

**Universidad Católica del Uruguay**  
**Facultad de Ciencias Humanas**

**Autotexto**

**Serie Estadística**

**Introducción al manejo de datos  
mediante SPSS**



**Universidad  
Católica**

DAMASO A. LARRAÑAGA • URUGUAY

**Laboratorio Metodológico**

Autores:

A. de León – M. Dodel – C. Rafaniello

**2008**

## ¿Cómo se organiza la información?

### La estructura tripartita de los datos

**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

Diagram illustrating the tripartite structure of data:

- Variables:** A box pointing to the column headers (A-F).
- Unidades de análisis:** A box pointing to the row numbers (1-9).
- Valores:** An oval pointing to the data cells within the matrix.

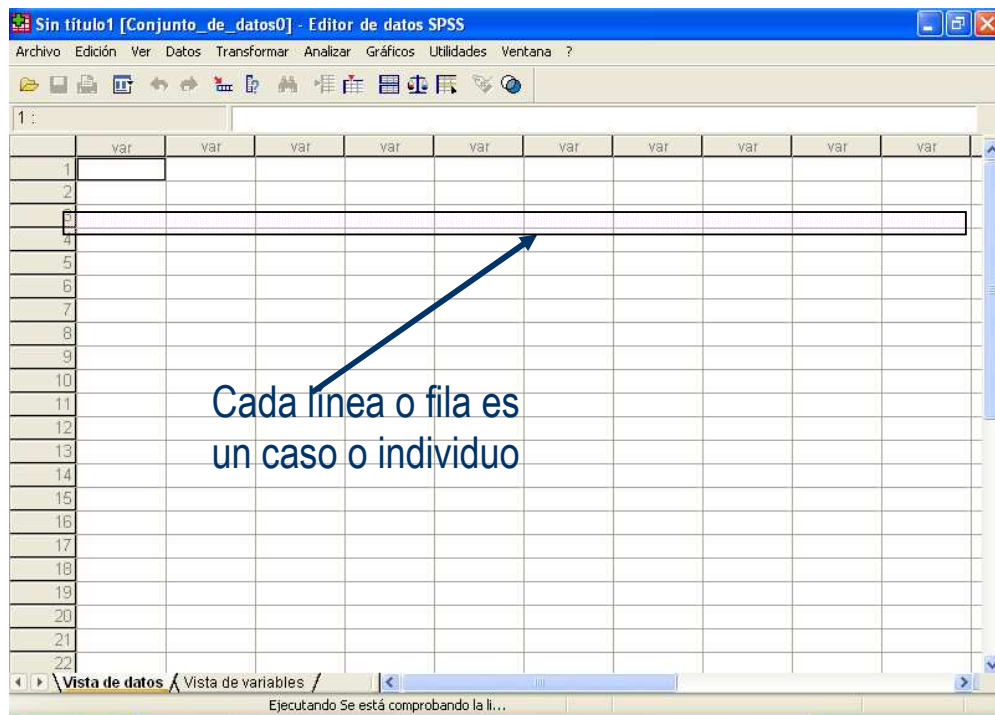
1. Unidades de análisis: Pueden ser personas, unidades colectivas (p.e., hogares, escuelas, sistemas políticos, países) o cualquier otra unidad de observación.

2. Variables: Son las dimensiones que vamos a observar en cada unidad: p.e., sexo, edad, ingresos, etc.

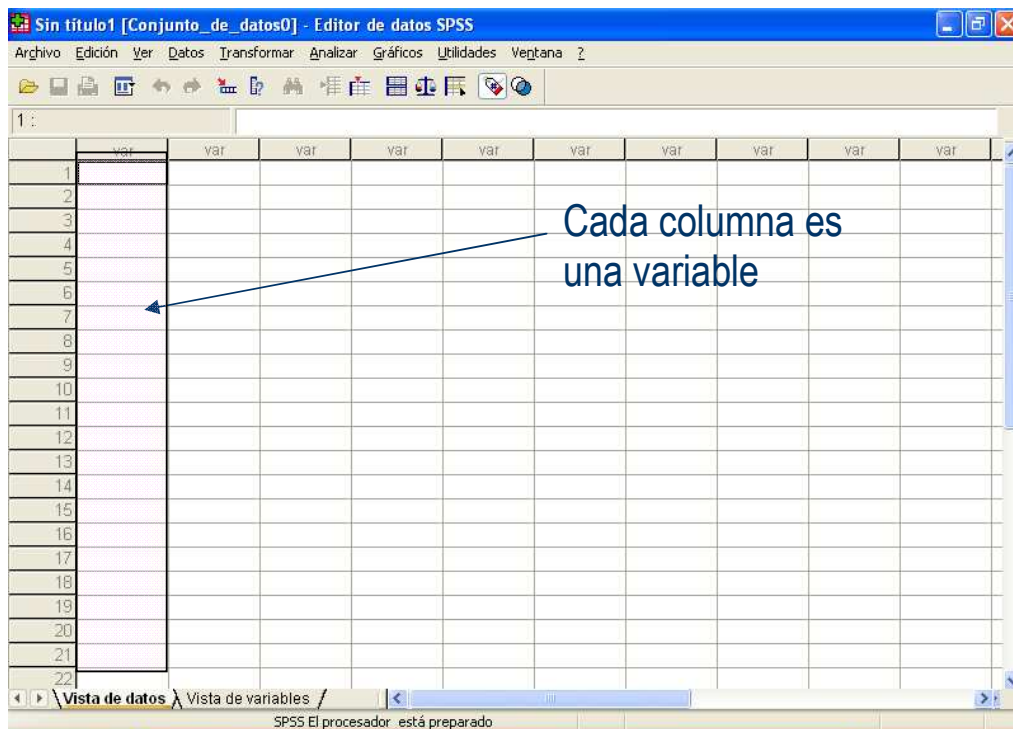
3. Valores: Son los resultados que asume cada unidad en las variables seleccionadas (p.e., varón, 13 años, etc)

## EDITOR DE DATOS: Database.sav

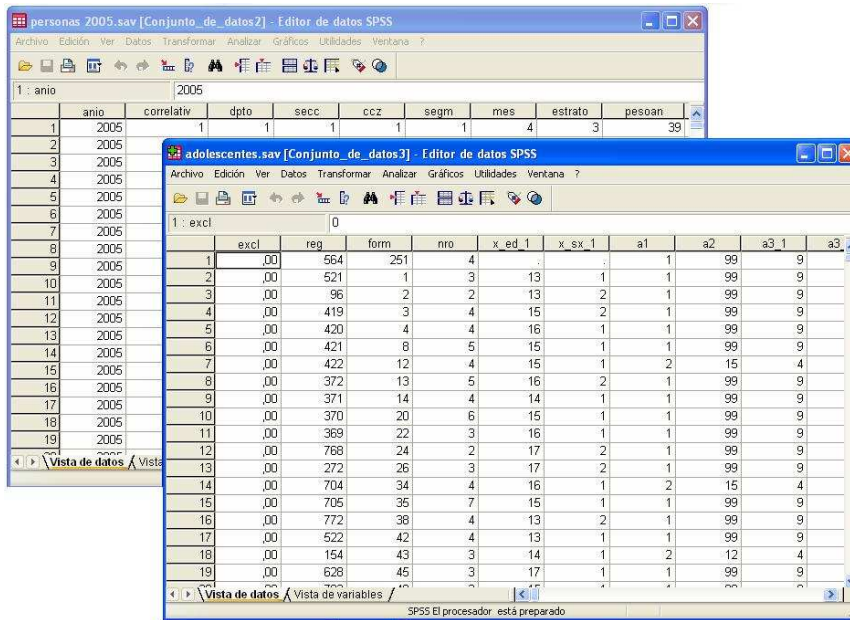
### Vista de datos:



La ventana del **Editor** de datos se abre automáticamente cuando se inicia una sesión de SPSS. Esta ventana muestra los contenidos del archivo de datos. Con el **Editor** se pueden crear nuevos archivos de datos o modificar los existentes, pero no se puede tener más de un archivo de datos abierto al mismo tiempo.



## A partir del SPSS versión 14.0



Con el **Editor** de esta versión del programa se pueden crear nuevos archivos de datos o modificar los existentes, **y también** se puede tener más de un archivo de datos abierto al mismo tiempo.

A pesar de ello, no es recomendable trabajar con más de un archivo de datos abierto a la vez, ya que esto aumenta la probabilidad de cometer errores al confundirse de archivos sobre los cuales se trabaja y enlentece al PC en general.

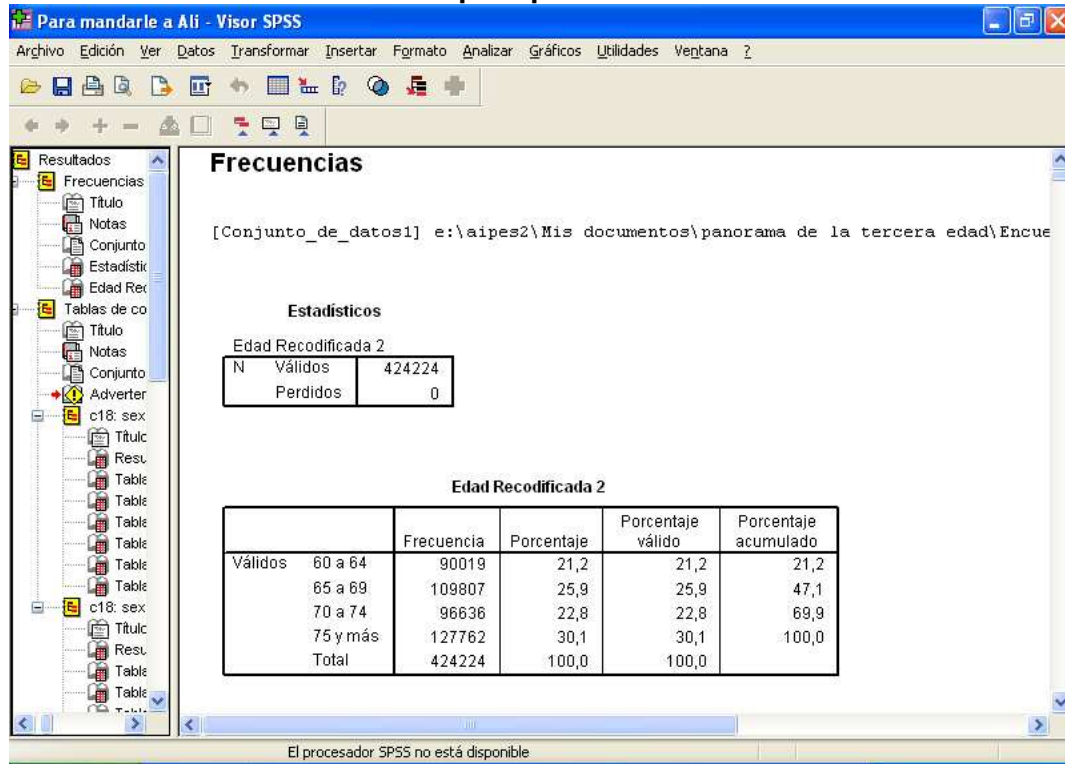
Con esta versión del programa se puede tener más de un archivo del mismo tipo -base, sintaxis y resultados- abierto al mismo tiempo.

Y si bien como ya mencionamos no es recomendable trabajar con más de un archivo a la vez, el SPSS distingue entre **Ventanas Activas** (con las cuales estamos trabajando en este momento) y **Ventanas Inactivas** (que se encuentran abiertas pero inactivas).

La forma de identificar cuando una ventana se encuentra **Activa** es a través del símbolo **+**



## EDITOR DE RESULTADOS: Output.spo



Los resultados estadísticos, tablas y gráficos se muestran en el **Editor de resultados**. Estos pueden editarse y guardarse para utilizarlos posteriormente. **La ventana** del Editor de resultados **se abre** automáticamente la primera vez que se ejecuta un procedimiento que genera resultados.

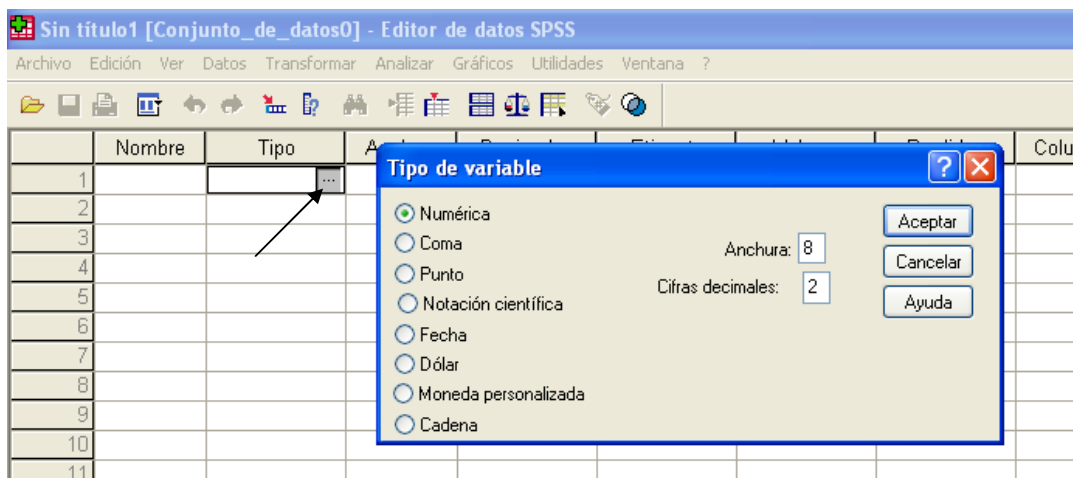
### Modificación de los gráficos y cuadros

- Se pueden modificar los gráficos de alta resolución y los diagramas en las ventanas de gráficos. Se pueden cambiar los colores, seleccionar diferentes tipos de fuentes y tamaños, intercambiar los ejes horizontal y vertical, rotar diagramas de dispersión 3-D e incluso cambiar el tipo de gráfico.
- Para esto es necesario hacer **doble click** con el mouse **sobre** el gráfico o cuadro que se desee modificar, habilitando de ese modo otra ventana que contiene un menú con las opciones correspondientes.

## VISTA DE VARIABLES:



## DEFINICIÓN DE VARIABLES

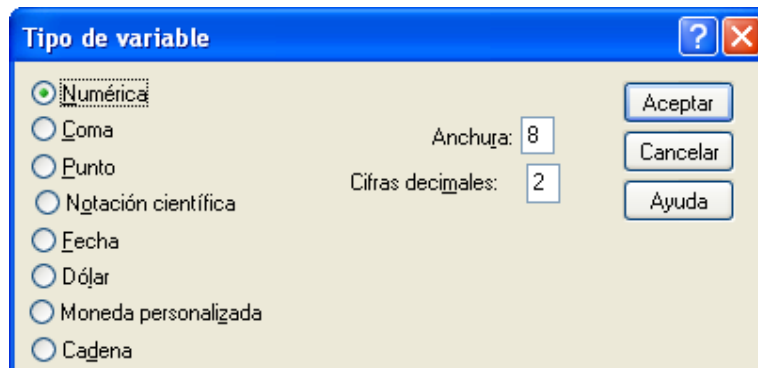


Se pueden definir variables nuevas o cambiar las definiciones de variables existentes. La información sobre la definición de datos incluye:

- **Nombre de variable** (máximo 8 caracteres en versiones antiguas)
- **Tipo de datos** (numérico, de cadena, fecha, etc.)
- **Etiquetas** descriptivas de variable y de valor
- **Códigos** especiales para valores perdidos

## A los nombres de variable se les aplican las siguientes reglas:

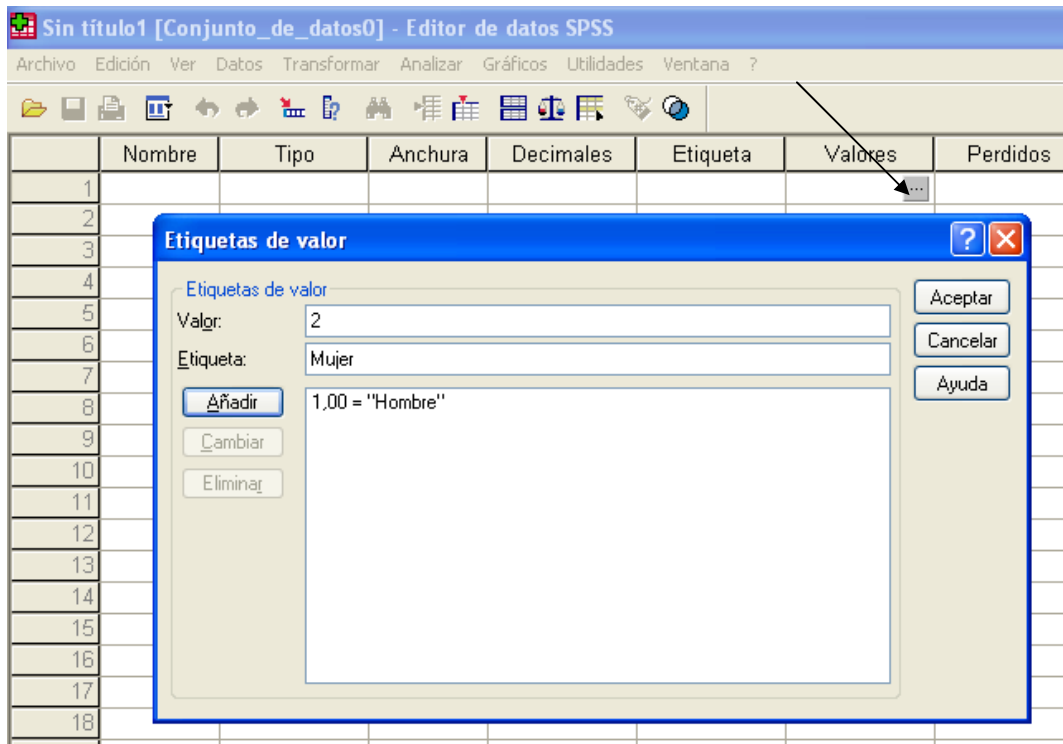
1. El nombre debe comenzar por una letra. El resto de los caracteres pueden ser letras, dígitos, puntos o los símbolos @, #, \_ o \$.
1. Los nombres de variable no pueden terminar en punto. Los espacios en blanco y los caracteres especiales (por ejemplo, !, ?, ' y \*) no se pueden utilizar.
1. La longitud del nombre no debe tener más de ocho caracteres (en versiones antiguas). Cada nombre de variable debe ser exclusivo; no se permiten duplicados. Los nombres de variable no distinguen a las mayúsculas y las minúsculas.
4. Las siguientes palabras claves reservadas no se pueden utilizar como nombres de variable: ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO, WITH



**Variables numéricas:** Son aquellas que admiten cualquier número como valor válido, pudiendo ir precedido del signo de “más” o de “menos”. Su longitud máxima es de 40 caracteres, de los cuales 16 pueden ser cifras decimales. La coma debe teclearse siempre que haya parte decimal.

**Variables alfanuméricas:** Admiten en su codificación cualquier carácter. En su definición debe especificarse únicamente su longitud máxima.

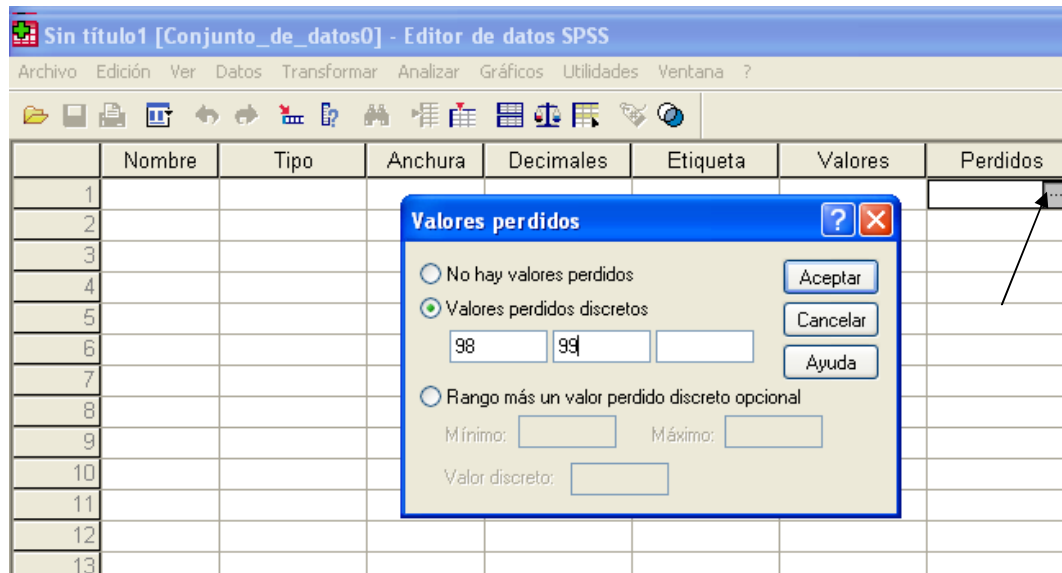
## Etiquetas y valores



Aunque los nombres de variable pueden ser sólo de 8 caracteres, las etiquetas de variable (**Var label**) pueden tener hasta 256 caracteres. Estas etiquetas descriptivas aparecen en los cuadros o gráficos de los resultados.

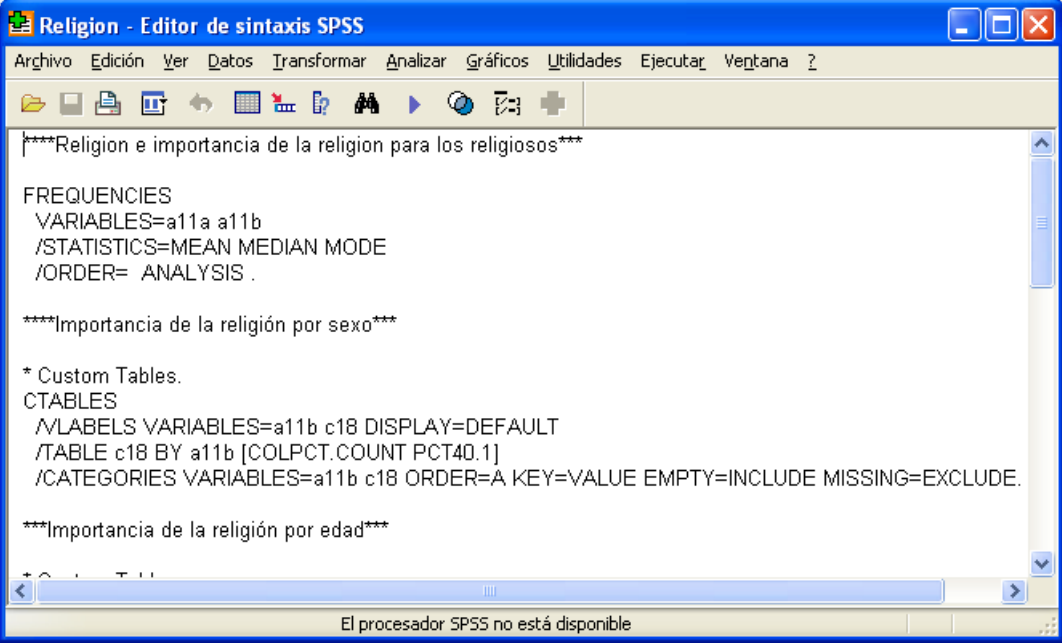
Asimismo se pueden asignar etiquetas descriptivas a cada valor de una variable (**Value label**). Esto es particularmente útil si el archivo de datos utiliza códigos numéricos para representar categorías no numéricas (Ej: *códigos 1 y 2 para valores hombre y mujer en variable Sexo*). Las etiquetas de valor pueden tener hasta 60 caracteres pero no están disponibles para variables de cadena larga (*variables de cadena de más de 8 caracteres*).

## Valores perdidos



Esta opción (**Missing**) define los valores de datos especificados como perdidos por el usuario. A menudo es útil para saber por qué se pierde información. Por ejemplo, puede desear distinguir entre datos perdidos porque un entrevistado se niega a responder y entre datos perdidos porque la pregunta no afectaba a dicho entrevistado. Los valores de datos especificados como perdidos por el usuario aparecen marcados para un tratamiento especial y se excluyen de la mayoría de los cálculos.

## EDITOR DE SINTAXIS: Otra forma de trabajar



```
Religion - Editor de sintaxis SPSS
Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ejecutar Ventana ?

|****Religion e importancia de la religion para los religiosos****

FRECUENCIAS
  VARIABLES=a11a a11b
  /STATISTICS=MEAN MEDIAN MODE
  /ORDER= ANALYSIS .

****Importancia de la religión por sexo****

* Custom Tables.
CTABLES
  /VARIABLES=a11b c18 DISPLAY=DEFAULT
  /TABLE c18 BY a11b [COLPCT.COUNT PCT40.1]
  /CATEGORIES VARIABLES=a11b c18 ORDER=A KEY=VALUE EMPTY=INCLUDE MISSING=EXCLUDE.

***Importancia de la religión por edad****

El procesador SPSS no está disponible
```

Existe otra forma -complementaria- para trabajar con el SPSS: a través de la redacción de **sintaxis**. La sintaxis, en pocas palabras, es el **lenguaje** con el cual trabaja el SPSS.

Si bien esta modalidad requiere mayores conocimientos y familiarización con el programa (por lo cual en este momento no es lo más recomendable) esta opción tiene la ventaja de **respaldar los procesos** y no sólo los resultados. Permitiendo así repetirlos y agilizar la tarea al trabajar con una cantidad importante y compleja de datos.