

# **Informática Aplicada a la Administración**

**CLAVE: MA**

**PROFESOR: MTRO. ALEJANDRO SALAZAR GUERRERO**

## **1. INTRODUCCIÓN**

- 1.1. El hardware
- 1.2. Componentes hardware
- 1.3. El software
- 1.4. Sistema operativo
- 1.5. Windows

## **2. OFFICE**

- 2.1. Procesador de palabras
  - 2.1.1. Escritura y formatos
  - 2.1.2. Menús en Word
- 2.2. Hoja de cálculo
  - 2.2.1. Menús Excel
  - 2.2.2. Fórmulas
  - 2.2.3. Funciones personalizadas
  - 2.2.4. Gráficos y macros
- 2.3. Presentaciones en PowerPoint
  - 2.3.1. Menús PowerPoint
  - 2.3.2. Inserción y modificación de imágenes, gráficas y textos
  - 2.3.3. Animación

## **3. SPSS**

- 3.1. Módulos del programa y normas generales
  - 3.1.1. Tipos de ficheros
  - 3.1.2. Tipos de ventanas
  - 3.1.3. Personalización menús y barras de herramientas
- 3.2. Definición de las variables
  - 3.2.1. Matriz de datos
  - 3.2.2. Definición formato de columnas
  - 3.2.3. Etiquetado de variables y valores
  - 3.2.4. Modificación de variables
  - 3.2.5. Definición y uso de los conjuntos de variables
- 3.3. Datos
  - 3.3.1. Introducción de datos
  - 3.3.2. Edición de datos
  - 3.3.3. Importación y exportación de ficheros de datos
  - 3.3.4. Administración y modificación de ficheros de datos
- 3.4. Tabulación de datos y análisis exploratorio de datos
- 3.5. Tablas de contingencias y representaciones gráficas

## **4. FLASH**

- 4.1. Introducción a Flash
- 4.2. Capas , dibujo símbolos y fotogramas
- 4.3. Animación
  - 4.3.1. Animación fotograma a fotograma
  - 4.3.2. Animación por interpolación de forma
  - 4.3.3. Animación por interpolación de movimiento
  - 4.3.4. Combinación de animaciones
- 4.4. Acciones en Flash
  - 4.4.1. Panel de acciones

## **5. INTERNET**

- 5.1. Historia del internet
- 5.2. Ventana de internet Explorer
- 5.3. Tipos de páginas web
- 5.4. Dreamweaver

### 3. SPSS

Statistical Package for the Social Sciences (SPSS) es un programa estadístico informático muy usado en las ciencias sociales y las empresas de investigación de mercado. En la actualidad, las siglas se usan tanto para designar el programa estadístico como la empresa que lo produce. Originalmente SPSS fue creado como el acrónimo de Statistical Package for the Social Sciences ya que se está popularizando la idea de traducir el acrónimo como "Statistical Product and Service Solutions". Sin embargo, aunque realizando búsquedas por internet estas pueden llevar a la página web de la empresa, dentro de la página misma de la empresa no se encuentra dicha denominación.

Fue creado en 1968 por Norman H. Nie, C. Hadlai (Tex) Hull y Dale H. Bent. Entre 1969 y 1975 la Universidad de Chicago por medio de su National Opinion Research Center estuvo a cargo del desarrollo, distribución y venta del programa. A partir de 1975 corresponde a SPSS Inc.

Originalmente el programa fue creado para grandes computadores. En 1970 se publica el primer manual de usuario del SPSS por Nie y Hall. Este manual populariza el programa entre las instituciones de educación superior en EE. UU. En 1984 sale la primera versión para computadores personales.

Como programa estadístico es muy popular su uso debido a la capacidad de trabajar con bases de datos de gran tamaño. En la versión 12 es de 2 millones de registros y 250.000 variables. Además, de permitir la recodificación de las variables y registros según las necesidades del usuario. El programa consiste en un módulo base y módulos anexos que se han ido actualizando constantemente con nuevos procedimientos estadísticos. Cada uno de estos módulos se compra por separado.

Actualmente, compete no solo con softwares licenciados como lo son SAS, MatLab, Statistica, Stata, sino también con software de código abierto y libre, de los cuales el más destacado es el Lenguaje R.

Desde la versión 14, pero más específicamente desde la versión 15 se ha implantado la posibilidad de hacer uso de las librerías de objetos del SPSS desde diversos lenguajes de programación. Aunque principalmente se ha implementado para Python, también existe la posibilidad de trabajar desde Visual Basic, C++ y otros lenguajes

#### 3.1. Módulos del programa y normas generales

El sistema de módulos de SPSS, como los de otros programas (similar al de algunos lenguajes de programación) provee toda una serie de capacidades adicionales a las existentes en el sistema base. Algunos de los módulos disponibles son:

- Modelos de Regresión
- Modelos Avanzados
  - Reducción de datos: Permite crear variables sintéticas a partir de variables colineales por medio del Análisis Factorial.
  - Clasificación: Permite realizar agrupaciones de observaciones o de variables (cluster analysis) mediante tres algoritmos distintos.

- Pruebas no paramétricas: Permite realizar distintas pruebas estadísticas especializadas en distribuciones no normales.
- Tablas: Permite al usuario dar un formato especial a las salidas de los datos para su uso posterior. Existe una cierta tendencia dentro de los usuarios y de los desarrolladores del software por dejar de lado el sistema original de TABLES para hacer uso más extensivo de las llamadas CUSTOM TABLES.
- Tendencias
- Categorías: Permite realizar análisis multivariados de variables normalmente categorías. También se pueden usar variables métricas siempre que se realice el proceso de recodificación adecuado de las mismas.
- Análisis Conjunto: Permite realizar el análisis de datos recogidos para este tipo específico de pruebas estadísticas.
- Mapas: Permite la representación geográfica de la información contenida en un fichero (descontinuado para SPSS 16).
- Pruebas Exactas: permite realizar pruebas estadísticas en muestras pequeñas.
- Análisis de Valores Perdidos: Regresión simple basada en imputaciones sobre los valores ausentes.
- Muestras Complejas: permite trabajar para la creación de muestras estratificadas, por conglomerados u otros tipos de muestras.
- SamplePower (cálculo de tamaños muestrales)
- Árboles de Clasificación: Permite formular árboles de clasificación y/o decisión con lo cual se puede identificar la conformación de grupos y predecir la conducta de sus miembros.
- Validación de Datos: Permite al usuario realizar revisiones lógicas de la información contenida en un fichero.sav. y obtener reportes de los valores considerados extraños. Es similar al uso de sintaxis o scripts para realizar revisiones de los ficheros. De la misma forma que estos mecanismos es posterior a la digitalización de los datos.
- SPSS Programmability Extensión (SPSS 14 en adelante). Permite utilizar el lenguaje de programación Python para un mejor control de diversos procesos dentro del programa que hasta ahora eran realizados principalmente mediante scripts (con el lenguaje SAX Basic). Existe también la posibilidad de usar las tecnologías .NET de Microsoft para hacer uso de las librerías del SPSS. Aunque algunos usuarios han cuestionado sobre la necesidad de incluir otros lenguajes, la empresa no tiene esto entre sus objetivos inmediatos.

Desde el SPSS/PC hay una versión adjunta denominada SPSS Student que es un programa completo de la versión correspondiente pero limitada en su capacidad en cuanto al número de registros y variables que puede procesar. Esta versión es para fines de enseñanza del manejo del programa

### 3.1.1. Tipos de ficheros

SPSS tiene un sistema de ficheros en el cual el principal son los archivos de datos (extensión. SAV). Aparte de este tipo existen otros dos tipos de uso frecuente:

Archivos de salida (output, extensión. SPO): en estos se despliega toda la información de manipulación de los datos que realizan los usuarios mediante las ventanas de comandos. Son susceptibles de ser exportados con varios formatos (originalmente HTML, RTF o TXT, actualmente la versión 15 incorpora la exportación a PDF junto a los formatos XLS y DOC que ya se encontraban en la versión 12)

Archivos de sintaxis (extensión. SPS): Casi todas las ventanas de SPSS cuentan con un botón que permite hacer el pegado del proceso que el usuario desea realizar. Lo anterior genera un archivo de sintaxis donde se van guardando todas las instrucciones que llevan a cabo los comandos del SPSS. Este archivo es susceptible de ser modificado por el usuario. Muchos de los primeros usuarios del SPSS suelen escribir estos archivos en vez de utilizar el sistema de pegado del programa.

Existe un tercer tipo de fichero: el fichero de scripts (extensión. SBS). Este fichero es utilizado por los usuarios más avanzados del software para generar rutinas que permiten automatizar procesos muy largos y/o complejos. Muchos de estos procesos suelen no ser parte de las salidas estándar de los comandos del SPSS, aunque parten de estas salidas. Buena parte de la funcionalidad de los archivos de scripts ha sido ahora asumida por la inserción del lenguaje de programación Python en las rutinas de syntax del SPSS. Procedimientos que antes solo se podían realizar mediante scripts ahora se pueden hacer desde el syntax mismo.

El programa cuando se instala trae un determinado número de ejemplos o utilidades de casi todos los ficheros en cuestión. Estos son usados para ilustrar algunos de los ejemplos de uso del programa.

### 3.1.2. Tipos de ventanas

El SPSS está organizado en dos grandes ventanas: el EDITOR DE DATOS y el VISOR DE RESULTADOS. En cada una de ellas tendremos: un menú principal con distintas opciones correspondientes al tipo de ventana; la barra de herramientas que contiene una serie de iconos con los que realizar tareas de una forma directa; y por último, la barra de estados, que nos indicará el estado del procesador de SPSS, dirá si está preparado.

#### Editor de datos

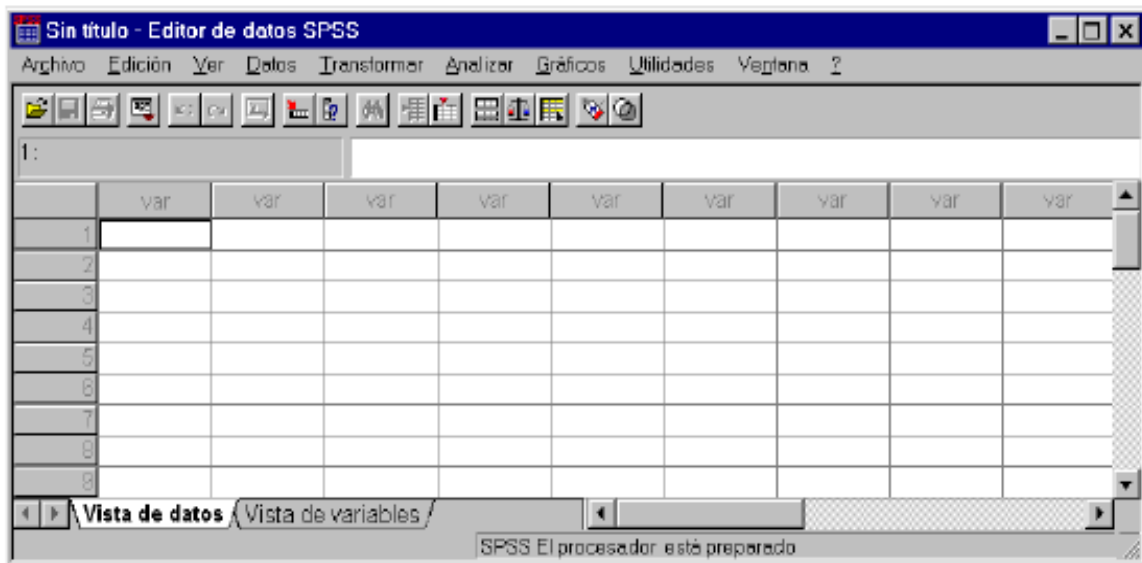
La primera ventana que encontraremos al entrar en SPSS nos permitirá, entre otras cosas, abrir un fichero de datos creado anteriormente o bien introducir nuevos datos. Si elegimos esta última opción o pulsamos CANCELAR apareceremos en el editor de datos. La ventana del editor de datos tiene una estructura similar a la de una hoja de cálculo y se utiliza para introducir los datos que se quieren analizar.

Los datos se introducen en la pestaña de vista de datos (en la parte inferior izquierda), dentro de variables (cada variable aparece como una columna en la pantalla). En la pestaña de vista de variables se explicarán las características de las mismas.

Las opciones que presenta el menú principal de la ventana del editor de datos son:

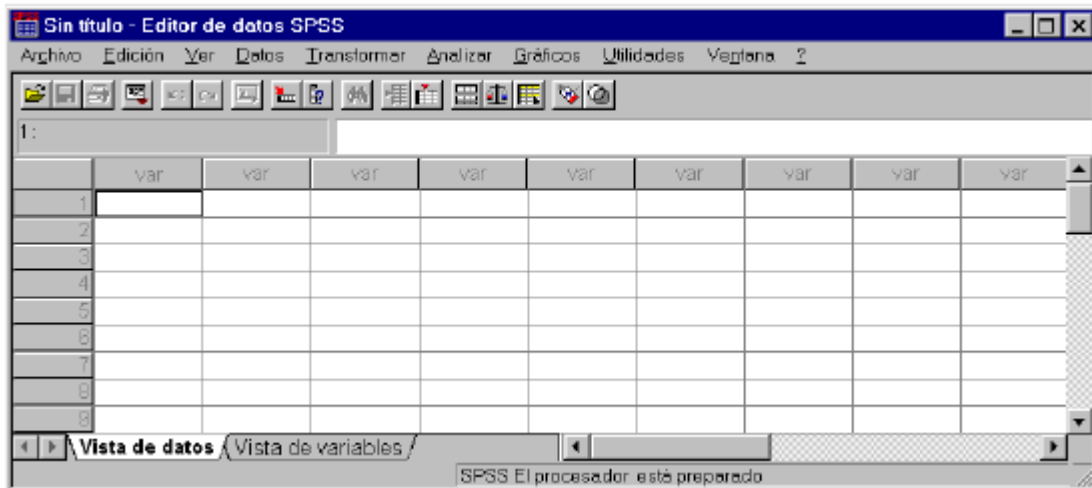
- ARCHIVO: se puede ABRIR, GUARDAR, IMPRIMIR archivos y SALIR de SPSS. Para abrir un archivo de datos de SPSS ya existente, se pulsa ARCHIVO → ABRIR →
- DATOS. En el cuadro que aparecerá a continuación, se selecciona la carpeta donde esté almacenado el archivo. De esta forma, aparecen los datos de este fichero en el editor. Para crear uno nuevo, ARCHIVO → NUEVO → DATOS.
- EDICIÓN: se puede copiar, pegar, eliminar, seleccionar y buscar texto. Contiene la opción
- DESHACER que sirve para anular la última operación realizada.

- VER: se pueden elegir distintas opciones sobre los que queremos que aparezca en la pantalla.
- DATOS: permite insertar una nueva variable; añadir una nueva observación, es decir, un nuevo caso; ordenar los casos mediante algún criterio, etc.



Muchas de estas opciones también se pueden realizar directamente sobre el fichero de datos.

- TRANSFORMAR: permite, entre otras cosas, construir nuevas variables que sean transformaciones de otras existentes, generar números aleatorios o agrupar valores en intervalos.
- ANALIZAR: contiene métodos estadísticos para ser aplicados sobre los datos incluidos en las variables.
- GRÁFICOS: permite realizar distintas representaciones gráficas (histogramas, diagramas de sectores, etc.) de los datos de las variables.
- UTILIDADES: contiene una serie de herramientas con las que se puede, por ejemplo, obtener información sobre variables y archivos, definir conjuntos de variables y utilizar sólo las variables seleccionadas en esos conjuntos, etc.
- VENTANA: permite activar cualquiera de las ventanas de SPSS o modificar el modo en que se disponen en la pantalla.
- El menú '?' contiene los comandos del sistema de ayuda de SPSS.



Muchas de las opciones del menú principal pueden realizarse directamente con los botones de la barra de herramientas. Para obtener información sobre la tarea que realiza cada botón se sitúa el ratón sobre el icono, así aparecerá una etiqueta con una breve descripción.

### 3.1.3. Personalización menús y barras de herramientas

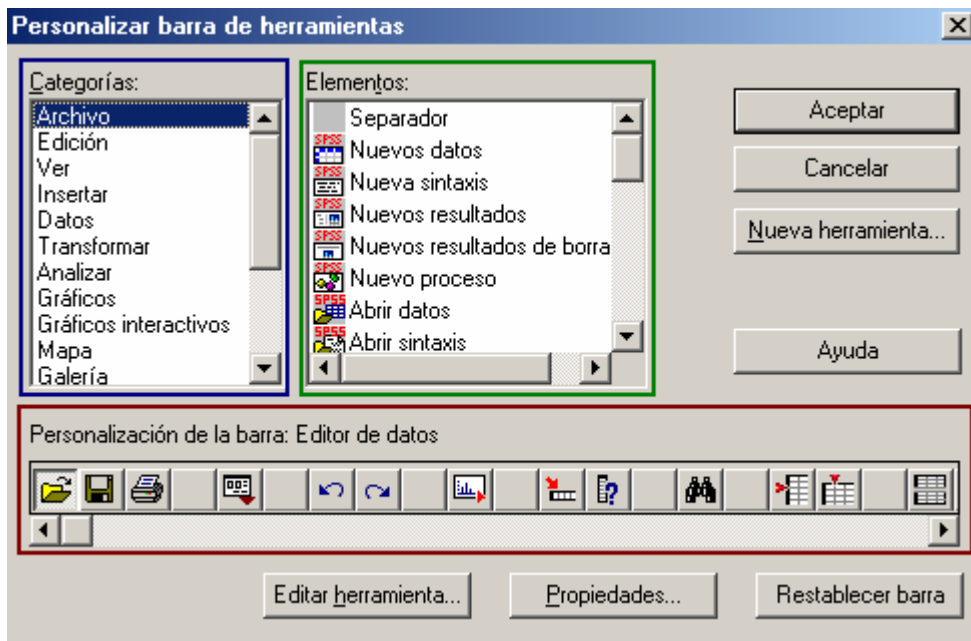
Los procedimientos que se incluyen en la barra de herramientas pueden ser modificados, extrayendo o ingresando los procedimientos que deseemos. Para realizar la personalización de la barra de herramientas, debemos ubicar el puntero del ratón sobre la barra de herramientas y hacer clic derecho sobre ella de manera que aparezca el menú desplegable [Fig.1-29].



Una vez aparece el menú, seleccionamos la opción personalizar con lo que aparece el cuadro de diálogo correspondiente. A través de este cuadro podemos personalizar las barras de herramientas existentes e incluso crear nuevas barras. En las barras de herramientas se puede incluir cualquier procedimiento disponible, o cualquier acción del menú.

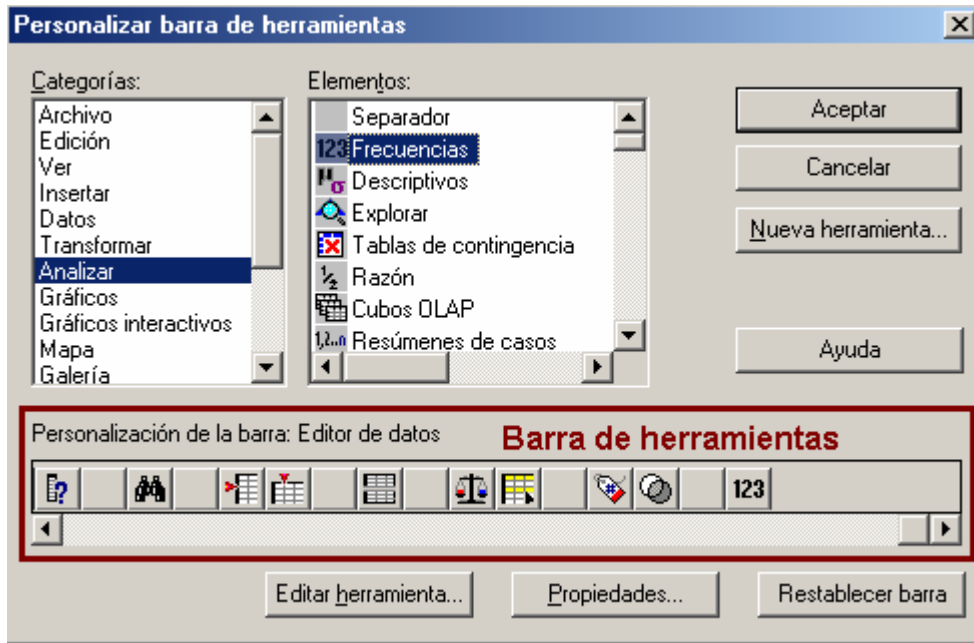
Para personalizar una barra de herramientas, debemos seleccionar en la lista de Categorías (Menús y opciones), la categoría en que se encuentre el procedimiento que deseamos incluir. Una vez se selecciona la Categoría, se actualizan en la lista de elementos los procedimientos que se incluyen dentro de ella. Para seleccionar el procedimiento basta con hacer clic sobre el y manteniendo oprimido el botón del ratón, arrastrarlo hasta la ubicación de la barra donde deseamos ingresarlo. Al soltar el botón del ratón, aparece en la barra el icono representativo del procedimiento seleccionado.





A manera de ejemplo ingresaremos en la barra de herramientas el procedimiento Frecuencias. Para realizarlo debemos seleccionar en la lista de categorías la opción Analizar, de manera que aparezca en la lista de elementos los procedimientos típicos de este menú. Una vez se actualiza el contenido, nos dirigimos a la barra de desplazamiento horizontal ubicada en la parte inferior del cuadro (Personalización de la barra Editor de datos) y la arrastramos hacia la derecha de manera que aparezca el extremo derecho de la barra de herramientas.

Después de aparecer el extremo de la barra, ubicamos en la lista de elementos la opción Separador ( ) en la parte superior de la lista de elementos; lo seleccionamos (Hacer clic) y manteniendo el botón del ratón oprimido lo arrastramos hacia el costado derecho de la barra del editor de datos, en donde lo soltamos. Una vez se suelta el separador, aparece en la barra un segmento sin icono; el objetivo de ingresar este separador, consiste en crear un espacio entre los botones usar conjuntos y Frecuencias que vamos a infiltrar. Después de ingresar el separador, introducimos el procedimiento Frecuencias, ubicándolo en la lista de elementos y llevándolo hasta el costado derecho de la barra de herramientas, en donde soltamos el botón de ratón y aparece el botón 123.



Una vez se ingresa el procedimiento a la barra de herramientas, hacemos clic en Aceptar con lo que se cierra el cuadro de diálogo y volvemos al editor de datos. Si nos fijamos en la barra de herramientas del editor de datos, notaremos que ahora aparece en ella el icono ( ), el cual representa el procedimiento Frecuencias; si hacemos clic en él se abrirá el cuadro de diálogo correspondiente. Este mismo procedimiento debe ser empleado para ingresar nuevas aplicaciones a la barra de herramientas.

### 3.2. Definición de las variables

Las variables en SPSS cuentan con una serie de propiedades que deben ser definidas por el investigador o usuario antes de realizar cualquier tipo de análisis con ella. De la correcta definición de las propiedades, depende en gran medida la calidad de los análisis que se realicen y por lo tanto la veracidad de los resultados o conclusiones que se generen. SPSS ha estipulado diez propiedades informativas de las variables entre las que encontramos:

#### I. Nombre:

Este parámetro nos permite identificar y diferenciar las variables que componen el archivo; para cada una de las variables se debe definir un nombre específico. El programa establece una serie de normas para los nombres de variables, entre las que encontramos:

Cada nombre de variable debe ser único; no se permiten duplicados.

La longitud del nombre no debe exceder los 64 bytes. Sesenta y cuatro bytes suelen equivaler a 64 caracteres en idiomas de un sólo byte (por ejemplo, inglés, francés, alemán, español, italiano, hebreo, ruso, griego, árabe, tailandés) y 32 caracteres en los idiomas de dos bytes (por ejemplo, japonés, chino, coreano).

El nombre debe comenzar por una letra. Los demás caracteres pueden ser letras, dígitos, puntos o los símbolos @, #, \_ o \$.

Los nombres de variable no pueden terminar en punto.

Se deben evitar los nombres de variable que terminan con subrayado (para evitar conflictos con las variables creadas automáticamente por algunos procedimientos). No se pueden utilizar espacios en blanco ni caracteres especiales (por ejemplo, !, ?, ' y \*). Las palabras reservadas (ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO, WITH) no se pueden utilizar como nombres de variable. Los nombres de variable se pueden definir combinando de cualquier manera caracteres en mayúsculas y en minúsculas, esta distinción entre mayúsculas y minúsculas se conserva en lo que se refiere a la visualización.

Para las versiones anteriores de SPSS (11.5, 11.0, 10.0, etc.) la longitud de las variables es de sólo ocho Bytes, lo cual generalmente no es suficiente para identificar una variable, por lo que es recomendable utilizar las tres primeras letras de cada palabra de la frase; es decir,

Estado Civil = estciv

Nivel de confianza = nivdecon

No necesariamente se debe seguir esta regla, lo realmente importante es que el nombre de la variable le permita identificar al usuario o investigador, el contenido a que se hace referencia; es decir, permitirle al usuario hacerse una idea del tema que abarca los datos de esa variable.

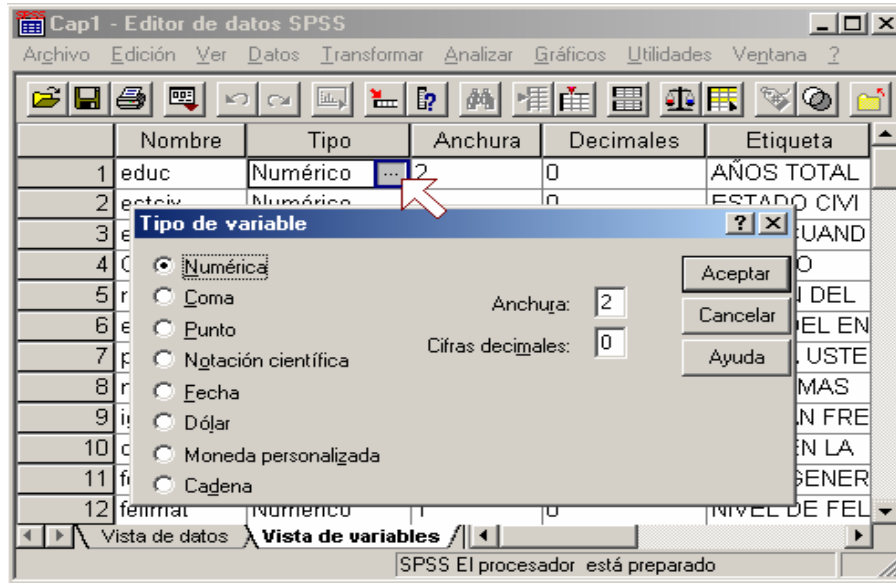
## II. Tipo:

La propiedad Tipo, nos permite especificarle al programa la naturaleza de los datos que se incluyen dentro de la variable; es decir, nos permite definir la forma y el significado de los caracteres que se encuentran en los registros de la variable. SPSS nos permite elegir entre ocho diferentes tipos de variables para representar Números (Magnitudes), Fechas (Tiempo), Monedas (Dinero) y Letras (Cadena). Desde luego es aconsejable trabajar las variables de forma numérica ya que el análisis estadístico es una ciencia matemática y para su correcto funcionamiento es necesario realizar las operaciones con números; ya que en algunos casos no es posible tener los datos de forma numérica, el paquete nos permite trabajarlos como una cadena de caracteres (Letras y Números).

Para definir el Tipo, debemos hacer clic en la casilla de la variable de interés, de manera que aparezca en el costado derecho de la casilla un pequeño cuadrado con puntos suspensivos (☰). Al seleccionar el botón (Hacer clic), aparece el cuadro de diálogo **Tipo de variable**, en donde aparecen los diferentes Tipos de variable que se pueden elegir para la variable seleccionada.

**Numérico:** Se emplea en una variable numérica cuyos valores representan magnitudes o cantidades y se asocian de forma estándar; es decir, asume la notación por defecto de Windows para la separación decimal (*Enteros* (,) *Decimales*) "1000,00"; este suele ser el tipo más usado.

**Coma y/o Punto:** Estos dos tipos de variables se emplean en una variable numérica cuyos valores representan magnitudes o cantidades. Al seleccionar la opción Coma los valores se asocian con comas que delimitan cada tres posiciones y con el punto como delimitador decimal "1,000.00". Cuando se selecciona el Punto los valores se asocian con puntos que delimitan cada tres posiciones y con la coma como delimitador decimal "1.000,00".



**Notación científica:** Se utiliza en una variable numérica cuyos valores son demasiado grandes o pequeños, por lo cual se emplea un exponente con signo que representa una potencia en base diez.  $1'000.000.00 = 1.0E+6$  ó  $0.000001 = 1.0E(-6)$ . SPSS nos permite representarlo de varias formas como 1000000, 1.0E6, 1.0D6, 1.0E+6, 1.0+6. La notación es útil cuando manejamos cifras extremas de lo contrario es mejor manejarlo de forma numérica.

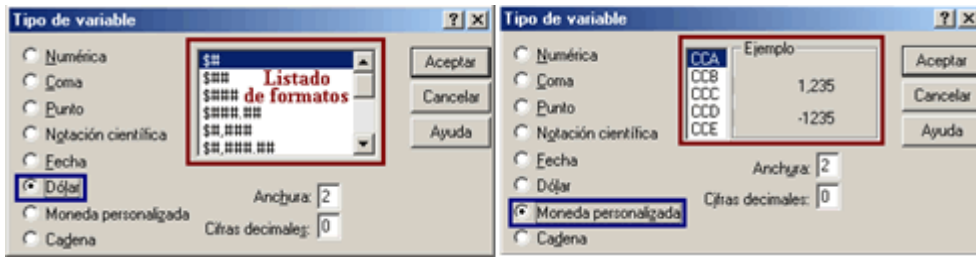
**Fecha:** Este tipo de variable se emplea cuando los valores de la variable representan fechas de calendario u horas de reloj; al seleccionarla aparece en el cuadro de diálogo una casilla con el listado de los diferentes formatos que el programa reconoce. Para elegir alguno de ellos basta con hacer clic sobre el formato y sucesivamente en Aceptar.

Anchura, Decimales y Etiquetas en SPSS

**Fecha:** Este tipo de variable se emplea cuando los valores de la variable representan fechas de calendario u horas de reloj; al seleccionarla aparece en el cuadro de diálogo una casilla con el listado de los diferentes formatos que el programa reconoce. Para elegir alguno de ellos basta con hacer clic sobre el formato y sucesivamente en Aceptar.



**Dólar:** se emplea en una variable numérica cuyos valores representan sumas de dinero en dólares. Al seleccionar este tipo de variable aparece en el cuadro de diálogo un listado de formatos monetarios, en donde debemos seleccionar el formato que más se acomode a los datos.



**Moneda personalizada:** Este tipo de variable se emplea cuando los valores de una variable representan sumas de dinero diferentes al dólar (Pesos, pesetas, Euros, etc.); al seleccionar esta opción aparece un nuevo listado, en el cual debemos seleccionar uno de los formatos existentes. Estos formatos no representan monedas específicas, sino que por el contrario el programa asume que la moneda es de origen distinto al dólar. La diferencia con el tipo dólar es que nos permite trabajar con cinco (5) diferentes tipos de moneda.

**Cadena:** Este tipo de variable se emplea cuando los valores no son numéricos o sencillamente no representan magnitudes o cantidades; estas variables no son utilizadas en los cálculos de los estadísticos. Las variables de cadena pueden contener cualquier tipo de caracteres siempre que no exceda la longitud máxima de 255; las mayúsculas y las minúsculas se consideran diferentes ya que el programa trabaja bajo el código ASCII. A este tipo de variables, también se le suele denominar como variable alfanumérica. Para definir alguno de los tipos de variable, basta con hacer clic sobre la opción que se desee y sucesivamente hacer clic en el botón *Aceptar*, con lo que se cierra la ventana y el tipo elegido aparece en la casilla seleccionada.

### III. Anchura:

Por medio de esta propiedad podemos definir el máximo de dígitos que contienen los registros de una variable; para el cálculo del ancho se incluyen los dígitos enteros y los decimales. Por ejemplo;

Anchura 5 = xxx.xx ó x,xxx.x ó xx,xxx donde x representa un número aleatorio.

No debemos cometer el error de pensar que una vez establecida la anchura, ya no podremos encontrar una cifra con mayor cantidad de números dentro de los registros. La opción Anchura se emplea para darle una idea al investigador, de las cifras que encontrará cuando le pida al paquete información de las variables, es decir, no restringe la cantidad de números sino que es un parámetro informativo, el cual le brinda a la persona que opere el programa una idea de los rangos máximos que puede tomar esta variable, pero no impide que se ingresen valores que sobrepasen esta longitud.

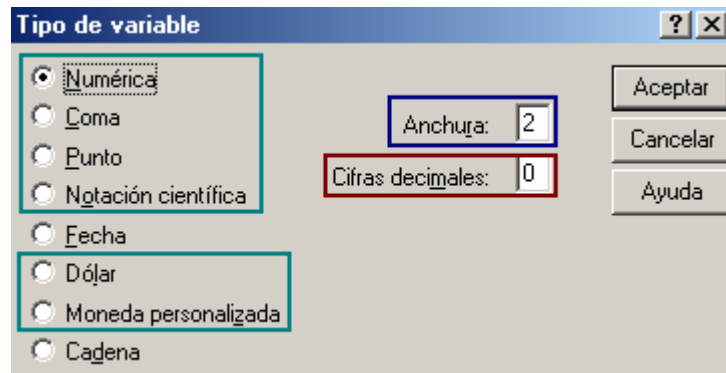
### IV. Decimales

A través de este parámetro se define el número de dígitos decimales que pueden contener los registros de la variable. Las cifras que superen esta longitud serán aproximadas por el programa. Cuando una cifra supera la longitud, el programa aproxima hacia arriba los dígitos que sobrepasen la longitud si el valor del último de ellos es igual o mayor que cinco, de lo contrario (menor que 5) se aproxima hacia abajo; es decir:

1.07X si  $X < 5$  entonces se aproxima a 0 es decir = 1.07

1.07X si  $X \Rightarrow 5$  entonces se aproxima a 10 es decir = 1.08

Las propiedades Anchura y Decimales pueden ser editadas directamente desde la ventana de *Tipo de variable* cuando se eligen los tipos numéricos de variables Numérica, Coma, Punto, Notación científica, Dólar o Moneda personalizada [Fig.1-40], ya que al seleccionar estas opciones se habilita en el cuadro de diálogo las casillas Anchura y Decimales.



Hay que notar que cuando seleccionamos los Tipos de variables como la *Fecha* y *Cadena* estas propiedades se desactivan; esto se debe a que para el tipo de formato Fecha el programa ha predefinido estos parámetros y no podemos alterarlo, la única opción que tenemos es escoger otro formato de fecha; mientras que para el tipo cadena no se puede tener números decimales.

### V. Etiqueta

Dado que generalmente los sesenta y cuatro (64) caracteres del nombre (Versiones anteriores ocho [8]) y las normas que se deben cumplir, no permiten describir de forma clara la variable y el contenido de ella; SPSS nos brinda la posibilidad de utilizar una etiqueta por medio de la cual podemos describir la variable mediante la utilización de un máximo de 255 caracteres.

Estas tablas contienen la frecuencia y el porcentaje de las categorías de la variable Estado civil (Casado y Soltero); la primera tabla cuenta con etiquetas para el nombre de la variable y para las categorías de la variable, mientras que la segunda tabla no cuenta con etiquetas. Si nos fijamos en la tablas notaremos que para interpretar la segunda tabla encontramos dificultades ya que no podemos determinar que categoría representan los números cero (0) y uno (1). Esta misma dificultad puede presentarse cuando nosotros realizamos un análisis de datos y entregamos los resultados a una persona que no haya participado en los procedimientos; para evitar estos inconvenientes se sugiere definir las etiquetas de variable y de valores.

#### 3.2.1. Matriz de datos

**La matriz de datos:** Es un archivo utilizado para el análisis estadístico, su construcción es quizás el paso más importante para el desarrollo estadístico a nivel informático. Construye la matriz de datos y ya esta hecho el 50% del trabajo.

**Matriz de datos**

	A	B	C	D	E	F
1	Departamento	Sexo	Edad	Orientación	Escolaridad	Ingresos
2	Montevideo	Varón	18	Ciencias Sociales	11,4	10.000
3	Montevideo	Mujer	20	Economía	8,0	8.000
4	Montevideo	Mujer	21	Comunicación	7,5	12.500
5	Montevideo	Mujer	22	Comunicación	11,0	6.000
6	Artigas	Varón	20	Ingeniería	9,8	5.500
7	Cerro Largo	Varón	19	Ciencias Sociales	8,6	8.500
8	Canelones	Mujer	24	Educación	7,5	10.200
9	San José	Varón	35	Educación	11,0	13.000

Diagrama de la matriz de datos:

- Una caja verde "Variables" apunta a las columnas A-F.
- Una caja verde "Unidades de análisis" apunta a las filas 1-9.
- Una caja verde "Valores" apunta a los datos dentro de las celdas.

### 3.2.2. Definición formato de columnas

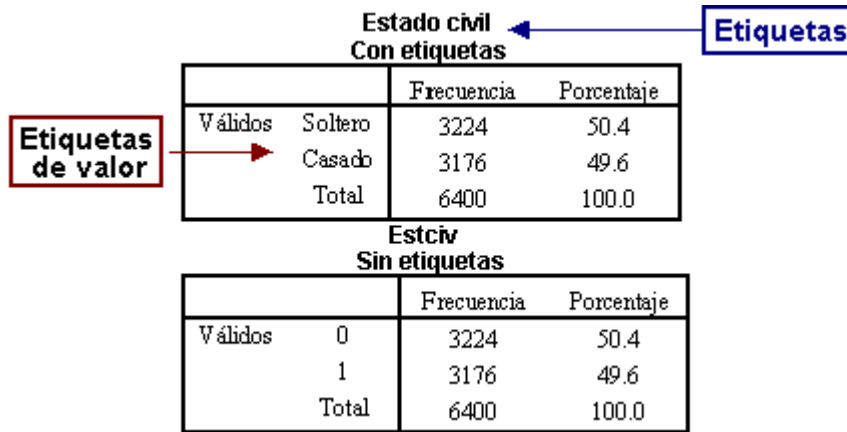
La mayoría de los procedimientos de SPSS generan resultados en forma de tablas, sin embargo, estos procedimientos no nos permiten modificar su estructura para adaptarlos a los propósitos de análisis. No obstante la opción Tablas de SPSS, nos ofrece una gran variedad de funciones especiales con las cuales podemos personalizar los informes, generando tablas específicas que cubran las necesidades o requerimientos del investigador.

A través del módulo de *Tablas* podemos crear diferentes tipos de tablas, involucrando variables Categóricas o de escala así como también conjuntos de respuesta múltiple (Dicotomías o Categóricas), de igual forma se le pueden aplicar a los datos una diversidad de estadísticos como la media, la mediana, el porcentaje de fila, el porcentaje de columna, etc.

### 3.2.3. Etiquetado de variables y valores

Dado que generalmente los sesenta y cuatro (64) caracteres del nombre (Versiones anteriores ocho [8]) y las normas que se deben cumplir, no permiten describir de forma clara la variable y el contenido de ella; SPSS nos brinda la posibilidad de utilizar una etiqueta por medio de la cual podemos describir la variable mediante la utilización de un máximo de 255 caracteres.

El uso de la etiqueta es bastante útil para facilitar la interpretación de los resultados (Tablas, Gráficos o estadísticos), para las personas que no han participado en la generación de los procedimientos y desconocen el significado del nombre de la variable. El uso de la etiqueta es opcional, el programa en caso de no existir una etiqueta utiliza el nombre de la variable para generar los resultados. Para saber si una variable tiene estipulada una etiqueta debemos ubicar el cursor del ratón sobre el nombre de la variable en la vista de datos, de manera que aparezca una leyenda informativa.



**Figura 1-41**

Estas tablas contienen la frecuencia y el porcentaje de las categorías de la variable Estado civil (Casado y Soltero); la primera tabla cuenta con etiquetas para el nombre de la variable y para las categorías de la variable, mientras que la segunda tabla no cuenta con etiquetas. Si nos fijamos en la tablas notaremos que para interpretar la segunda tabla encontramos dificultades ya que no podemos determinar que categoría representan los números cero (0) y uno (1). Esta misma dificultad puede presentarse cuando nosotros realizamos un análisis de datos y entregamos los resultados a una persona que no haya participado en los procedimientos; para evitar estos inconvenientes se sugiere definir las etiquetas de variable y de valores.

Antes de definir la propiedad *Valores* debemos ver primero las propiedades *Perdidos* y *Medida*, ya que la utilización de la etiquetas de valor está determinado por estos dos parámetros y en este momento no sería muy clara su definición.

### 3.2.4. Modificación de variables

Este es el parámetro más importante de las variables, de su definición depende el tipo de análisis que podemos realizar con el programa. Dentro de la estadística se han catalogado cuatro diferentes *escalas de medida*, pero para SPSS estas escalas se resumen en sólo tres:

- **Nominal:** son variables numéricas cuyos valores (Números) indican una categoría de pertenencia. Para este tipo de medida, las categorías no cuentan con un orden lógico que nos permita establecer una comparación de superioridad entre ellas. Un ejemplo de variable nominal puede ser el género, la raza, el estado civil, etc.
- **Ordinal:** son variables numéricas cuyos valores indican una categoría de pertenencia y a su vez las categorías poseen un orden lógico que nos indica una superioridad o prelación. Un ejemplo de variable ordinal puede ser el nivel de ingresos, categoría del vehículo, nivel educativo, etc.
- **Escala:** son variables numéricas cuyos valores representan una magnitud o cantidad y no una categoría; los valores de este tipo de medida pueden ser empleados en operaciones aritméticas como la suma, la resta, la multiplicación y la división ya que los intervalos (Distancia entre los números) cuentan con la misma longitud. Un ejemplo de variable de escala puede ser la edad, las ventas, la distancia en metros, la altura, etc.

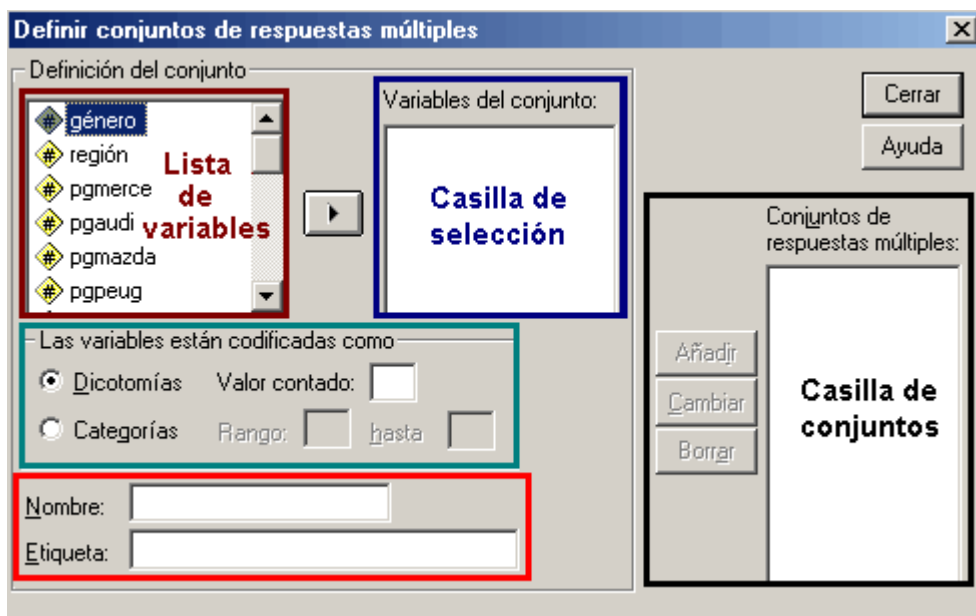


Para los archivos de datos con formato SPSS creados en versiones anteriores se aplican las siguientes reglas.

- Las variables de cadena (alfanuméricas) se establecen en nominales.
- Las variables de cadena y numéricas con etiquetas de valor definidas se establecen en ordinales.
- Las variables numéricas sin etiquetas de valor definidas que no superen un número específico de valores únicos (24), se establecen como ordinales, mientras que si el número de valores supera los 24 se definen como de Escala.

### 3.2.5. Definición y uso de los conjuntos de variables

Al seleccionar la opción *Definir Conjuntos* aparece el cuadro de diálogo correspondiente; a través de este cuadro se definen las propiedades del conjunto de variables que contienen las respuestas de una pregunta *Abierta* o *Cerrada*. Este cuadro se divide en cinco secciones entre las que encontramos la lista de variables, la casilla de selección, la definición del tipo de pregunta, las propiedades del conjunto y la casilla de conjuntos.



Al igual que en la mayoría de procedimientos del programa, en la casilla de selección se deben ingresar todas las variables de interés, que para el caso corresponden a las variables que contienen las respuestas de una misma pregunta. Las variables que se introduzcan en esta casilla definirán un conjunto, por lo que es de gran importancia verificar que no se agregue por error una variable que no contenga respuestas de la misma pregunta, ya que esto implicaría que no cuenta con las mismas propiedades de las variables de interés (Tipo, Medida, Valores, Perdidos), lo que puede causar que se presenten problemas al momento de generar las tablas de frecuencia o de contingencia.

Una vez se han ingresado las variables, es necesario definirle al programa el tipo de respuestas que contienen. En la sección *Las variables están codificadas como*, encontramos las opciones *Dicotomías* y *Categorías*. SPSS denomina *Dicotomías* a las respuestas de una pregunta cerrada y *Categorías* a las respuestas de preguntas Abiertas. Cuando se selecciona la opción *Dicotomías*, se activa la casilla *Valor contado*, en esta casilla se debe ingresar el valor que representa a las

respuestas validas (Si, existencia), el cual será tenido en cuenta para los cálculos. Cuando se selecciona la opción Categorías se activan las casillas del rango (Desde, Hasta); en estas casillas se deben ingresar los valores (Mínimo y Máximo) de los datos.

Después de especificar la codificación de las variables, se debe definir el nombre y la etiqueta del conjunto de datos, ingresando los caracteres en las respectivas casillas. Para el nombre del conjunto SPSS nos permite ingresar hasta un máximo de siete (7) caracteres, teniendo en cuenta no utilizar palabras clave como casenum, sysmis, jdate, date, time, length y width. Para la etiqueta del conjunto, por el contrario se puede ingresar hasta un máximo de 40 caracteres.

La última sección que encontramos en el cuadro de diálogo corresponde a la *casilla de conjuntos*, en ella aparecerán todos los conjuntos de variables que hayan sido definidos satisfactoriamente. Para facilitar su identificación cada nombre de conjunto esta precediendo con un signo de dólar (\$); es necesario recalcar que el programa no crea una nueva súper-variable en el editor de datos para guardar toda la información del conjunto, sino por el contrario cuando el nombre del conjunto es empleado en un procedimiento, SPSS lo sustituye por las variables que lo componen al momento de realizar los cálculos de los estadísticos (Tablas de frecuencia o Tablas de contingencia).

### 3.3. Datos

#### 3.3.1. Introducción de datos

SPSS debemos seleccionar el menú de Archivo/Nuevo/Datos:

Si ya estamos en el editor de datos vacío, no hace falta. Si, previamente, teníamos abierto otro archivo de datos, por ejemplo MUNDO.sav, y hemos realizado alguna modificación de este archivo, antes de cerrar este y abrir uno nuevo, nos preguntará si queremos guardar los cambios realizados. La respuesta es NO, de este modo, el archivo original no será modificado.

Una vez tengas la ventana del Editor de datos del SPSS abierta y vacía, lee atentamente las siguientes instrucciones antes de introducir datos:

- Debemos definir las variables adecuadamente (en Vista de variables) de forma que después podamos empezar a introducir los datos en cualquier orden (en Vista de datos).

Para introducir los datos, nos situamos en una casilla, tecleamos su valor y pulsamos Intro o el Tabulador o nos situamos con el ratón en otra casilla. Podemos introducir los datos por filas o columnas indistintamente.

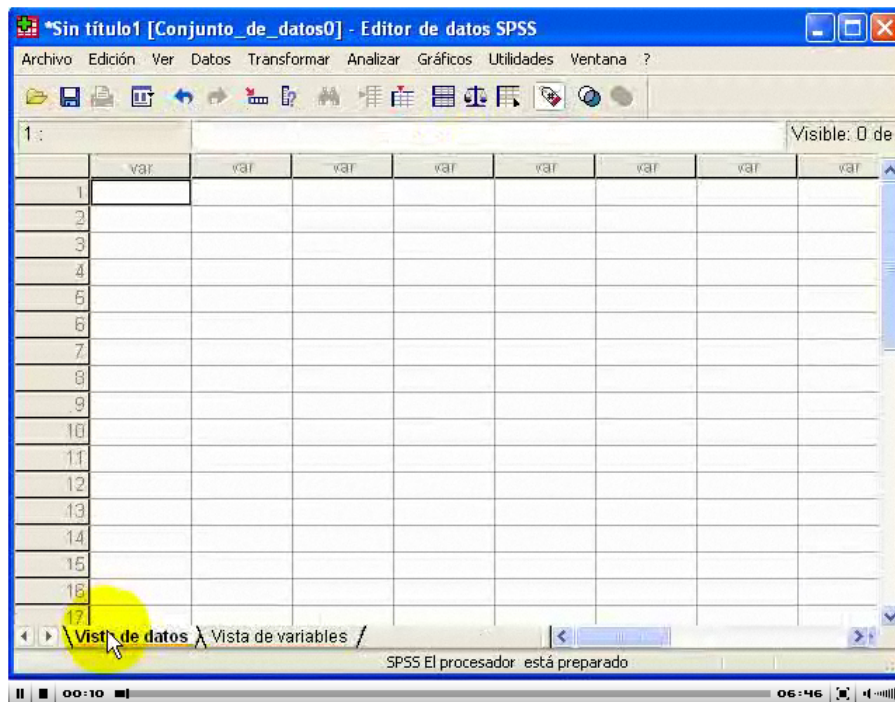
Conviene ir actualizando el archivo con la opción Guardar Datos o con el botón de grabar habitual en otras aplicaciones Windows.

- si ya lo habíamos guardado con anterioridad, para guardar las últimas modificaciones, seleccionaremos la opción de Guardar dentro de Archivo,
- si es la primera vez que lo guardamos, como todavía no le hemos asignado un nombre al archivo, seleccionaremos la opción Guardar como... dentro de Archivo. Nos pedirá el nombre que queremos darle al archivo, su extensión (por defecto utilizará .sav para datos del SPSS) y la carpeta en que queremos guardarlo.

### 3.3.2. Edición de datos

Una vez introducidos los datos podemos considerar el modificarlos, ya sea para introducir o ampliar con nuevos valores o para rectificar alguno ya introducido. La edición de datos en SPSS contiene comandos para cortar, copiar, pegar, eliminar y buscar (similar a otras aplicaciones Windows), insertar variables o casos, ir al caso, ordenar casos,... y se encuentran en el menú Edición y Datos.

Una vez elaborado el archivo de datos, y previamente al análisis estadístico de los datos es posible modificar los datos originales presentes en el archivo. Tales cambios pueden referirse a las variables, a los casos, o al archivo en su conjunto. Puede interesarnos tanto la recodificación de los valores de una variable como la generación de nuevas mediante transformaciones de las ya existentes.

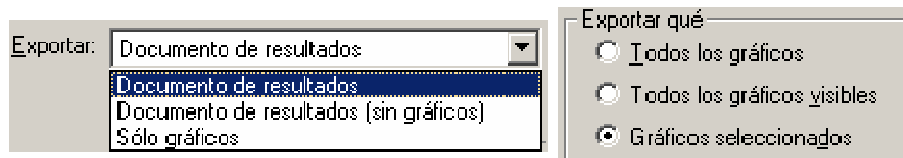


### 3.3.3. Importación y exportación de ficheros de datos

- Importar archivos de Excel a SPSS
  - Es muy fácil, sólo hay que tener en cuenta si la primera fila pone los nombres de las variables.
- Importar archivos de texto a SPSS
  - Los separadores entre datos pueden ser espacios en blanco, comas u otros caracteres.
  - No obstante, hoy en día, el standard más normal es utilizar tabuladores. Los tabuladores son caracteres invisibles que sólo marcan la separación.

Para exportar resultados de SPSS, debemos ir al menú *Archivo* y escoger la opción *Exportar*, de modo que surja el cuadro de diálogo correspondiente. A través de este cuadro se definen los parámetros que serán exportados, así como las propiedades del archivo resultante.

Para exportar los resultados, es necesario elegir en la lista de exportación el tipo de elementos que van a ser exportados. Se puede exportar los resultados y los gráficos, los resultados sin los gráficos ó sólo los gráficos. Una vez seleccionado el tipo de elementos, definimos el nombre del archivo resultante (Introduzca un nombre de archivo para los documentos de resultados o un nombre clave para los gráficos [si está seleccionada la opción sólo gráficos]). Por lo general los resultados son guardados bajo el nombre OUTPUT. Si deseamos cambiar el nombre es necesario ingresar en la casilla *Exportar archivo* una nueva ruta o un nuevo nombre para el archivo resultante.



Después de definir el nombre del archivo, determinamos en la sección *Exportar qué*, los elementos que vamos a remitir. En esta sección encontramos las opciones Todos los objetos (Tablas y gráficos), todos los objetos visibles y la opción objetos seleccionados.

- Exportar archivos de SPSS a Excel

Una de las alternativas más sobresalientes que se puede apreciar en el Visor de resultados corresponde a *Exportar*. A través de este procedimiento podemos enviar los resultados obtenidos mediante SPSS a una gran diversidad de formatos como Html (Paginas Web), de texto, Word/RTF y Excel. Esta opción nos permite compartir los resultados del paquete con nuestros colaboradores o incluso subirlos a Internet, a través de la creación de archivos de resultados en otros formatos de mayor difusión.

- Exportar archivos de SPSS a texto

- Hay que tener en cuenta que SPSS utiliza la terminación .dat en lugar de .txt para los archivos separados por tabuladores.

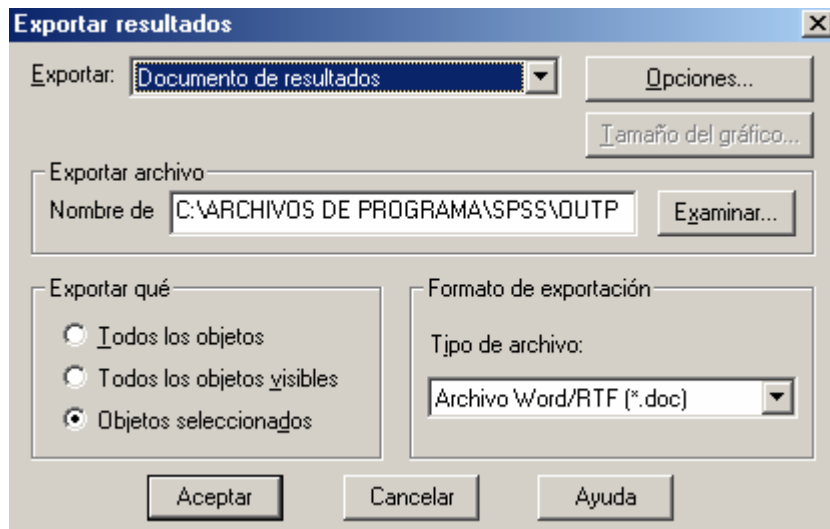
- Otros formatos

- SPSS es capaz de manejar otros formatos de manera apropiada

- ¿Qué hay de copiar y pegar?

- Copiar y pegar entre programas a veces funciona correctamente y a veces no. Hay que tener cuidado y asegurarse que todo ha funcionado bien.

- La importación de datos con el SPSS realiza comprobaciones que no se hacen con copiar y pegar.



Figuras 1-59

### 3.3.4. Administración y modificación de ficheros de datos

Consiste en la serie de actividades que se llevan a cabo con los archivos en forma automática, tales como: creación, grabación, fusión, borrar.

### 3.4. Tabulación de datos y análisis exploratorio de datos

La finalidad del Análisis Exploratorio de Datos (AED) es examinar los datos previamente a la aplicación de cualquier técnica estadística. De esta forma el analista consigue un entendimiento básico de sus datos y de las relaciones existentes entre las variables analizadas.

El AED proporciona métodos sencillos para organizar y preparar los datos, detectar fallos en el diseño y recogida de datos, tratamiento y evaluación de datos ausentes, identificación de casos atípicos y comprobación de los supuestos subyacentes en la mayor parte de las técnicas multivariantes.

El Análisis Exploratorio de Datos (A.E.D.) es un conjunto de técnicas estadísticas cuya finalidad es conseguir un entendimiento básico de los datos y de las relaciones existentes entre las variables analizadas. Para conseguir este objetivo el A.E.D. proporciona métodos sistemáticos sencillos para organizar y preparar los datos, detectar fallos en el diseño y recogida de los mismos, tratamiento y evaluación de datos ausentes (missing), identificación de casos atípicos (outliers) y comprobación de los supuestos subyacentes en la mayor parte de las técnicas multivariantes (normalidad, linealidad, homocedasticidad). El examen previo de los datos es un paso necesario, que lleva tiempo, y que habitualmente se descuida por parte de los analistas de datos. Las tareas implícitas en dicho examen pueden parecer insignificantes y sin consecuencias a primera vista, pero son una parte esencial de cualquier análisis estadístico.

#### ETAPAS DEL A.E.D.

Para realizar un A.E.D. conviene seguir las siguientes etapas:

- 1) Preparar los datos para hacerlos accesibles a cualquier técnica estadística.

- 2) Realizar un examen gráfico de la naturaleza de las variables individuales a analizar y un análisis descriptivo numérico que permita cuantificar algunos aspectos gráficos de los datos.
- 3) Realizar un examen gráfico de las relaciones entre las variables analizadas y un análisis descriptivo numérico que cuantifique el grado de interrelación existente entre ellas.
- 4) Evaluar, si fuera necesario, algunos supuestos básicos subyacentes a muchas técnicas estadísticas como, por ejemplo, la normalidad, linealidad y homocedasticidad.
- 5) Identificar los posibles casos atípicos (outliers) y evaluar el impacto potencial que puedan ejercer en análisis estadísticos posteriores.
- 6) Evaluar, si fuera necesario, el impacto potencial que pueden tener los datos ausentes (missing) sobre la representatividad de los datos analizados.

### 3.5. Tablas de contingencias y representaciones gráficas

Las tablas de contingencia nos permiten realizar comparaciones de Relación / Independencia entre dos o más variables categóricas, ya sean de tipo Nominal u Ordinal. En este tipo de tablas se definen las categorías de una variable a través de la frecuencia o el porcentaje de las categorías de una segunda variable.

Las tablas de contingencia pueden servir a diferentes propósitos como el análisis descriptivo, en cuyo caso el objetivo es proporcionar información condensada que describa las categorías de las variables involucradas, sin realizar comparaciones explícitas que generen conclusiones de una población. A su vez este tipo de tablas también pueden ser empleadas para la generación de análisis de Inferencia, en donde se intenta extraer conclusiones sobre las relaciones de las variables que puedan ser aplicadas a una población.

Para este objetivo se emplean pruebas estadísticas a los datos de una muestra (por lo general la prueba de independencia de Chi-cuadrado). Para generar una tabla de contingencia es necesario tener un archivo de datos abierto en el Editor de datos de SPSS.

En la sección Casillas de selección, encontramos las tres ubicaciones que podemos emplear para las variables de la tabla (Filas, Columnas o Capas); al ingresar una variable a la casilla Filas, sus categorías definen cada una de las filas de datos de la tabla; a su vez al ingresar una variable en la casilla Columnas, sus categorías definen cada una de las columnas de la tabla.

Si se ingresan múltiples variables a una de las casillas, el programa genera una tabla de contingencia para cada combinación de variables de fila y columna. Por último encontramos la casilla de *Capa*; al ingresar una variable en esta casilla, el programa genera una tabla de contingencia para cada una de las categorías de la variable de Capa.