudiantes universitarios (dos masculinos y dos femeninos) que ordenaran estos grupos según las similitudes. A dos grupos (uno masculino y otro femenino) se les pidió que realizaran la ordenación dos veces, pero que la segunda ordenación la hicieran según criterios distintos a los de la primera. Así, se obtuvo un total de seis "orígenes". Cada origen se corresponde con una matriz de proximidades de 15 x 15, cuyas casillas son iguales al número de personas de un origen menos el número de veces que se partitionaron los objetos en ese origen.
- **kinship_ini.sav.** Este archivo de datos contiene una configuración inicial para una solución tridimensional de *kinship_dat.sav*.
- **kinship_var.sav.** Este archivo de datos contiene variables independientes *sexo*, *gener(ación)*, y *grado* (de separación) que se pueden usar para interpretar las dimensiones de una solución para *kinship_dat.sav*. Concretamente, se pueden usar para restringir el espacio de la solución a una combinación lineal de estas variables.
- **marketvalues.sav.** Este archivo de datos se refiere a las ventas de viviendas en una nueva urbanización en Algonquin, Ill., durante los años 1999-2000. Los datos de estas ventas son públicos.
- **nhis2000_subset.sav.** La National Health Interview Survey (NHIS, encuesta del Centro Nacional de Estadísticas de Salud de EE.UU.) es una encuesta detallada realizada entre la población civil de Estados Unidos. Las encuestas se realizaron en persona a una muestra representativa de las unidades familiares del país. Se recogió tanto la información demográfica como las observaciones acerca del estado y los hábitos de salud de los integrantes de cada unidad familiar. Este archivo de datos contiene un subconjunto de información de la encuesta de 2000. National Center for Health Statistics. National Health Interview Survey, 2000. Archivo de datos y documentación de uso público. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/. Fecha de acceso: 2003.
- **ozone.sav.** Los datos incluyen 330 observaciones de seis variables meteorológicas para pronosticar la concentración de ozono a partir del resto de variables. Investigadores anteriores^{12,13}, entre otras, encontraron no linealidades entre estas variables, lo que dificulta los enfoques de regresión estándar.
- **pain_medication.sav.** Este archivo de datos hipotéticos contiene los resultados de una prueba clínica sobre medicación antiinflamatoria para tratar el dolor artrítico crónico. Resulta de particular interés el tiempo que tarda el fármaco en hacer efecto y cómo se compara con una medicación existente.
- **patient_los.sav.** Este archivo de datos hipotéticos contiene los registros de tratamiento de pacientes que fueron admitidos en el hospital ante la posibilidad de sufrir un infarto de miocardio (IM o "ataque al

¹¹ Rosenberg, S., y M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, 489-502.

¹² Breiman, L., y J. H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80, 580-598.

¹³ Hastie, T., y R. Tibshirani. 1990. *Generalized additive models*. London: Chapman and Hall.

corazón"). Cada caso corresponde a un paciente distinto y registra diversas variables relacionadas con su estancia hospitalaria.

- **patlos_sample.sav.** Este archivo de datos hipotéticos contiene los registros de tratamiento de una muestra de pacientes que recibieron trombolíticos durante el tratamiento del infarto de miocardio (IM o "ataque al corazón"). Cada caso corresponde a un paciente distinto y registra diversas variables relacionadas con su estancia hospitalaria.
- **poll_cs.sav.** Archivo de datos hipotéticos sobre las iniciativas de los encuestadores para determinar el nivel de apoyo público a una ley antes de una asamblea legislativa. Los casos corresponden a votantes registrados. Cada caso registra el condado, la población y el vecindario en el que vive el votante.
- **poll_cs_sample.sav.** Este archivo de datos hipotéticos contiene una muestra de los votantes enumerados en *poll_cs.sav*. La muestra se tomó según el diseño especificado en el archivo de plan *poll_csplan* y este archivo de datos registra las probabilidades de inclusión y las ponderaciones muestrales. Sin embargo, tenga en cuenta que debido a que el plan muestral hace uso de un método de probabilidad proporcional al tamaño (PPS), también existe un archivo que contiene las probabilidades de selección conjunta (*poll_jointprob.sav*). Las variables adicionales que corresponden a los datos demográficos de los votantes y sus opiniones sobre la propuesta de ley se recopilaron y añadieron al archivo de datos después de tomar la muestra.
- **property_assess.sav.** Archivo de datos hipotéticos sobre las iniciativas de un asesor del condado para mantener actualizada la evaluación de los valores de las propiedades utilizando recursos limitados. Los casos corresponden a las propiedades vendidas en el condado el año anterior. Cada caso del archivo de datos registra la población en que se encuentra la propiedad, el último asesor que visitó la propiedad, el tiempo transcurrido desde la última evaluación, la valoración realizada en ese momento y el valor de venta de la propiedad.
- **property_assess_cs.sav.** Archivo de datos hipotéticos sobre las iniciativas de un asesor de un estado para mantener actualizada la evaluación de los valores de las propiedades utilizando recursos limitados. Los casos corresponden a propiedades del estado. Cada caso del archivo de datos registra el condado, la población y el vecindario en el que se encuentra la propiedad, el tiempo transcurrido desde la última evaluación y la valoración realizada en ese momento.
- **property_assess_cs_sample.sav.** Este archivo de datos hipotéticos contiene una muestra de las propiedades recogidas en *property_assess_cs.sav*. La muestra se tomó en función del diseño especificado en el archivo de plan *property_assess_csplan*, y este archivo de datos registra las probabilidades de inclusión y las ponderaciones muestrales. La variable adicional *Valor actual* se recopiló y añadió al archivo de datos después de tomar la muestra.
- **recidivism.sav.** Archivo de datos hipotéticos sobre las iniciativas de una agencia de orden público para comprender los índices de reincidencia en su área de jurisdicción. Cada caso corresponde a un infractor anterior y registra su información demográfica, algunos detalles de su primer delito y, a continuación, el tiempo transcurrido desde su segundo arresto, si ocurrió en los dos años posteriores al primer arresto.
- **recidivism_cs_sample.sav.** Archivo de datos hipotéticos sobre las iniciativas de una agencia de orden público para comprender los índices de reincidencia en su área de jurisdicción. Cada caso corresponde a un infractor anterior, liberado de su primer arresto durante el mes de junio de 2003, y registra información demográfica, algunos detalles de su primer delito y los datos de su segundo arresto, si se ha producido antes de finales de junio de 2006. Los acusados fueron seleccionados de un muestreo de departamentos en el plan de muestreo especificado en *recidivism_csplan*; puesto que utiliza el método de probabilidad proporcional al tamaño (PPS), también es un archivo que contiene las probabilidades de selección conjunta (*recidivism_cs_jointprob.sav*).
- **rfm_transactions.sav.** Archivo de datos hipotéticos que contiene datos de transacciones de compra, incluida la fecha de compra, los artículos adquiridos y el importe de cada transacción.
- **salesperformance.sav.** Archivo de datos hipotéticos sobre la evaluación de dos nuevos cursos de formación de ventas. Sesenta empleados, divididos en tres grupos, reciben formación estándar. Además, el grupo 2 recibe formación técnica; el grupo 3, un tutorial práctico. Cada empleado se sometió a un examen al final del curso de formación y se registró su puntuación. Cada caso del archivo de datos representa a un alumno distinto y registra el grupo al que fue asignado y la puntuación que obtuvo en el examen.

- **satisf.sav.** Archivo de datos hipotéticos sobre una encuesta de satisfacción llevada a cabo por una empresa minorista en cuatro tiendas. Se encuestó a 582 clientes en total y cada caso representa las respuestas de un único cliente.
- **screws.sav.** Este archivo de datos contiene información sobre las características de tornillos, pernos, tuercas y tachuelas¹⁴.
- **shampoo_ph.sav.** Archivo de datos hipotéticos sobre el control de calidad en una fábrica de productos para el cabello. Se midieron seis lotes de resultados distintos en intervalos regulares y se registró su pH. El intervalo objetivo es 4,5–5,5.
- **ships.sav.** Un conjunto de datos presentado y analizado en otro lugar¹⁵ que afecta a los daños en los buques de carga causados por las olas. Los recuentos de incidentes se pueden modelar como si ocurrieran con una tasa de Poisson dado el tipo de barco, el período de construcción y el período de servicio. Los meses de servicio agregados para cada casilla de la tabla formados por la clasificación cruzada de factores proporcionan valores para la exposición al riesgo.
- **site.sav.** Archivo de datos hipotéticos sobre las iniciativas de una compañía para seleccionar sitios nuevos para sus negocios en expansión. Se ha contratado a dos consultores para evaluar los sitios de forma independiente, quienes, además de un informe completo, han resumido cada sitio como una posibilidad "buena", "media" o "baja".
- **smokers.sav.** Este archivo de datos se ha extraído de la Encuesta Nacional sobre Consumo de Drogas de 1998 y es una muestra probabilística de los hogares estadounidenses. (<http://dx.doi.org/10.3886/ICPSR02934>) Por lo tanto, el primer paso en un análisis de este archivo de datos debería ser ponderar los datos para reflejar las tendencias de la población.
- **stocks.sav** Este archivo de datos hipotéticos contiene precios de acciones y volumen de un año.
- **stroke_clean.sav.** Este archivo de datos hipotéticos contiene el estado de una base de datos médica después de haberla limpiado mediante los procedimientos de la opción Preparación de datos.
- **stroke_invalid.sav.** Este archivo de datos hipotéticos contiene el estado inicial de una base de datos médica que incluye contiene varios errores de entrada de datos.
- **stroke_survival.** Este archivo de datos hipotéticos registra los tiempos de supervivencia de los pacientes que finalizan un programa de rehabilitación tras un ataque isquémico. Tras el ataque, la ocurrencia de infarto de miocardio, ataque isquémico o ataque hemorrágico se anotan junto con el momento en el que se produce el evento registrado. La muestra está truncada a la izquierda ya que únicamente incluye a los pacientes que han sobrevivido al final del programa de rehabilitación administrado tras el ataque.
- **stroke_valid.sav.** Este archivo de datos hipotéticos contiene el estado de una base de datos médica después de haber comprobado los valores mediante el procedimiento Validar datos. Sigue conteniendo casos potencialmente anómalos.
- **survey_sample.sav.** Este archivo de datos contiene datos de encuestas, incluyendo datos demográficos y diferentes medidas de actitud. Se basa en un subconjunto de variables de NORC General Social Survey de 1998, aunque algunos valores de datos se han modificado y que existen variables ficticias adicionales se han añadido para demostraciones.
- **tcm_kpi.sav.** Es un archivo de datos hipotético que contiene valores de indicadores clave de rendimiento semanales para una empresa. También contiene datos semanales para diversas métricas controlables durante el mismo periodo de tiempo.
- **tcm_kpi_upd.sav.** Este archivo de datos es idéntico a *tcm_kpi.sav* pero contiene datos para cuatro semanas adicionales.
- **telco.sav.** Archivo de datos hipotéticos sobre las iniciativas de una compañía de telecomunicaciones para reducir el abandono de clientes en su base de clientes. Cada caso corresponde a un cliente distinto y registra diversa información demográfica y de uso del servicio.
- **telco_extra.sav.** Este archivo de datos es similar al archivo de datos *telco.sav*, pero las variables de gasto del cliente de "tenencia" y convertidas en logaritmos se han eliminado y sustituido por variables de gasto de cliente convertidas en logaritmos estandarizados.

¹⁴ Hartigan, J. A. 1975. *Clustering algorithms*. New York: John Wiley and Sons.

¹⁵ McCullagh, P., y J. A. Nelder. 1989. *Modelos lineales generalizados*, 2.^a ed. Londres: Chapman & Hall.

- **telco_missing.sav.** Este archivo de datos es un subconjunto del archivo de datos *telco.sav*, pero algunos valores de datos demográficos se han sustituido con valores perdidos.
- **testmarket.sav.** Archivo de datos hipotéticos sobre los planes de una cadena de comida rápida para añadir un nuevo artículo a su menú. Hay tres campañas posibles para promocionar el nuevo producto, por lo que el artículo se presenta en ubicaciones de varios mercados seleccionados aleatoriamente. Se utiliza una promoción diferente en cada ubicación y se registran las ventas semanales del nuevo artículo durante las primeras cuatro semanas. Cada caso corresponde a una ubicación semanal diferente.
- **testmarket_1month.sav.** Este archivo de datos hipotéticos es el archivo de datos *testmarket.sav* con las ventas semanales "acumuladas" para que cada caso corresponda a una ubicación diferente. Como resultado, algunas de las variables que cambiaban semanalmente desaparecen y las ventas registradas se convierten en la suma de las ventas realizadas durante las cuatro semanas del estudio.
- **tree_car.sav.** Archivo de datos hipotéticos que contiene datos demográficos y de precios de compra de vehículos.
- **tree_credit.sav.** Archivo de datos hipotéticos que contiene datos demográficos y de historial de créditos bancarios.
- **tree_missing_data.sav** Archivo de datos hipotéticos que contiene datos demográficos y de historial de créditos bancarios con un elevado número de valores perdidos.
- **tree_score_car.sav.** Archivo de datos hipotéticos que contiene datos demográficos y de precios de compra de vehículos.
- **tree_textdata.sav.** Archivo de datos sencillos con dos variables diseñadas principalmente para mostrar el estado predeterminado de las variables antes de realizar la asignación de nivel de medición y etiquetas de valor.
- **tv-survey.sav.** Archivo de datos hipotéticos sobre una encuesta dirigida por un estudio de TV que está considerando la posibilidad de ampliar la emisión de un programa de éxito. Se preguntó a 906 encuestados si verían el programa en distintas condiciones. Cada fila representa un encuestado diferente; cada columna es una condición diferente.
- **ulcer_recurrence.sav.** Este archivo contiene información parcial de un estudio diseñado para comparar la eficacia de dos tratamientos para prevenir la reaparición de úlceras. Proporciona un buen ejemplo de datos censurados por intervalos y se ha presentado y analizado en otro lugar¹⁶.
- **ulcer_recurrence_recoded.sav.** Este archivo reorganiza la información de *ulcer_recurrence.sav* para permitir modelar la probabilidad de eventos de cada intervalo del estudio en lugar de sólo la probabilidad de eventos al final del estudio. Se ha presentado y analizado en otro lugar¹⁷.
- **verd1985.sav.** Este archivo de datos se refiere a una encuesta¹⁸. Se registraron las respuestas de 15 sujetos a 8 variables. Se han dividido las variables de interés en tres grupos. El conjunto 1 incluye *edad* y *ecivil*, el conjunto 2 incluye *mascota* y *noticia*, mientras que el conjunto 3 incluye *música* y *vivir*. Se escala *mascota* como nominal múltiple y *edad* como ordinal; el resto de variables se escalan como nominal simple.
- **virus.sav.** Archivo de datos hipotéticos sobre las iniciativas de un proveedor de servicios de Internet (ISP) para determinar los efectos de un virus en sus redes. Se ha realizado un seguimiento (aproximado) del porcentaje de tráfico de correos electrónicos infectados en sus redes a lo largo del tiempo, desde el momento en que se descubre hasta que la amenaza se contiene.
- **wheeze_steubenville.sav.** Se trata de un subconjunto de un estudio longitudinal de los efectos sobre la salud de la contaminación atmosférica en los niños¹⁹. Los datos contienen medidas binarias repetidas

¹⁶ Collett, D. 2003. *Modelización de datos de supervivencia en la investigación médica*, 2 ed. Boca Ratón: Chapman & Hall/CRC.

¹⁷ Collett, D. 2003. *Modelización de datos de supervivencia en la investigación médica*, 2 ed. Boca Ratón: Chapman & Hall/CRC.

¹⁸ Verdegaal, R. 1985. *Meer sets analyse voor kwalitatieve gegevens (in Dutch)*. Leiden: Department of Data Theory, University of Leiden.

¹⁹ Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer y B. G. Ferris Jr. 1984. Passive smoking, gas cooking, and respiratory health of children living in six cities. *American Review of Respiratory Diseases*, 129, 366-374.

del estado de sibilancias para niños de Steubenville, Ohio, con edades de 7, 8, 9 y 10 años, junto con un registro fijo de si la madre era fumadora o no durante el primer año del estudio.

- **workprog.sav.** Archivo de datos hipotéticos sobre un programa de obras del gobierno que intenta colocar a personas desfavorecidas en mejores trabajos. Se siguió una muestra de participantes potenciales del programa, algunos de los cuales se seleccionaron aleatoriamente para entrar en el programa, mientras que otros no siguieron esta selección aleatoria. Cada caso representa un participante del programa diferente.
- **worldsales.sav** Este archivo de datos hipotéticos contiene ingresos por ventas por continente y producto.

Avisos

Esta información se ha desarrollado para productos y servicios ofrecidos en Estados Unidos. Este material puede estar disponible en otros idiomas en IBM. Sin embargo, deberá poseer una copia del producto o de la versión del producto en ese idioma para poder acceder a él.

Es posible que IBM no ofrezca los productos, servicios o características que se tratan en este documento en otros países. Consulte al representante local de IBM para obtener información sobre los productos y servicios que actualmente pueden adquirirse en su zona. Las referencias a programas, productos o servicios de IBM no pretenden establecer ni implicar que sólo puedan utilizarse dichos productos, programas o servicios de IBM. En su lugar, se puede utilizar cualquier producto, programa o servicio funcionalmente equivalente que no infrinja ninguno de los derechos de propiedad intelectual de IBM. Sin embargo, es responsabilidad del usuario evaluar y comprobar el funcionamiento de todo producto, programa o servicio que no sea de IBM.

IBM puede tener patentes o solicitudes de patente en tramitación que cubran la materia descrita en este documento. La posesión de este documento no le otorga ninguna licencia sobre dichas patentes. Puede enviar preguntas acerca de las licencias, por escrito, a:

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
EE.UU.*

Para consultas sobre licencias relacionadas con información de doble byte (DBCS), póngase en contacto con el departamento de propiedad intelectual de IBM de su país o envíe sus consultas, por escrito, a:

*Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokio 103-8510, Japón*

INTERNATIONAL BUSINESS MACHINES CORPORATION PROPORCIONA ESTA PUBLICACIÓN "TAL CUAL" SIN GARANTÍAS DE NINGUNA CLASE, NI EXPLÍCITAS NI IMPLÍCITAS, QUE INCLUYEN, PERO NO SE LIMITAN A, LAS GARANTÍAS IMPLÍCITAS DE NO VULNERACIÓN, MERCANTIBILIDAD O ADECUACIÓN A UN FIN DETERMINADO. Algunas jurisdicciones no permiten la renuncia a las garantías explícitas o implícitas en ciertas transacciones; por lo tanto, es posible que esta declaración no sea aplicable en su caso.

Esta información puede incluir imprecisiones técnicas o errores tipográficos. Periódicamente se efectúan cambios en la información aquí contenida; estos cambios se incorporarán en nuevas ediciones de la publicación. IBM puede realizar mejoras y/o cambios en los productos y/o programas descritos en esta publicación en cualquier momento sin previo aviso.

Las referencias hechas en esta publicación a sitios web que no son de IBM se proporcionan sólo para la comodidad del usuario y no constituyen de modo alguno un aval de esos sitios web. Los materiales de estos sitios web no forman parte de los materiales para este producto IBM, por lo que la utilización de dichos sitios web es a cuenta y riesgo del usuario.

IBM puede utilizar o distribuir la información que se le proporcione del modo que considere adecuado sin incurrir por ello en ninguna obligación con el remitente.

Los tenedores de licencias de este programa que deseen obtener información acerca de éste con el fin de permitir: (i) el intercambio de información entre programas creados independientemente y otros programas (incluido el presente) y (ii) la utilización mutua de la información que se ha intercambiado, deben ponerse en contacto con:

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
EE.UU.*

Dicha información puede estar disponible, sujeta a los términos y condiciones correspondientes, incluyendo, en algunos casos, el pago de una tarifa.

IBM proporciona el programa bajo licencia descrito en este documento y todo el material bajo licencia disponible para el mismo según los términos del acuerdo de cliente de IBM, el acuerdo internacional de licencia de programa de IBM o cualquier otro acuerdo equivalente entre las partes.

Los ejemplos de datos de rendimiento y de clientes citados se presentan solamente a efectos ilustrativos. Los resultados reales de rendimiento pueden variar en función de las configuraciones específicas y condiciones de operación.

La información acerca de productos que no son de IBM se ha obtenido de los proveedores de dichos productos, sus publicaciones anunciadas u otras fuentes disponibles públicamente. IBM no ha probado esos productos y no puede confirmar la exactitud del rendimiento, la compatibilidad ni ninguna otra afirmación relacionada con productos no IBM. Las preguntas relacionadas con las prestaciones de los productos que no son de IBM deben dirigirse a los proveedores de esos productos.

Las declaraciones sobre el futuro rumbo o intención de IBM están sujetas a cambio o retirada sin previo aviso y representan únicamente metas y objetivos.

Esta información contiene ejemplos de datos e informes utilizados en operaciones empresariales cotidianas. Para ilustrarlos de la forma más completa posible, los ejemplos incluyen nombres de personas, empresas, marcas y productos. Todos estos nombres son ficticios y cualquier parecido con los nombres de personas o empresas reales es pura coincidencia.

LICENCIA DE DERECHOS DE AUTOR:

Esta información contiene programas de aplicación de muestra en lenguaje fuente, que ilustran técnicas de programación en distintas plataformas operativas. Puede copiar, modificar y distribuir estos programas de ejemplo en cualquier formato sin abonar ninguna cantidad a IBM, con el fin de desarrollar, utilizar, comercializar o distribuir programas de aplicaciones de acuerdo con la interfaz de programación de aplicaciones correspondiente a la plataforma operativa para la que se han escrito los programas de ejemplo. Estos ejemplos no se han probado exhaustivamente bajo todas las condiciones. Por lo tanto, IBM no puede garantizar ni dar por supuesta la fiabilidad, la capacidad de servicio ni la funcionalidad de estos programas. Los programas de ejemplo se proporcionan "TAL CUAL" sin garantías de ningún tipo. IBM no será responsable de ningún daño resultante del uso de los programas de ejemplo.

Cada copia, parcial o completa, de estos programas de ejemplo, o cualquier trabajo obtenido a partir de los mismos, debe incluir el siguiente aviso de copyright:

© Copyright IBM Corp. 2021. Las partes de este código se derivan de IBM Corp. Programas de ejemplo

© Copyright IBM Corp. 1989 - 2021. Reservados todos los derechos.

Marcas comerciales

IBM, el logotipo de IBM e ibm.com son marcas comerciales o marcas registradas de International Business Machines Corp., registradas en numerosas jurisdicciones de todo el mundo. Otros nombres de productos y servicios pueden ser marcas registradas de IBM o de otras empresas. En la web hay disponible una lista actualizada de las marcas registradas de IBM en "Copyright and trademark information" en www.ibm.com/legal/copytrade.shtml.

Adobe, el logotipo Adobe, PostScript y el logotipo PostScript son marcas registradas o marcas comerciales de Adobe Systems Incorporated en Estados Unidos y/o otros países.

Intel, el logotipo de Intel, Intel Inside, el logotipo de Intel Inside, Intel Centrino, el logotipo de Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium y Pentium son marcas comerciales o marcas registradas de Intel Corporation o sus filiales en Estados Unidos y otros países.

Linux es una marca registrada de Linus Torvalds en los Estados Unidos o en otros países.

Microsoft, Windows, Windows NT, y el logotipo de Windows son marcas comerciales de Microsoft Corporation en Estados Unidos, otros países o ambos.

UNIX es una marca registrada de The Open Group en Estados Unidos y otros países.

Java y todas las marcas registradas y logotipos basados en Java son marcas registradas de Oracle o sus filiales.

Índice

A

Access (Microsoft) [11](#)
archivos de bases de datos
lectura [11](#)
archivos de datos de texto
lectura [14](#)
archivos de Excel
lectura [8](#)
archivos de hoja de cálculo
lectura [8](#)
leer nombres de variables [8](#)
archivos de muestra
ubicación [81](#)
archivos de sintaxis
apertura [63](#)
Asistente de fecha y hora [69](#)
Asistente de importación de texto [14](#)
Asistente para bases de datos [11](#)

C

cálculo de nuevas variables [67](#)
capas
crear en tablas dinámicas [42](#)
casos
ordenación [73](#), [75](#)
selección [75](#)
crear etiquetas de variables [21](#)

D

datos categóricos
medidas de resumen [27](#)
datos continuos [27](#)
datos cualitativos [27](#)
datos cuantitativos [27](#)
datos de cadena
introducir datos [20](#)
datos de escala [27](#)
datos de intervalo [27](#)
datos de razón [27](#)
datos nominales [27](#)
datos numéricos [19](#)
datos ordinales [27](#)
desplazamiento
elementos en el Visor [39](#)
elementos en tablas dinámicas [41](#)

E

editar tablas dinámicas [42](#)
Editor de datos
introducción de datos numéricos [19](#)
introducir datos no numéricos [20](#)
entrada de datos [19](#), [20](#)

etiquetas de valor
asignación [23](#)
controlar la visualización en el Visor [23](#)
variables numéricas [23](#)
etiquetas de variable
creación [21](#)
Excel (Microsoft)
exportar resultados a [52](#)
exportación de resultados
a Excel [52](#)
a PowerPoint [52](#)
a Word [52](#)
HTML [59](#)
expresiones condicionales [69](#)

F

funciones en expresiones [68](#)

G

gráficos
barras [28](#), [33](#)
creación de gráficos [33](#)
histogramas [30](#)
gráficos de barras [28](#)

H

herramienta Ayuda de la sintaxis [62](#)
histogramas [30](#)
HTML
exportación de resultados [59](#)

I

incorporación (inversión) de filas y columnas en tablas dinámicas [41](#)
introducir datos
no numéricos [20](#)
numéricos [19](#)

M

medidas de resumen
datos categóricos [27](#)
variables de escala [29](#)

N

nivel de medición [27](#)
nueva codificación de valores [65](#)

O

ocultar filas y columnas en tablas dinámicas [43](#)
ordenación de casos [73](#)

P

pegar sintaxis
desde un recuadro de diálogo [61](#)
PowerPoint (Microsoft)
exportar resultados a [52](#)
procesamiento de archivos segmentados [73](#)

R

recuentos
tablas de recuentos [27](#)

S

selección de casos [75](#)
sintaxis [61](#)
subconjuntos de casos
basado en fechas y horas [78](#)
expresiones condicionales [76](#)
filtrado de casos no seleccionados [78](#)
muestra aleatoria [77](#)
selección [75](#)
Si se satisface la condición [76](#)
suprimir casos no seleccionados [78](#)

T

tablas de frecuencias [27](#)
tablas dinámicas
acceder a definiciones [40](#)
bandejas de pivote [41](#)
capas [42](#)
edición [42](#)
formato [42](#)
formatos de casilla [44](#)
ocultar filas y columnas [43](#)
ocultar puntos decimales [44](#)
tipos de datos de celda [44](#)
transposición de filas y columnas [41](#)
tipos de datos
para variables [22](#)

V

valores perdidos
para variables no numéricas [24](#)
para variables numéricas [24](#)
perdidos del sistema [23](#)
valores perdidos del sistema [23](#)
variables
etiquetas [21](#)
tipos de datos [22](#)
variables de escala
medidas de resumen [29](#)
variables de fecha y hora [69](#)
ventanas de sintaxis
autocompletar [62](#)

ventanas de sintaxis (*continuación*)

codificación de colores [62](#)
ejecutar mandatos [61](#), [63](#)
mandatos de edición [62](#)
mandatos de pegar [61](#)
puntos de corte [63](#)

Visor

desplazamiento de los resultados [39](#)
ocultar y mostrar la salida [39](#)

W

Word (Microsoft)

exportar resultados a [52](#)

