

Tema: Tablas con SPSS

1.- Introducción

Muchos procedimientos de SPSS generan resultados en forma de tablas. La opción Tablas de SPSS, no obstante, ofrece funciones especiales diseñadas para admitir un gran número de posibilidades de generación de informes personalizados. Muchas de las funciones personalizadas son especialmente útiles para el análisis de encuestas y estudios de mercado.

Hay que comenzar advirtiendo que el Procedimiento “Tablas” no se incluye en el sistema Base de SPSS, por lo que solo estarán disponibles si se ha adquirido dicha opción adicional.

En la versión 12 de SPSS se han remodelado los menús correspondientes a la generación de tablas proporcionando:

- Una sencilla interfaz de generación de tablas mediante el sistema de arrastrar y colocar que permite consultar una vista previa de la tabla a medida que se seleccionan variables y opciones.
- Una única interfaz unificada de generación de tablas en vez de varias opciones de menú y cuadros de diálogo para diferentes tipos de tablas.
- Una nueva sintaxis del comando CTABLES, más sencilla y fácil de entender, para sustituir a la sintaxis del comando TABLES. (Se incluye un programa de conversión que convierte trabajos anteriores de TABLES a CTABLES.)
- Subtotales para subconjuntos de categorías de una variable categórica.
- Control personalizado del orden de presentación de categorías y posibilidad de ocultar o mostrar categorías de manera selectiva.

Las tablas generadas mediante esta aplicación aparecen como tablas pivot en la ventana del Visor. Las tablas pivot ofrecen una gran flexibilidad a la hora de definir el formato y la presentación de las tablas, limitada por el nivel de medida de la variable considerada: básicamente categóricas y de escala. Más adelante se describirán pormenorizadamente las aplicaciones disponibles para cada tipo de variable.

2.- Generación de tablas

Para acceder al menú de Tablas, pulsaremos Analizar-Tablas-Tablas Personalizadas

demo.sav - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

1 : edad 55

edad ecivil dir

1	55	Casado
2	56	Soltero
3	28	Casado
4	24	Casado
5	25	Soltero
6	45	Casado
7	42	Soltero
8	35	Soltero
9	46	Soltero
10	34	Casado
11	55	Casado
12	28	Soltero

Vista de datos Vista de variables

Tablas personalizadas SPSS El procesador está preparado

Analizar ► Tablas ► Tablas personalizadas...

Comparar medias
Modelo lineal general
Modelos mixtos
Correlaciones
Regresión
Loglineal
Clasificar
Reducción de datos
Escalas
Pruebas no paramétricas
Series temporales
Supervivencia
Respuesta múltiple
Análisis de valores perdidos...

11,30 Económico Secundaria
37,20 Lujo Parte del b
19,80 Medio Parte del b
28,20 Medio Secundaria
12,20 Económico Primaria
46,10 Lujo Parte del b
35,50 Lujo Parte del b
11,80 Económico Bachillerat

En la pestaña Tabla del generador de tablas, se seleccionan las variables y las medidas de resumen que aparecerán en la tabla.

Tablas personalizadas

Tabla Títulos Estadísticos de contraste Opciones

Variables

Edad en a... Estado civ... Años en la... Ingresos f... Categoría ... Precio del ... Categoría ... Nivel ... Años ... Retira ... Años ...

Normal Compa... Capas

Columns

Tablas personalizadas

Para utilizar óptimamente este cuadro de diálogo, es necesario definir etiquetas para todas las variables categóricas y establecer correctamente sus niveles de medida.

Definir propiedades de variables puede ayudarle a realizar esto.

No volver a mostrar este mensaje

Aceptar Definir propiedades de vars. Cancelar

Definir Estadísticos de resumen Posición: Columnas Ocultar

N% Estadísticos de resumen... Categorías y totales... Origen: Por defecto

Aceptar Pegar Restablecer Cancelar Ayuda

Observamos que aparece una ventana que nos avisa de la correcta definición de variables y etiquetas, para el mejor funcionamiento del módulo.

2.1.- Definición de variables

En el panel superior izquierdo de la ventana se muestran las variables del archivo de datos. Tablas personalizadas distingue entre dos niveles de medida diferentes para las variables y trata de manera distinta las variables en función de su nivel de medida:

- Categóricas. Datos con un número limitado de valores o categorías distintas (por ejemplo, sexo o religión). Las variables categóricas pueden ser variables de cadena (alfanuméricas) o variables numéricas que utilizan códigos numéricos para representar las categorías (por ejemplo, 0=hombre y 1=mujer). También se hace referencia a estos datos como datos cualitativos.
- Escala. 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 ambos valores es 19.797€. También se hace referencia a estos datos como datos cuantitativos o continuos.

Las variables categóricas definen categorías (filas, columnas y capas) en la tabla y el estadístico de resumen por defecto es el recuento (número de casos en cada categoría). Por ejemplo, una tabla por defecto de la variable categórica sexo sólo mostrará el número de hombres y el número de mujeres. La tabla que se propone generar en el siguiente gráfico, muestra el cruce de las categorías “Ingresos en miles” y “Años en la empresa actual”:

■ Tablas personalizadas

Tabla | Títulos | Estadísticos de contraste | Opciones |

Variables

Edad en a... Estado civ... Años en la... Ingresos f... Categoría ... Precio del ... Categoría ... Nivel educ... Años en al... Retirado [r... Años en al... Categoría ...

Normal Compa... Capas

Columns

Categoría de ingresos en miles

		Menos de 2...	25\$ - 49\$	50\$ - 74\$	75\$ o +
		Recuento	Recuento	Recuento	Recuento
Años en al empresa actual	Menos de 5	nnnn	nnnn	nnnn	nnnn
	5 a 15	nnnn	nnnn	nnnn	nnnn
	Más de 15	nnnn	nnnn	nnnn	nnnn

Filas

Categorías:

Sin categorías (variable de escala)

Definir

Estadísticos de resumen

Posición: Columnas Ocult.

Origen: Variables de columna

Estadísticos de resumen

Posición de categorías:

Por defecto

Aceptar Pegar Restablecer Cancelar Ayuda

Las variables de escala se resumen normalmente dentro de las categorías de las variables categóricas y el estadístico de resumen por defecto es la media. Por ejemplo, sobre la tabla anterior, podemos incluir la edad media de los individuos:

Tablas personalizadas

Tabla | Títulos | Estadísticos de contraste | Opciones |

Variables

Normal Compa... Capas

Edad en a... Estado civ... Años en la... Ingresos f... Categoría ... Precio del ... Categoría ... Nivel educ... Años en al... Años en al... Retirado [r... Años en al...

Edad en años

Columnas

Edad en años

Categoría de ingresos en miles

	Menos de 2...	25\$ - 49\$	50\$ - 74\$	75\$ o +
Media	nnnn	nnnn	nnnn	nnnn
Años en al empresa actual	Menos de 5	nnnn	nnnn	nnnn
	5 a 15	nnnn	nnnn	nnnn
	Más de 15	nnnn	nnnn	nnnn

Estadísticos de resumen: Variables de escala

Variable seleccionada: Edad en años

Estadísticos: Recuento Recuento no ponderado Máximo Mediana Mínimo Perdido

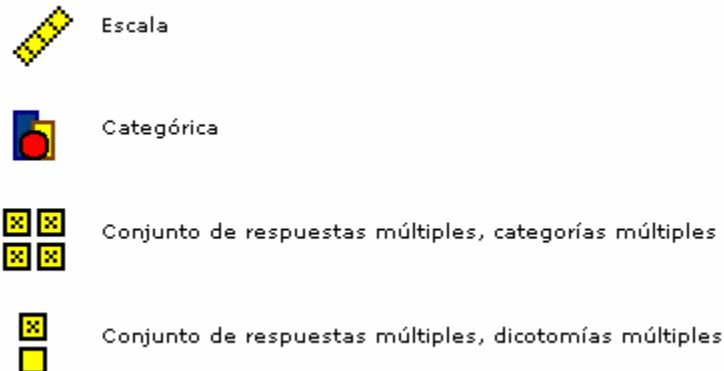
Visualización:

Estadístico	Etiqueta	Formato	Decimales
Media	Media	Automático	

Aplicar a selección Aplicar a todo Cerrar Ayuda Aceptar Pegar Restablecer Cancelar Ayuda

Obsérvese que podemos seleccionar otros estadísticos de resumen sobre la lista que proporciona el programa. También puede resumir las variables de escala por sí mismas, sin utilizar una variable categórica para definir grupos. Esto resulta especialmente útil para apilar resúmenes de varias variables de escala. Esta aplicación se describirá en el apartado correspondiente.

Un ícono junto a cada variable de la lista de variables indica el tipo el variable.



Es posible cambiar el nivel de medida de una variable en el generador de tablas, pulsando con el botón derecho del ratón en ella y seleccionando Categórica o Escala en el menú contextual (también se puede Vista de variables del Editor de datos, aunque este cambio es permanente). Hay que observar que el procedimiento Tablas

personalizadas considera como variables categóricas las variables definidas con la métrica nominal u ordinal.

Si se selecciona una variable categórica en la lista de variables, las categorías definidas para la variable se visualizarán en la lista Categorías. Estas categorías también se visualizarán en el panel de lienzo cuando utilice la variable en una tabla. Si la variable no tiene categorías definidas, la lista Categorías y el panel de lienzo mostrarán dos categorías de marcadores de posición: Categoría 1 y Categoría 2.

Las categorías definidas mostradas en el generador de tablas se basan en las etiquetas de valor, las etiquetas descriptivas asignadas a los diferentes valores de los datos (por ejemplo, los valores numéricos de 0 y 1, con las etiquetas de valor de hombre y mujer). Puede definir las etiquetas de valor en la Vista de variables del Editor de datos o con Definir propiedades de variable, en el menú Datos de la ventana Editor de datos.

2.2.- Panel de Lienzo

Una tabla se genera arrastrando variables y colocándolas en las filas y columnas del panel de lienzo. El panel de lienzo muestra una vista previa de la tabla que se va a crear, no muestra en las casillas los valores reales de los datos, pero proporciona una representación bastante exacta del diseño de la tabla final. Para las variables categóricas, la tabla real puede contener más categorías que la vista previa si el archivo de datos contiene valores particulares para los que no se han definido etiquetas de valor.

- La vista normal muestra todas las filas y columnas que se incluirán en la tabla, incluyendo las filas y/o columnas para los estadísticos de resumen y las categorías de las variables categóricas.
- La vista compacta muestra sólo las variables que habrá en la tabla, sin la representación de las filas y columnas que contendrá la tabla.

2.3.- Reglas básicas y limitaciones para generar una tabla

- Para las variables categóricas, los estadísticos de resumen se basan en la variable más al interior de la dimensión origen de los estadísticos.
- Para las variables categóricas, la dimensión origen de los estadísticos por defecto (las filas o las columnas) depende del orden en el que se arrastren y coloquen las variables en el panel de lienzo. Por ejemplo, si arrastra una variable primero al panel de filas, la dimensión de las filas es la dimensión por defecto origen de los estadísticos.
- Las variables de escala sólo se pueden resumir dentro de las categorías de la variable más al interior de la dimensión de las filas o de las columnas. (Puede colocar la variable de escala en cualquier nivel de la tabla, pero se resume en el nivel más al interior).
- Las variables de escala no se pueden resumir dentro de otras variables de escala. Puede apilar resúmenes de varias variables de escala o resumir variables de escala dentro de las categorías de las variables categóricas. No puede anidar una

variable de escala dentro de otra ni colocar una variable de escala en la dimensión de las filas y otra en la dimensión de las columnas.

La pestaña Tabla del generador de tablas ofrece un alto grado de control sobre el diseño y contenido de las tablas, incluyendo:

2.4.- Apilación y anidación de variables

La apilación se puede entender como la extracción de tablas diferentes y su unión en la misma presentación. Por ejemplo, puede mostrar la información sobre sexo y categoría de edad en secciones diferentes de la misma tabla.

Para apilar variables, seleccionaremos todas las variables que deseé apilar, y las arrastraremos todas juntas, colocándolas en las filas o columnas del panel de lienzo. También podemos arrastrar y colocar las variables por separado, colocando cada una por encima o por debajo de las variables existentes en las filas o a la derecha o la izquierda de las variables existentes en las columnas.

Por su parte, la anidación, al igual que las tablas de contingencia, puede mostrar la relación existente entre dos variables categóricas, con la diferencia de que una variable se anida dentro de la otra en la misma dimensión. Por ejemplo, podemos anidar la variable “Sexo” dentro de la categoría de “Años en la empresa actual” en la dimensión de las filas, mostrando el número de hombres y mujeres que pertenecen a cada una de las categorías de la antigüedad en la empresa:

También puede anidar una variable de escala dentro de una variable categórica. Por ejemplo, puede anidar la variable de Ingresos dentro de Sexo, mostrando por separado, para hombres y mujeres, los valores de los ingresos medios (o la mediana u otra medida de resumen).

Para anidar variables, arrastraremos una primera variable categórica al área de filas o columnas del panel de lienzo. Seguidamente arrastramos una variable categórica o de escala y la colocamos a la derecha o la izquierda de la variable de fila categórica o por encima o por debajo de la variable de columna categórica.

2.5.- Capas

Puede utilizar capas para añadir una dimensión de profundidad a las tablas, creando "cubos" tridimensionales. Las capas son similares a la anidación o la apilación. La diferencia fundamental es que sólo resulta visible simultáneamente una categoría de capa. Por ejemplo, si utiliza Categoría de edad como variable de fila y Sexo como variable de capas se generará una tabla en la que la información correspondiente a hombres y mujeres aparecerá en diferentes capas de la tabla.

Para crear capas pulsamos en Capas en la pestaña Tabla del generador de tablas para visualizar la lista Capas. Arrastramos y colocamos las variables de escala o categóricas que definirán las capas en la lista Capas. También podemos arrastrar y soltar variables en el botón Capas sin ver la lista Capas.

Es interesante notar que en la lista Capas se no puede mezclar variables de escala y categóricas, esto es, todas las variables deben ser del mismo tipo. Además en la lista Capas, los conjuntos de respuestas múltiples se tratarán como categóricas, como veremos posteriormente. Las variables de escala de las capas siempre están apiladas.

En el siguiente gráfico se muestra un ejemplo en el que se han incluido las categorías "Sexo" y "Categoría de precio del vehículo principal":

The screenshot shows the 'Tablas personalizadas' (Table Personalization) dialog box in SPSS. The 'Capas' (Layers) tab is selected. On the left, the 'Variables' (Variables) list includes 'Edad en a...', 'Estado civi...', 'Años en la...', 'Ingresos f...', 'Categoría ...', 'Precio del ...', 'Categoría ...', 'Nivel educ...', 'Años en al...', 'Retirado [r...', and 'Años en al...'. The 'Filas' (Rows) list is empty. The 'Columns' (Columns) list includes 'Categoría de ingresos en miles' with categories 'Menos de 2...', '25\$ - 49\$', '50\$ - 74\$', and '75\$ o+'. The 'Capas' (Layers) list contains 'Género' and 'Catego...'. The 'Resultado de capa' (Layer Result) section has two radio buttons: 'Mostrar cada categoría como una capa' (selected) and 'Mostrar cada combinación de categorías como una capa'. The 'Estadísticos de resumen' (Summary Statistics) section shows 'Estadísticos de resumen' selected, with 'Posición: Columnas' and 'Origen: Variables de columna'. The 'Posición de categorías' (Category Position) section shows 'Por defecto'. At the bottom are buttons for 'Aceptar' (Accept), 'Pegar' (Paste), 'Restablecer' (Reset), 'Cancelar' (Cancel), and 'Ayuda' (Help).

Obsérvese que al haber 2 variables categóricas de capas, el programa da la opción de que se puedan apilar o anidar (opción "Resultado de capa"). Por tanto, se podrá:

- Mostrar cada categoría como una capa (equivalente a la apilación). Por tanto, aparecerá una capa distinta para cada categoría de la variable de capas: el número total de capas es simplemente la suma del número de categorías para cada variable de capas. Por ejemplo, si hay tres variables de capas, cada una con tres categorías, la tabla tendrá nueve capas.
- Mostrar cada combinación de categorías como una capa (equivalente a anidar o cruzar las capas). El número total de capas es el producto del número de categorías para cada variable de capas. Por ejemplo, si hay tres variables, cada una con tres categorías, la tabla tendrá 27 capas.

2.6.- Visualización y ocultación de nombres de variable y/o etiquetas

Las siguientes opciones están disponibles para la presentación de nombres de variable y etiquetas:

- Mostrar sólo las etiquetas de variable. Para las variables que no tengan etiquetas de variable definidas se mostrará el nombre de la variable. Ésta es la configuración por defecto.
- Mostrar sólo las etiquetas de variable.
- Mostrar etiquetas y nombres de variable.
- No mostrar etiquetas ni nombres de variable. Aunque la columna/fila que contiene la etiqueta o el nombre de la variable seguirá visualizándose en la vista previa de la tabla del panel de lienzo, esta columna/fila no se visualizará en la tabla real.

Para mostrar u ocultar etiquetas o nombres de variable pulsaremos con el botón derecho del ratón en la variable de la vista previa de la tabla del panel de lienzo. A continuación seleccionamos Mostrar etiqueta de variable o Mostrar nombre de variable en el menú contextual emergente para activar o desactivar la presentación de etiquetas o nombres. Una marca de verificación junto a la selección indica que será visualizada.

3.- Estadísticos de resumen

El cuadro de diálogo Estadísticos de resumen le permite:

- Añadir y eliminar estadísticos de resumen de una tabla.
- Cambiar las etiquetas para los estadísticos.
- Cambiar el orden de los estadísticos.
- Cambiar el formato de los estadísticos, incluyendo el número de posiciones decimales.

Los estadísticos de resumen (y las demás opciones) disponibles aquí dependen del nivel de medida de la variable origen de los estadísticos de resumen. El origen de los estadísticos de resumen (la variable en la que se basan los estadísticos de resumen) está determinado por:

- Nivel de medida. Si una tabla (o una sección de una tabla en una tabla apilada) contiene una variable de escala, los estadísticos de resumen se basarán en la variable de escala.
- Orden de selección de las variables. La dimensión origen de los estadísticos por defecto (fila o columna) para las variables categóricas depende del orden en el que se arrastren y coloquen las variables en el panel de lienzo. Por ejemplo, si arrastramos una variable primero al área Filas, la dimensión de las filas es, por defecto, la dimensión origen de los estadísticos.
- Anidar. Para las variables categóricas, los estadísticos de resumen se basan en la variable más al interior de la dimensión origen de los estadísticos.

Una tabla apilada puede tener varias variables origen de los estadísticos de resumen (tanto de escala como categóricas), pero cada sección de la tabla sólo puede tener un origen de los estadísticos de resumen.

3.1.- Estadísticos de resumen para variables categóricas

Los estadísticos básicos disponibles para las variables categóricas son los recuentos y porcentajes. También se pueden especificar estadísticos de resumen personalizados para los totales y subtotales. Estos estadísticos de resumen personalizados incluyen medidas de tendencia central (como la media y la mediana) y dispersión (como la desviación típica) que pueden ser adecuadas para algunas variables categóricas ordinales.

- Recuento. Número de casos incluido en cada casilla de la tabla o número de respuestas para conjuntos de respuestas múltiples.
- Recuento no ponderado. El número de casos no ponderados incluido en cada casilla de la tabla. Sólo varía respecto al recuento si está activado ponderación.
- Porcentajes de columna. Porcentajes incluidos dentro de cada columna. La suma de los porcentajes de cada columna de una subtabla (para los porcentajes simples) da como resultado 100%. Los porcentajes de columna suelen ser de gran utilidad sólo si tiene una variable categórica de filas.
- Porcentajes de fila. Porcentajes incluidos dentro de cada fila. La suma de los porcentajes de cada fila de una subtabla (para los porcentajes simples) da como resultado 100%. Los porcentajes de fila suelen ser de gran utilidad sólo si tiene una variable categórica de columnas.
- Porcentajes de fila de capa y de columna de capa. La suma de los porcentajes de fila o columna (para los porcentajes simples) da como resultado 100% en todas las subtablas de una tabla anidada. Si la tabla contiene capas, la suma de los porcentajes de fila o columna da como resultado 100% en todas las subtablas anidadas de cada capa.
- Porcentajes de capa. Porcentajes incluidos dentro de cada capa. Para los porcentajes simples, la suma de los porcentajes de casilla incluidos dentro de la capa que está visible da como resultado 100%. Si no tiene ninguna variable de capas, esto equivale a los porcentajes de tabla.
- Porcentajes de tabla. Los porcentajes de cada casilla se basan en la tabla entera. Todos los porcentajes de casilla se basan en el mismo número total de casos y su suma da como resultado 100% (para los porcentajes simples) en la tabla entera.
- Porcentajes de subtabla. Los porcentajes de cada casilla se basan en la subtabla. Todos los porcentajes de casilla de la subtabla se basan en el mismo número

total de casos y su suma da como resultado 100% en la subtabla. En las tablas anidadas, la variable que precede al nivel de anidación más al interior define las subtablas. Por ejemplo, en una tabla de Estado civil dentro de Sexo y dentro de Categoría de edad, Sexo define las subtablas.

Respecto de los porcentajes comentados, merece la pena advertir que en el caso de variables apiladas, para los cálculos de porcentajes, cada sección de tabla definida por una variable de apilación se tratará como una tabla independiente. La suma de los porcentajes de fila de capas, columna de capas y tabla dan como resultado 100% (para los porcentajes simples) dentro de cada sección de tabla apilada. La base de los porcentajes, para los diferentes cálculos de los porcentajes, se basa en los casos de cada sección de tabla apilada.

Además, los porcentajes se pueden calcular de tres maneras diferentes, determinadas por el tratamiento de los valores perdidos en la base computacional:

- Porcentaje simple. Los porcentajes se basan en el número de casos que se utilizan en la tabla y su suma siempre da como resultado 100%. Si se excluye una categoría de la tabla, se excluirán de la base los casos de esa categoría. Los casos con valores perdidos del sistema se excluyen siempre de la base. Se excluirán los casos con valores perdidos definidos por el usuario si las categorías perdidas definidas por el usuario se excluyen de la tabla (el valor por defecto) o se incluirán si las categorías perdidas definidas por el usuario se incluyen en la tabla. Los porcentajes que no tengan "N válido" o "N total" en su nombre son porcentajes simples.
- Porcentaje del N total. Los casos con valores perdidos del sistema y valores perdidos definidos por el usuario se añaden a la base simple de los porcentajes. La suma de los porcentajes **puede ser inferior** a 100%.
- Porcentaje del N válido. Los casos con valores perdidos definidos por el usuario se eliminan de la base simple de los porcentajes incluso si las categorías perdidas definidas por el usuario se incluyen en la tabla.

Finalmente, los casos de las categorías excluidas manualmente distintas de las perdidas definidas por el usuario se excluyen siempre de la base.

3.2.- Estadísticos de resumen para variables de escala y totales categóricos personalizados

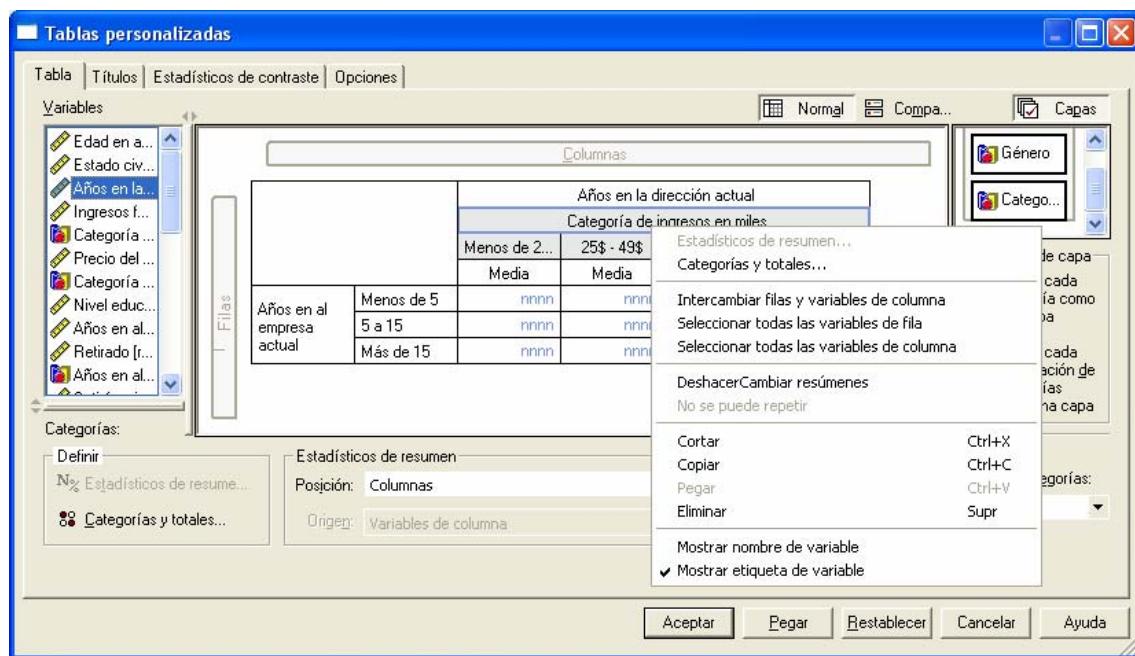
Además de los recuentos y porcentajes disponibles para las variables categóricas, están también disponibles los siguientes estadísticos de resumen para las variables de escala y como resúmenes del total y el subtotal personalizados para las variables categóricas. Estos estadísticos de resumen no están disponibles para los conjuntos de respuestas múltiples o las variables de cadena (alfanuméricas).

- Media. Media aritmética; la suma dividida por el número de casos.
- Mediana. Es el valor por encima y por debajo del cual se encuentran la mitad de los casos, el percentil 50.
- Moda. El valor más frecuente. Si hay un empate, se muestra el valor menor.
- Mínimo. Es el valor menor (más pequeño).

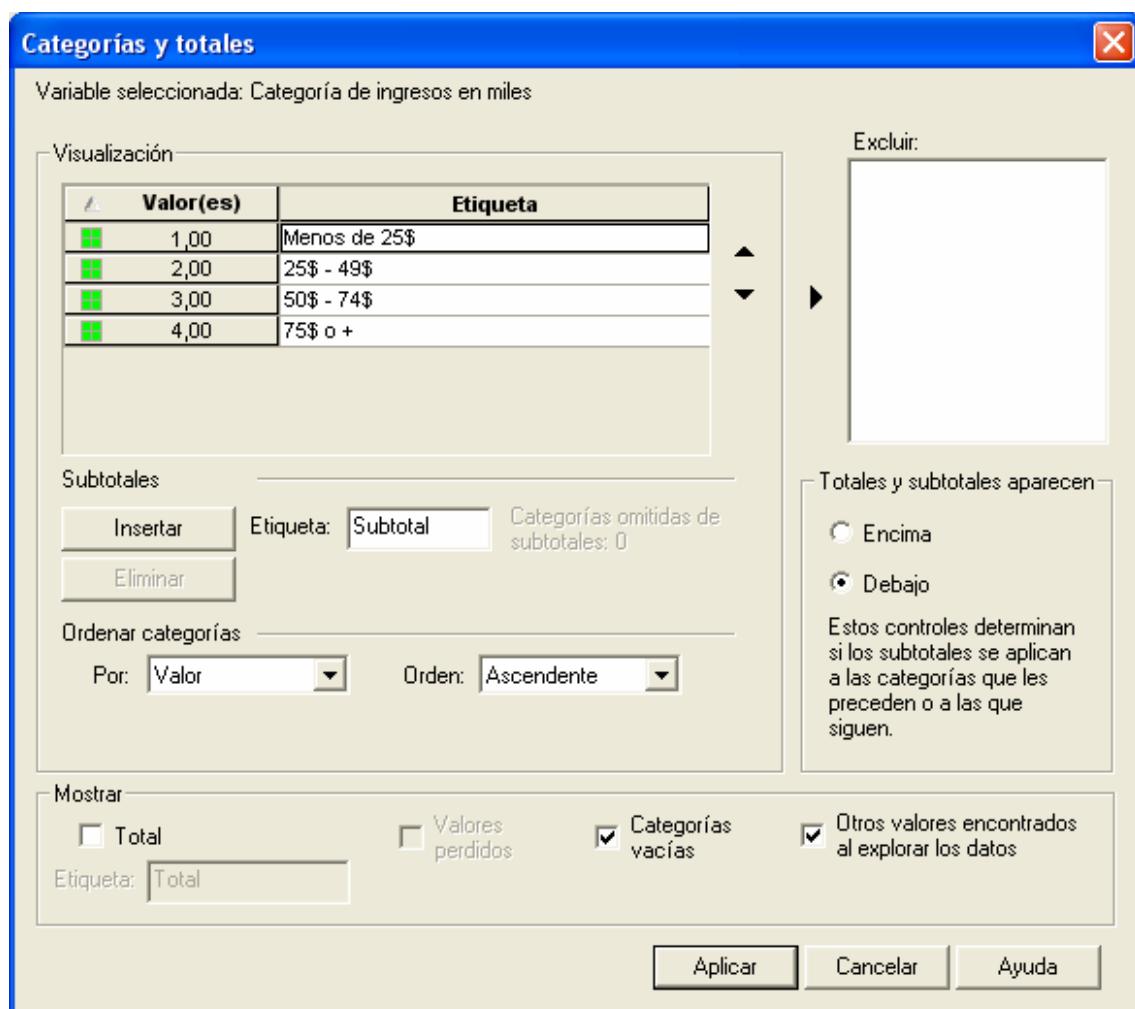
- Máximo. El valor mayor (más grande).
- Perdidos. Es el recuento de los valores perdidos (definidos por el usuario y del sistema, ambos).
- Percentil. Puede incluir los percentiles 5, 25, 75, 95 y/o 99.
- Rango. Es la diferencia existente entre los valores mínimo y máximo.
- Error típico de la media. Es una medida de cuánto puede variar el valor de la media entre varias muestras tomadas de la misma distribución.
- Desviación típica. Es una medida de la dispersión en torno a la media.
- Suma. La suma de los valores.
- Porcentaje de la suma. Porcentajes basados en sumas. Está disponible para las filas y columnas (dentro de las subtablas), filas y columnas enteras (de todas las subtablas), capas, subtablas y tablas enteras.
- N total. Es el recuento de los valores no perdidos, perdidos definidos por el usuario y perdidos del sistema. No incluye los casos correspondientes a las categorías excluidas manualmente, que sean distintas de perdidas definidas por el usuario.
- N válido. Es el recuento de los valores no perdidos. No incluye los casos correspondientes a las categorías excluidas manualmente, que sean distintas de las perdidas definidas por el usuario.
- Varianza (realmente cuasi-varianza). Es una medida de dispersión en torno a la media, igual a la suma de las desviaciones al cuadrado respecto a la media, dividida por el número de casos menos 1.

En el caso apilado, como ocurría para los resúmenes para categorías, cada sección se trata como una tabla independiente, por lo que los estadísticos de resumen se calculan para cada sección en consecuencia.

Para las tablas de variables categóricas que contienen totales o subtotales, es posible incluir estadísticos de resumen diferentes de los resúmenes mostrados para cada categoría. Por ejemplo, se pueden mostrar recuentos y porcentajes de columna para una variable de fila categórica ordinal y mostrar la mediana para el estadístico "total". Para ello, seleccionamos la variable categórica y pulsamos con el botón izquierdo del ratón. Aparecerá el siguiente menú contextual:



Pulsando ahora en Categorías y Totales, aparece el cuadro de diálogo siguiente:



en la que podemos modificar las categorías que representarán los estadísticos de resumen y otras opciones. Entre otras, podremos:

- Reordenar y excluir categorías.
- Insertar totales y subtotales
- Incluir o excluir categorías vacías.
- Incluir o excluir categorías que contienen valores perdidos.
- Incluir o excluir categorías que no tienen etiquetas de valor definidas.

Este cuadro de diálogo sólo está disponible para las variables categóricas y los conjuntos de respuestas múltiples (no está disponible, por tanto, para las variables de escala). Si se seleccionan varias variables con distintas categorías, no se pueden insertar subtotales, excluir categorías o reordenar las categorías manualmente. Esto sucede si seleccionamos varias variables en la vista previa del panel de lienzo y accedemos a este cuadro de diálogo para todas las variables seleccionadas de forma simultánea. No obstante, puede realizar estas acciones para cada variable por separado.

Para las variables sin ninguna etiqueta de valor definida, sólo se pueden ordenar categorías e insertar totales.

Los totales y subtotales se pueden mostrar por encima o por debajo de las categorías incluidas en cada total.

- Si está seleccionada la opción Debajo en el grupo Totales y subtotales aparecen, los totales aparecen por encima de cada subtabla y todas las categorías situadas por encima, incluyendo la categoría seleccionada (pero por debajo de cualquier subtotal precedente), se incluyen en todos los subtotales.
- Si está seleccionada la opción Encima en el grupo Totales y subtotales aparecen, los totales aparecen por debajo de cada subtabla y todas las categorías situadas por debajo, incluyendo la categoría seleccionada (pero por encima de cualquier subtotal precedente), se incluyen en cada subtotal.

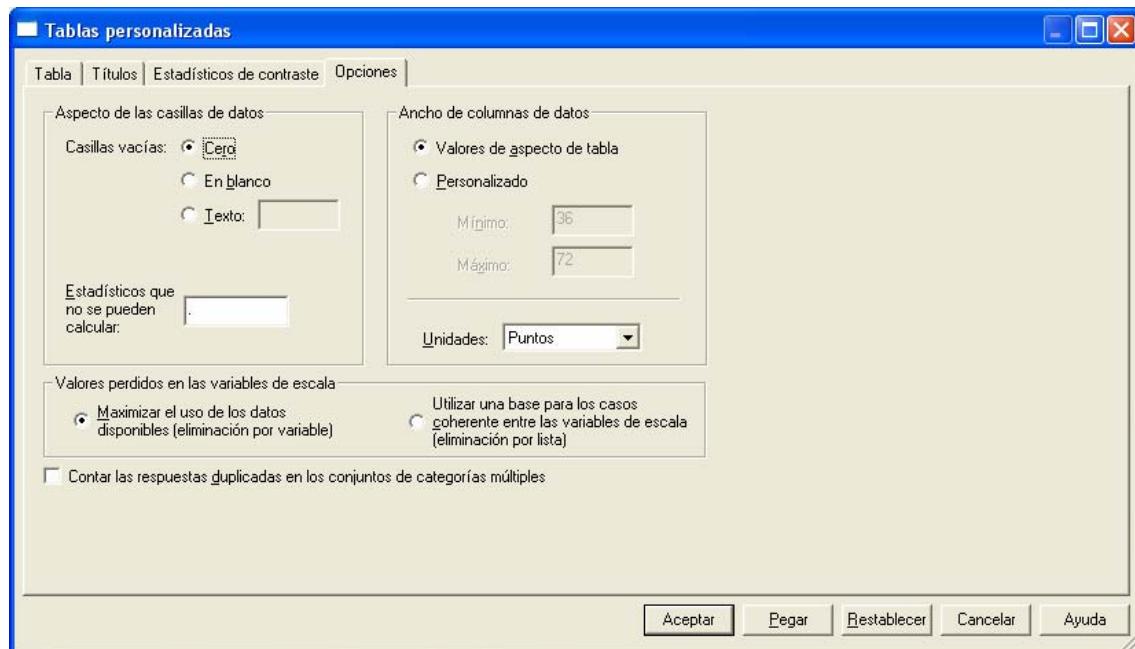
Importante: Se debe seleccionar la posición de presentación de los subtotales antes de definir los subtotales. El cambio en la posición de presentación puede afectar a todos los subtotales (y no únicamente al subtotal seleccionado actualmente) y, del mismo modo, cambiará las categorías incluidas en los subtotales.

Con la opción Valores perdidos se controla la presentación de valores perdidos definidos por el usuario o de valores que contienen valores perdidos (por ejemplo, un código de 99 para representar "no procede" para el embarazo en el caso de los hombres). Por defecto, se excluyen los valores perdidos definidos por el usuario. Si marcamos esta opción, incluiremos en las tablas las categorías perdidas definidas por el usuario, y aunque la variable puede contener más de una categoría de valores perdidos, la vista previa de la tabla del panel de lienzo sólo mostrará una categoría genérica de valor perdido.

Con la opción Categorías vacías incluiremos en las tablas estas categorías. Además con la opción Otros valores encontrados al explorar los datos, incluiremos automáticamente en las tablas los valores de las categorías del archivo de datos que no tienen definida la etiqueta de valor.

3.3.- Opciones adicionales

Algunas opciones adicionales se encuentran en la pestaña Opciones de Tablas personalizadas:



Como vemos, este cuadro de diálogo permite:

- Especificar lo que se visualiza en las casillas vacías y en las casillas para las que no se pueden calcular estadísticos.
- Controlar cómo se van a tratar los valores perdidos en el cálculo de los estadísticos de las variables de escala.
- Establecer los anchos mínimo y/o máximo de las columnas de datos.
- Controlar el tratamiento de las respuestas duplicadas en los conjuntos de categorías múltiples.

Aspecto de las casillas de datos. Controla lo que se visualiza en las casillas vacías y en las casillas para las que no se pueden calcular estadísticos.

- Casillas vacías. Para las casillas de la tabla que no contienen casos (frecuencia de casilla 0), se puede seleccionar una de las siguientes tres opciones de presentación: Cero, en blanco o un texto que se especifique. El valor de texto puede tener hasta 255 caracteres.
- Estadísticos que no se pueden calcular. Texto que aparece si no se puede calcular un estadístico (por ejemplo, la media para una categoría sin casos). El valor de texto puede tener hasta 255 caracteres. El valor por defecto es un punto (.).

Ancho de las columnas de datos. Controla los anchos mínimo y/o máximo de las columnas de datos. Esta opción no afecta a los anchos de columna para las etiquetas de fila.

- Valores de aspecto de tabla. Utiliza la especificación del ancho de las columnas de datos definido en el aspecto de tabla por defecto actual. Mediante un Aspecto de tabla, podemos crear nuestro propio aspecto de tabla por defecto, utilizado al crear nuevas tablas, y podemos controlar los anchos tanto de la columna de las etiquetas de las filas como de las columnas de datos.
- Personalizado. Anula los valores por defecto del aspecto de tabla para el ancho de las columnas de datos. Habremos de especificar los anchos mínimo y/o máximo de las columnas de datos para la tabla y la unidad de medida: puntos, pulgadas o centímetros.

Valores perdidos en las variables de escala. Para las tablas con dos o más variables de escala, controla el tratamiento de los datos perdidos para los estadísticos de las variables de escala.

- Maximizar el uso de los datos disponibles (eliminación por variable). Todos los casos con valores válidos para cada variable de escala se incluyen en los estadísticos de resumen de esa variable de escala.
- Utilizar una base coherente para los casos entre las variables de escala (eliminación por lista). Los casos con valores perdidos para cualquier variable de escala de la tabla se excluyen de los estadísticos de resumen de todas las variables de escala de la tabla.

Contar las respuestas duplicadas en los conjuntos de categorías múltiples. Una respuesta duplicada es una misma respuesta para dos o más variables del conjunto de categorías múltiples. Por defecto, las respuestas duplicadas no se cuentan, pero puede ser perfectamente válido que desee incluirlas en el recuento (como, por ejemplo, en un conjunto de categorías múltiples que represente al fabricante de los tres últimos coches adquiridos por un encuestado).

Por su parte, en la pestaña Títulos podemos modificar:

- Título. Texto que aparece por encima de la tabla.
- Pie. Texto que aparece por debajo de la tabla y por encima de las notas al pie.
- Esquina. Texto que aparece en la esquina superior izquierda de la tabla. El texto de la esquina sólo aparece si la tabla contiene variables de fila y si la propiedad de las etiquetas de la dimensión de las filas de la tabla pivote está establecida en Anidadas. Este no es el valor por defecto del aspecto de tabla.

Además podemos incluir los siguientes valores generados automáticamente en el título de la tabla, el texto al pie o la etiqueta de esquina:

- Fecha. El formato de año, mes y día actuales depende de la configuración regional actual de Windows.
- Hora. El formato de horas, minutos y segundos actuales depende de la configuración regional actual de Windows.
- Expresión de tabla. Variables que se utilizan y cómo se utilizan en la tabla. Si una variable tiene una etiqueta de variable definida, se mostrará la etiqueta. En

la tabla generada, los siguientes símbolos indican cómo se utilizan las variables en la tabla:

- + indica variables apiladas.
- > indica anidación.
- BY indica cruce de variables o capas.

Por su parte, en la pestaña Estadísticos de contraste podemos obtener varias pruebas de significación para las tablas personalizadas, incluyendo:

- Pruebas chi-cuadrado de independencia.
- Pruebas sobre la igualdad de las medias de columna.
- Pruebas sobre la igualdad de las proporciones de columna.

Estas pruebas no estarán disponibles para variables de respuestas múltiples o tablas en las que las etiquetas de categoría se desplacen de su dimensión de tabla por defecto.

Pruebas de independencia (Chi-cuadrado). Esta opción genera una prueba chi-cuadrado sobre la independencia para las tablas en las que existe como mínimo una variable categórica en las filas y otra en las columnas. También se puede especificar el nivel alfa de la prueba, que debe tener un valor mayor que 0 e inferior a 1.

Comparar las medias de columna (pruebas t). Esta opción genera pruebas por parejas sobre la igualdad de las medias de columna, para tablas en las que existe como mínimo una variable categórica en las columnas y una variable de escala en las filas. Se puede seleccionar que los valores p de las pruebas sean corregidas mediante el método Bonferroni. También se puede especificar el nivel alfa de la prueba, que debe tener un valor mayor que 0 e inferior a 1.

Comparar las proporciones de columna (pruebas z). Esta opción genera pruebas por parejas sobre la igualdad de las proporciones de columna para las tablas en las que existe como mínimo una variable categórica en las columnas y otra en las filas. Se puede seleccionar que los valores p de las pruebas sean corregidas mediante el método Bonferroni. También se puede especificar el nivel alfa de la prueba, que debe tener un valor mayor que 0 e inferior a 1.

4.- Conjuntos de respuestas múltiples

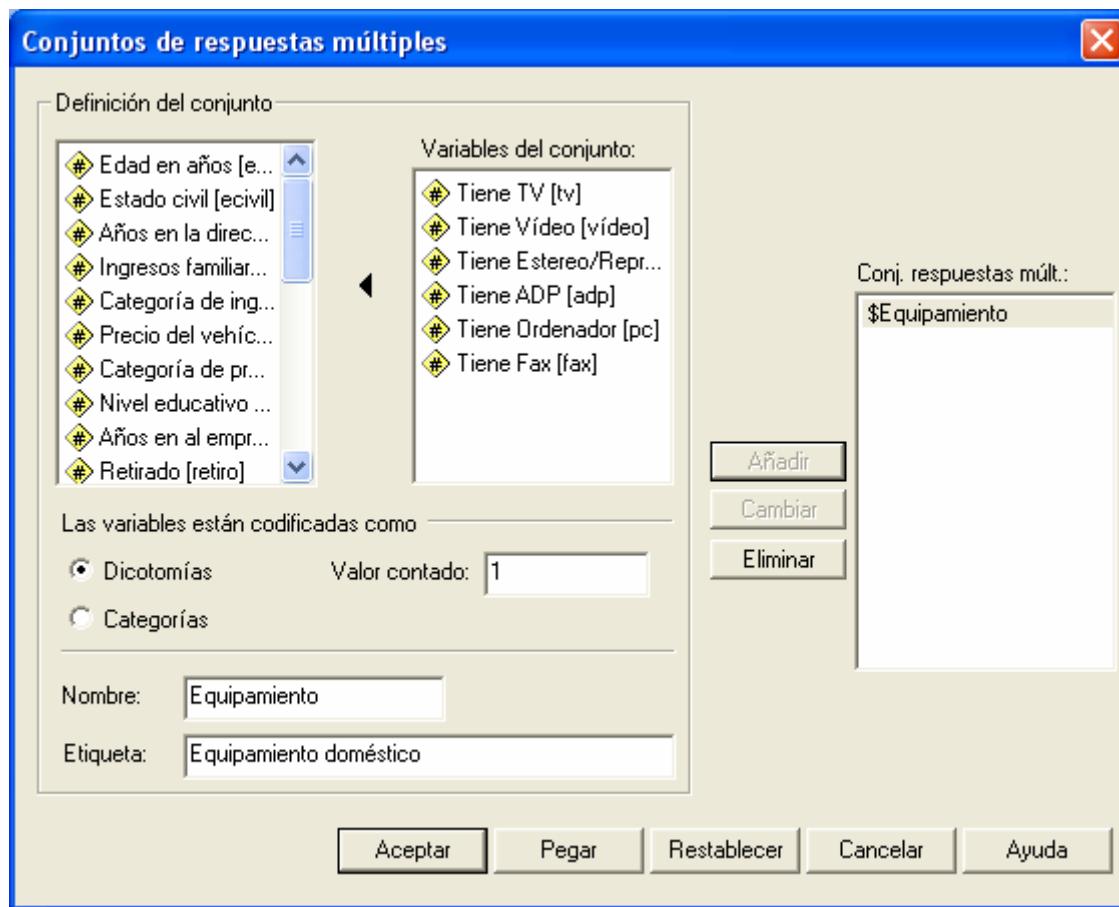
Las Tablas personalizadas admiten un tipo especial de "variable" al que se denomina conjunto de respuestas múltiples. En realidad, los conjuntos de respuestas múltiples no son, en sentido estricto, "variables". No aparecen en el Editor de datos y los demás procedimientos no los reconocen. Los conjuntos de respuestas múltiples utilizan varias variables para registrar respuestas a preguntas en las que el encuestado puede ofrecer más de una respuesta. Los conjuntos de respuestas múltiples se tratan como variables categóricas y, por consiguiente, la mayoría de las tareas que se pueden realizar con variables categóricas también se pueden llevar a cabo con los conjuntos de respuestas múltiples.

Los conjuntos de respuestas múltiples se crean a partir de múltiples variables del archivo de datos. Un conjunto de respuestas múltiples es un constructo especial

perteneciente a un archivo de datos con formato SPSS. Se pueden definir y guardar varios conjuntos en un archivo de datos con formato SPSS, pero no se pueden importar o exportar conjuntos de respuestas múltiples desde o a otros formatos de archivo. (Se pueden copiar conjuntos de respuestas múltiples de otros archivos de datos de SPSS mediante Copiar propiedades de datos en el menú Datos en la ventana Editor de datos.

Nota: Las Tablas personalizadas no admiten pruebas de significación para tablas que contienen conjuntos de respuestas múltiples.

Para definir conjuntos de respuestas múltiples accedemos a Analizar-Tablas-Conjuntos de respuestas múltiples. A continuación seleccionamos el conjunto de preguntas relativas al mismo tema, que definen la respuesta múltiple. Como ejemplo consideraremos la variable Equipamiento del modo siguiente:



El nombre asignado para el conjunto de respuestas múltiples es Equipamiento. El nombre puede tener hasta siete caracteres. De forma automática, se añade un signo dólar al comienzo del nombre del conjunto.

El cuadro de diálogo distingue:

- **Dicotomías:** Un conjunto de dicotomías múltiples se compone normalmente de múltiples variables dicotómicas: variables con sólo dos valores posibles del tipo sí/no, presente/ausente, seleccionado/no seleccionado. Si bien las variables pueden no ser estrictamente dicotómicas, todas las variables del conjunto se

codifican de la misma manera y el Valor contado representa la condición correspondiente a afirmativo/presente/seleccionado.

- Categorías: Un conjunto de categorías múltiples se compone de varias variables, todas ellas codificadas de la misma manera, a menudo con muchas posibles categorías de respuestas. Por ejemplo, un elemento de la encuesta pregunta, "Nombre hasta tres nacionalidades que mejor describan su herencia étnica." Puede haber cientos de respuestas posibles, pero por cuestiones de codificación se ha limitado la lista a las 40 nacionalidades más comunes, con cualquier otra opción relegada a la categoría "otras". En el archivo de datos, las tres opciones se convierten en tres variables, cada una con 41 categorías (40 nacionalidades codificadas más la categoría "otras").

Por tanto, en nuestro caso habremos de marcar dicotomías puesto que cada una de las variables tienen respuesta SÍ/NO, con la categoría SI de código 1. De este modo, el valor contado es 1, ya que, como decimos, es el código asignado a la tenencia de ese equipamiento (TV, vídeo, etc.).

Con el ejemplo anterior, observamos que habrá de tenerse en cuenta ciertas reglas para definir conjuntos de respuestas múltiples:

- Todas las variables del conjunto deben estar codificadas de la misma manera.
- Las etiquetas de valor se deben utilizar de forma consistente. Si una variable tiene definidas las etiquetas de valor, todas las demás variables deberán tener el mismo valor asignado a las mismas etiquetas de valor.
- En los conjuntos de dicotomías múltiples, cualquier etiqueta de variable definida para cada una de las variables del conjunto debe ser exclusiva. Dos o más variables del conjunto no deben tener la misma etiqueta de variable.

En cuanto a los estadísticos de resumen, en este procedimiento se encuentran disponibles:

- % de respuestas de la columna/fila/capa. Porcentaje basado en las respuestas.
- % de respuestas de la columna/fila/capa (Base: Recuento). Las respuestas son el numerador y el recuento total es el denominador.
- % de recuento de columnas/filas/capas (Base: Respuestas). El recuento es el numerador y las respuestas totales son el denominador.
- % de respuestas de la columna/fila de capas. Porcentaje de todas las subtablas. Porcentaje basado en las respuestas.
- % de respuestas de la columna/fila de capas (Base: Recuento). Porcentajes de todas las subtablas. Las respuestas son el numerador y el recuento total es el denominador.
- % de respuestas de la columna/fila de capas (Base: Respuestas). Porcentajes de todas las subtablas. El recuento es el numerador y las respuestas totales son el denominador.

Para el recuento de las respuestas, tenemos:

- % de respuestas de la subtabla/tabla. Porcentaje basado en las respuestas.
- % de respuestas de la subtabla/tabla (Base: Recuento). Las respuestas son el numerador y el recuento total es el denominador.

- % del recuento de subtablas/tablas (Base: Respuestas). El recuento es el numerador y las respuestas totales son el denominador.