

Práctica 3.

Cálculo de estadísticos de tendencia central y dispersión a partir de ficheros de encuesta.

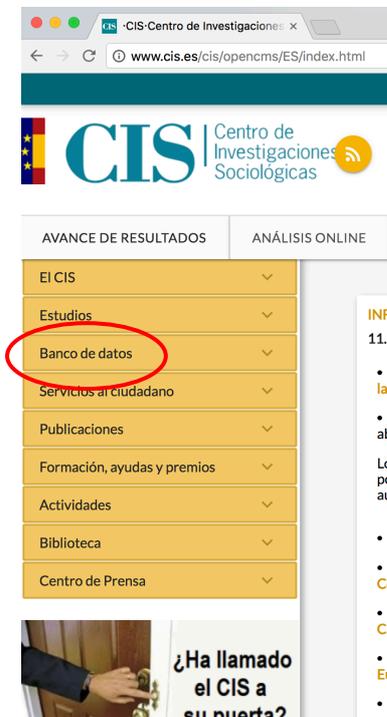
En esta práctica se obtendrán estadísticos de resumen de variables de intervalo procedentes de encuestas de opinión. La práctica tiene dos fases, en la primera se explicará cómo descargar un fichero con microdatos de una encuesta del CIS mientras que en la segunda parte se calcularán e interpretarán los principales estadísticos de tendencia central y de dispersión. La realización de esta práctica resulta central para conocer los instrumentos básicos para el tratamiento de grandes cantidades de información. En concreto analizaremos algunas variables del Barómetro de Opinión de Enero de 2012 (Estudio CIS 2927).

Primera Parte: Elaboración del fichero activo

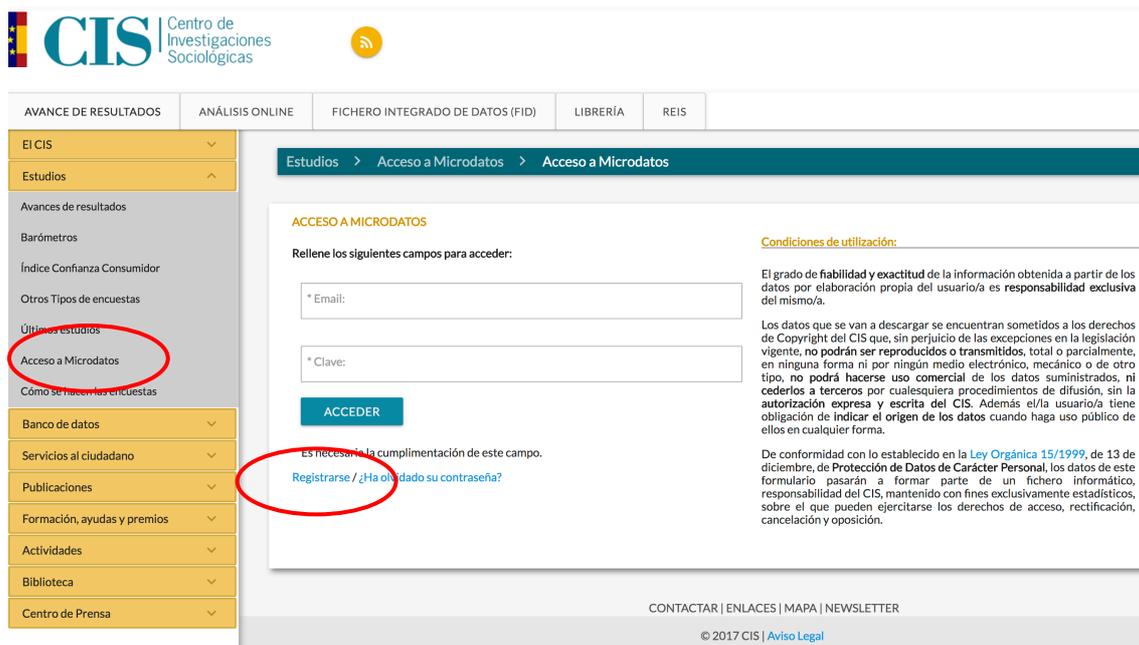
En la práctica anterior se explicó como grabar manualmente los datos de una variable. Ahora nos situaremos en el caso en que los datos procedentes de los cuestionarios ya han sido grabados en soporte informático. Por una parte tenemos los datos de los cuestionarios grabados en un fichero, en nuestro caso será el fichero DA2927, y por otra parte tenemos el programa para realizar los cálculos: el PSPP. El fichero de datos brutos tiene que convertirse en un fichero de trabajo, que será el que lea el programa PSPP. Este fichero de trabajo se conoce con el nombre de fichero activo –tiene extensión SAV-, y lo visualizaremos en el formato de hoja de cálculo. Para realizar dicha operación -la conversión del fichero bruto de datos en un fichero de trabajo- necesitamos un fichero de sintaxis que contiene las instrucciones para la lectura de los datos. El fichero denominado de “sintaxis” que alberga el conjunto de instrucciones que permite localizar los datos y especificar las posiciones de las variables. La ejecución de dichas instrucciones permitirá la generación de un fichero de trabajo, sobre el que haremos las distintas demandas de información. De forma esquemática podemos representar el funcionamiento de los programas de análisis de encuestas:

a) Obtención y preparación de los ficheros

Comenzaremos obteniendo los datos desde el Banco de Datos del CIS. Desde la dirección www.cis.es podemos llegar a localizar los datos desde diversos lugares (Estudios, Barómetros...). En este caso recomendamos el acceso por **Banco de Datos**.



Dentro de **Banco de datos**, vamos **Acceso a Microdatos**.



Deberemos cumplimentar el registro y una vez recibida la confirmación por mail, podremos autenticarnos en la página.

Estudios > Acceso a Microdatos > Registro

REGISTRO

Rellene los siguientes campos para acceder:

Nombre:

Apellidos:

Univ./Centro/ Empresa:

Correo electrónico:

Acepta las condiciones de uso

Es necesaria la cumplimentación de este campo.

Condiciones de utilización:

El grado de **fiabilidad y exactitud** de la información obtenida a partir de los datos por elaboración propia del usuario/a es **responsabilidad exclusiva** del mismo/a.

Los datos que se van a descargar se encuentran sometidos a los derechos de Copyright del CIS que, sin perjuicio de las excepciones en la legislación vigente, **no podrán ser reproducidos o transmitidos**, total o parcialmente, en ninguna forma ni por ningún medio electrónico, mecánico o de otro tipo, **no podrá hacerse uso comercial** de los datos suministrados, ni **cederlos a terceros** por cualesquiera procedimientos de difusión, sin la **autorización expresa y escrita** del CIS. Además el/la usuario/a tiene obligación de **indicar el origen de los datos** cuando haga uso público de ellos en cualquier forma.

De conformidad con lo establecido en la **Ley Orgánica 15/1999**, de 13 de diciembre, de **Protección de Datos de Carácter Personal**, los datos de este formulario pasarán a formar parte de un fichero informático, responsabilidad del CIS, mantenido con fines exclusivamente estadísticos, sobre el que pueden ejercitarse los derechos de acceso, rectificación, cancelación y oposición.

Una vez autenticados. El siguiente paso es ir a la zona de **Banco de Datos** y abrir **Catálogo de Encuestas**. En dicha página seleccionaremos **Índice anual**, y en los correspondientes al año 2012 buscaremos el **Barómetro Enero 2012**.

04/01/12	2927	BARÓMETRO DE ENERO 2012	<ul style="list-style-type: none"> Resultados_PDF Resultados_HTML Cruces Ficha_técnica Cuestionario Fichero_datos 	Añadir
----------	------	-------------------------	---	--------

Una vez localizado el estudio en el Banco de Datos, seleccionamos: **Añadir**
 En la parte superior de la página aparecerá al lado de nuestra identidad la posibilidad de **Ver**.
 Accediendo a esta herramienta podremos comenzar la descarga de los datos.

Banco de datos > Catálogo de encuestas > Índice anual

Bienvenido/a: Luis Camarero Estudios seleccionados: **1 Ver**

[Cerrar Sesión](#)

Para descargar el fichero el CIS nos solicitará nuestra conformidad de las condiciones de utilización. Deberemos indicar nuestros datos, marcar la casilla de aceptación y pulsar en **Fichero_datos**. (De momento el CIS sólo garantiza la descarga para ciertos navegadores).

Estudios > Acceso a Microdatos > Estudios seleccionados

Bienvenido/a: Luis Camarero Estudios seleccionados: 1 [Ver](#)

[Cerrar Sesión](#)

ESTUDIOS SELECCIONADOS

Estos son los estudios que ha seleccionado

Fecha	Número	Título	Ficheros	Acción
01/02/12	2932	BARÓMETRO DE FEBRERO 2012	<ul style="list-style-type: none"> Resultados_PDF Resultados_HTML Cruces Ficha_técnica Cuestionario Fichero_datos 	Eliminar

Motivo de uso

Investigación, informe o publicación
 Trabajo en Curso
 Docencia
 Tesis, Tesina, Master
 Medios de comunicación
 Proyectos, planes o campañas
 Conferencias, ponencias, seminarios
 Otros

Acepta las condiciones de uso

[DESCARGAR TODOS](#)

[Contactar](#) | [Enlaces](#) | [Mapa](#) | [Newsletter](#) |
 © 2017 CIS | [Aviso Legal](#)

El fichero que descargamos es un fichero comprimido MD2927.ZIP. La mejor opción es que guardemos esta carpeta en un dispositivo USB, de forma que conozcamos claramente el “*path*” o lugar en el que está localizable dicha carpeta.

El “*path*” literalmente inglés es el camino nos indica la ruta de acceso física a los ficheros. Los discos duro de los ordenadores suelen denominarse con la letra C seguida de dos puntos. Así “C:” significa disco duro C. Los dispositivos USB utilizan letras por lo general desde la E hasta la K.

Después de la unidad se indica la carpeta o fichero –siempre utilizando un *slash* o barra inversa “\” como separador, por ejemplo si grabamos nuestro fichero en una memoria flash denominada K, el *path* será: **K:\MD2927.ZIP**

Al abrir la carpeta MD2927 encontraremos los siguientes ficheros:

codigo2927.pdf	Es el libro de códigos del estudio. Muestra la posición de las variables (columnas) en el fichero de datos. También nos muestra los valores permitidos para cada una de las variables.
cues2927.pdf	Es la reproducción del cuestionario utilizado
DA2927	Son los datos de la encuesta grabados (en formato ASCII)
ES2927	Es el fichero de sintaxis para SPSS o PSPP. Básicamente es el libro de códigos convertido en comandos para la lectura del fichero de datos.
Ft2927.pdf	Es la Ficha Técnica del Estudio
Sas2927	Es el fichero de sintaxis para el programa SAS (Para nuestra práctica este fichero no tiene utilidad ninguna)
tarjetas2927.pdf	Es el fichero que contiene una reproducción de las tarjetas que utilizaron los encuestadores durante la realización de la entrevista.

Antes de continuar MD conviene que se familiarice con el cuestionario utilizado. Abra el cuestionario, seleccione alguna pregunta y busque su correspondencia en el libro de códigos.

Para empezar a trabajar vamos a trasladar el fichero de datos DA2927 y el fichero de sintaxis

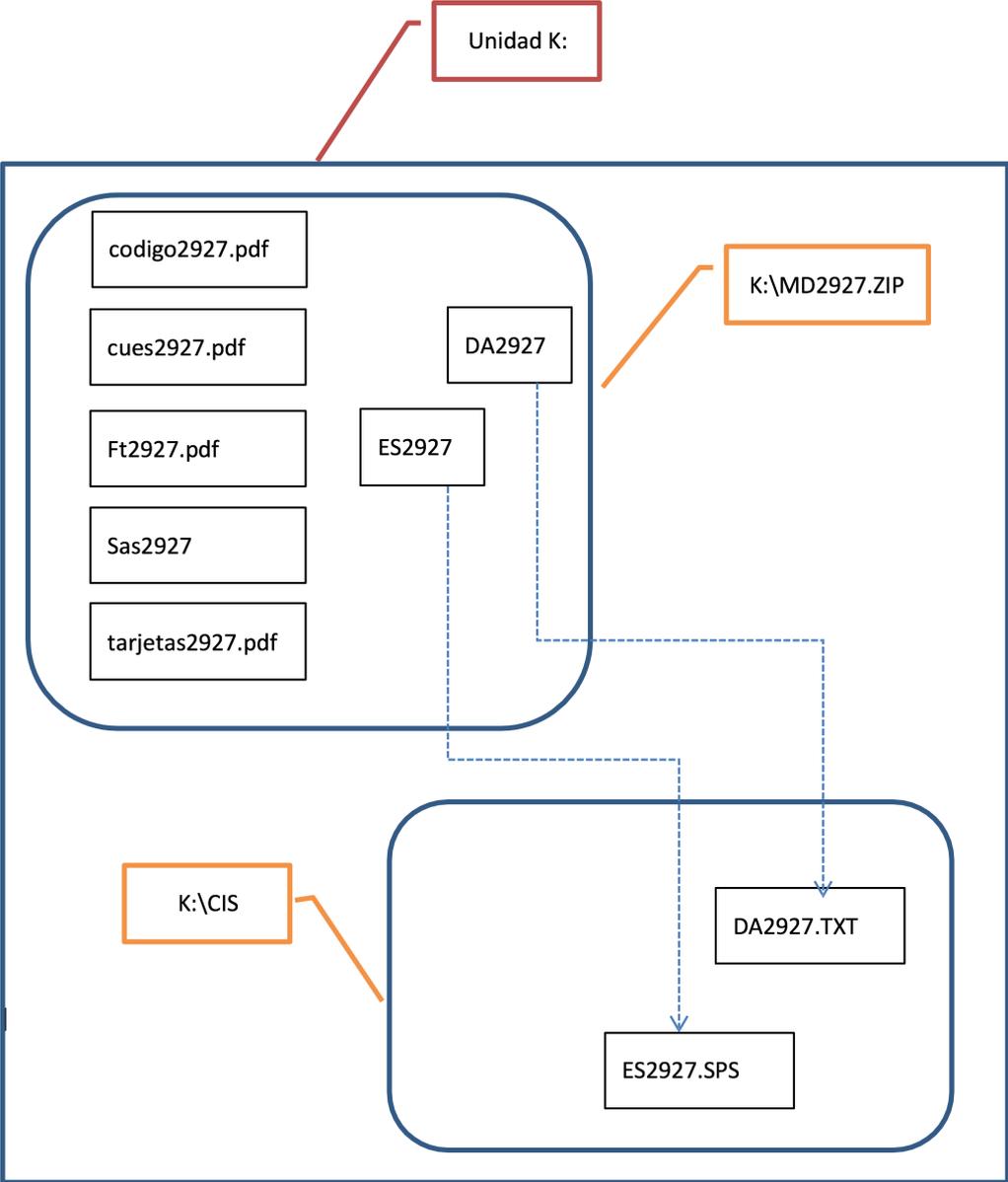
ES2927 a una nueva carpeta, alojada en el directorio raíz [K:]¹. La nueva carpeta la denominaremos CIS – por ejemplo- y en ella alojaremos los dos ficheros. Aprovecharemos también para añadir la extensión al nombre de los ficheros.

(En cada sistema operativo o versión de Windows varía el procedimiento por ello indicaremos los pasos de forma genérica):

- a) Crear en el directorio raíz de (por ejemplo en [K:]) una carpeta denominada CIS.
[K:\CIS]
- b) Trasladar el fichero DA2927 desde la carpeta [K:\DA2927.ZIP] a [K:\CIS]
- c) Trasladar el fichero ES2927 desde la carpeta [K:\DA2927.ZIP] a [K:\CIS]
- d) Renombrar el fichero [K:\CIS\DA2927] con la extensión TXT de forma que quede como DA2927.TXT situado como [K:\CIS\DA2927.TXT]
- e) Renombrar el fichero [K:\CIS\ES2927] con la extensión SPS de forma que quede como [K:\CIS\ES2927.SPS]

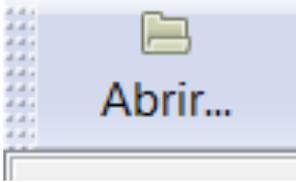
¹ El símbolo de los corchetes [] se utiliza para diferenciar dentro del texto las expresiones de denominación de los *paths* donde están alojados los ficheros

El siguiente esquema ayuda a comprender las operaciones realizadas:

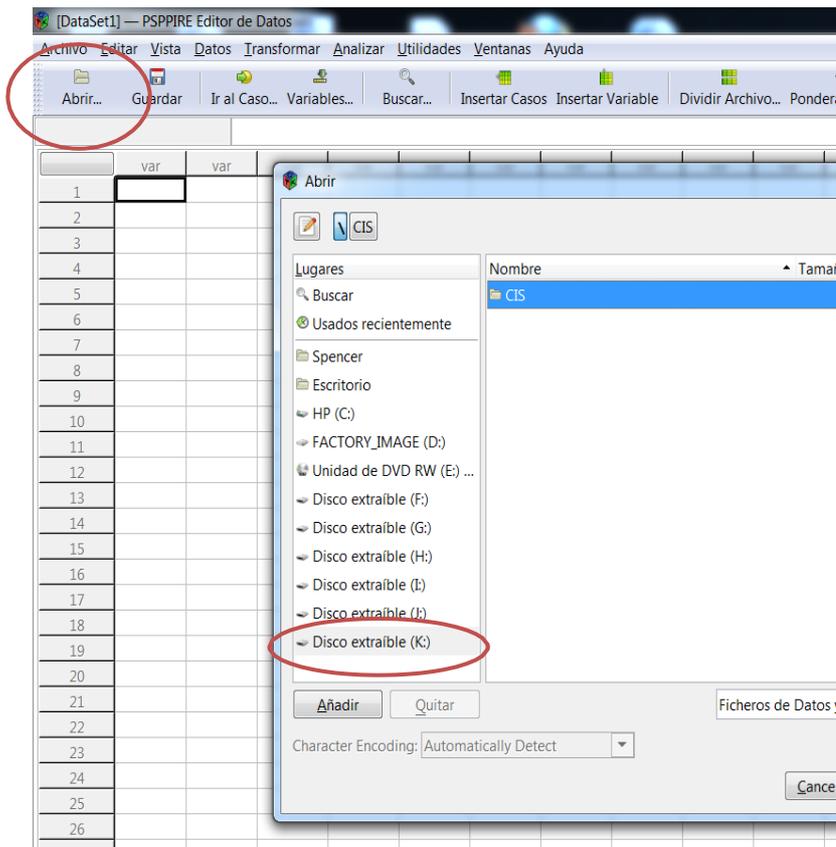


b) Generación del fichero activo

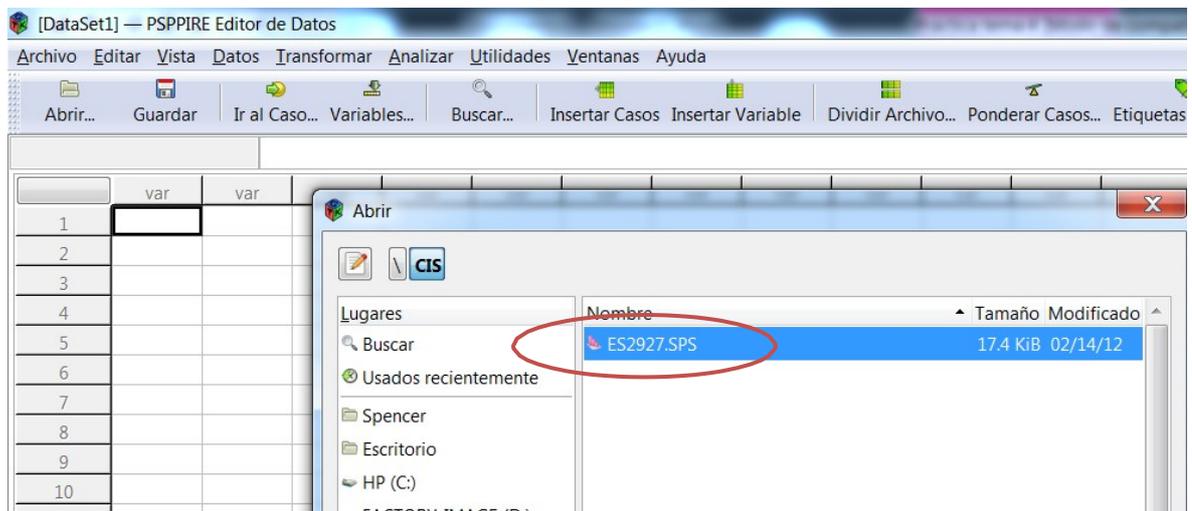
Una vez descargados y descomprimidos los ficheros necesarios, abriremos el programa PPSS. En primer lugar, llamaremos mediante el botón **Abrir...**



al fichero que contiene las instrucciones de sintaxis ES2927.SPS que está en: [K:\CIS\ES2927.SPS]



Navegaremos por los menús, seleccionando la unidad [K:], la carpeta [\CIS] y el fichero: ES2927.SPS



Después de seleccionar el fichero se abrirá una nueva ventana con las instrucciones que contiene el fichero de sintaxis.

```

*ES2927.sps — PSPPIRE Editor de sintaxis
Archivo  Editar  Ejecutar  Ventanas  Ayuda
1
2  DATA LIST FILE= 'DA2927'
3  /ESTU 1-4 CUES 5-9 CCAA 10-11 PROV 12-13 MUN 14-16 TAMUNI 17 AREA 18 DISTR 19-20 SECCION 21-23
4  ENTREV 24-27 P0 28 P0A 29-31 P1 32 P2 33 P3 34 P4 35 P5 36 P6 37 P701 38-39 P702 40-41 P703 42-43
5  P801 44-45 P802 46-47 P803 48-49 P901 50 P902 51 P903 52 P904 53 P905 54 P906 55 P907 56 P908 57
6  P909 58 P910 59 P10 60 P11 61 P12 62 P13 63 P1401 64 P1402 65 P1403 66 P1404 67 P1405 68 P1406 69
7  P1407 70 P1408 71 P1409 72 P1501 73 P1502 74 P1503 75 P1504 76 P1505 77 P1506 78 P1507 79 P1508 80
8  P16 81-82 P16A 83-84 P1701 85-86 P1702 87-88 P1703 89-90 P1704 91-92 P1705 93-94 P1706 95-96
9  P1707 97-98 P1708 99-100 P1709 101-102 P1710 103-104 P1711 105-106 P1712 107-108 P1713 109-110
10 P1714 111-112 P18 113 P1901 114 P1902 115 P1903 116 P1904 117 P1905 118 P1906 119 P20 120 P2101 121
11 P2102 122 P2103 123 P21A01 124 P21A02 125 P21A03 126 P2201 127-128 P2202 129-130 P2203 131-132
12 P2204 133-134 P2205 135-136 P2206 137-138 P2207 139-140 P2208 141-142 P2209 143-144 P2210 145-146
13 P2211 147-148 P2212 149-150 P2213 151-152 P23 153 P24 154 P25 155-156 P26 157-158 P27 159-160
14 P28 161 P28A 162-163 P29 164 P30 165-166 P31 167 P31A 168-169 P32 170 P32A 171 P33 172 P34 173
15 P35 174 P36 175 P37 176 P37A 177 P37B 178 P38 179-181 P39 182 P39A 183 P40 184-185 P41 186 P41A 187
16 P41B 188 P42 189 P4301 190 P4302 191 P4303 192 P4304 193 P4401 194-196 P4402 197-199 P4403 200-202
17 P4404 203-205 P4405 206-208 P4501 209-211 P4502 212-214 P4503 215-217 P4504 218-220 P4505 221-223
18 P46 224 P47 225 P48 226 P49 227 I1 228-230 I2 231-233 I3 234-236 I4 237-239 I5 240-242 I6 243-245
19 I7 246-248 I8 249-251 I9 252-254 E101 255-256 E102 257-258 E103 259-260 E2 261 E3 262-264 E4 265
20 C1 266 C1A 267-268 C2 269 C2A 270 C2B 271-272 C3 273 C4 274-275 P26A 276-277 VOTOSIM 278-279
21 RECUERDO 280-281 ESTUDIOS 282 OCUPA 283-284 CONDICION 285-286 ESTATUS 287 .
22
23 VARIABLE LABELS
24 ESTU 'Estudio'
25 /CCAA 'Comunidad autónoma'
26 /PROV 'Provincia'
27 /MUN 'Municipio'
28 /TAMUNJ 'Tamaño de municipio'
29 /AREA 'Área metropolitana'
30 /P0 'Nacionalidad del entrevistado'
31 /P0A 'Nacionalidad de la persona entrevistada (extranjeros)!'
32 /P1 'Valoración de la situación económica general de España'
33 /P2 'Valoración retrospectiva de la situación económica de España (1 año)'
34 /P3 'Valoración prospectiva de la situación económica de España (1 año)'
35 /P4 'Valoración de la situación política general de España'
36 /P5 'Valoración retrospectiva de la situación política de España (1 año)'
37 /P6 'Valoración prospectiva de la situación política de España (1 año)'
38 /P701 'Problemas principales que existen actualmente en España (3) (ordinal)'
39 /P702 'Problemas principales que existen actualmente en España (3) (ordinal)'
40 /P703 'Problemas principales que existen actualmente en España (3) (ordinal)'
41 /P801 'Problemas sociales que personalmente afectan más (3)'
42 /P802 'Problemas sociales que personalmente afectan más (3)'
43 /P803 'Problemas sociales que personalmente afectan más (3)'
44 /P901 'El paro'
45 /P902 'La inseguridad ciudadana'
46 /P903 'El terrorismo'
47 /P904 'Las drogas'
48 /P905 'El medio ambiente'

```

Si nos desplazamos por el fichero vamos a encontrar seis instrucciones distintas. Las instrucciones comienzan con las siguientes expresiones y terminan con “.” (un punto):

Línea 2	DATA LIST
Línea 23	VARIABLE LABELS
Línea 177	VALUE LABELS
Línea 721 y 726	EXECUTE
Línea 724	MISSING VALUES
Línea 727	FREQUENCIES

Instrucción **DATA LIST**

La primera instrucción DATA LIST sirve para indicar al programa la posición “path” del fichero que contiene los datos, y la posición de cada una de los datos en el fichero.

Lo primero que vamos a hacer es indicar el *path* correcto, que recordemos que es [K:\CIS\DA2927.TXT]

```

*ES2927.sps — PSPPIRE Editor de sintaxis
Archivo  Editar  Ejecutar  Ventanas  Ayuda
1
2  DATA LIST FILE='K:\CIS\DA2927.TXT'
3  /ESTU 1-4 CUES 5-9 CCAA 10-11 PROV 12-13 MUN 14-16 TAMUNI 17 AREA 18 DISTR 19-20 SECCION 21-23
4  ENTREV 24-27 P0 28 P0A 29-31 P1 32 P2 33 P3 34 P4 35 P5 36 P6 37 P701 38-39 P702 40-41 P703 42-43
5  P801 44-45 P802 46-47 P803 48-49 P901 50 P902 51 P903 52 P904 53 P905 54 P906 55 P907 56 P908 57
6  P909 58 P910 59 P10 60 P11 61 P12 62 P13 63 P1401 64 P1402 65 P1403 66 P1404 67 P1405 68 P1406 69
7  P1407 70 P1408 71 P1409 72 P1501 73 P1502 74 P1503 75 P1504 76 P1505 77 P1506 78 P1507 79 P1508 80
8  P16 81-82 P16A 83-84 P1701 85-86 P1702 87-88 P1703 89-90 P1704 91-92 P1705 93-94 P1706 95-96
9  P1707 97-98 P1708 99-100 P1709 101-102 P1710 103-104 P1711 105-106 P1712 107-108 P1713 109-110
10 P1714 111-112 P18 113 P1901 114 P1902 115 P1903 116 P1904 117 P1905 118 P1906 119 P20 120 P2101 121
11 P2102 122 P2103 123 P21A01 124 P21A02 125 P21A03 126 P2201 127-128 P2202 129-130 P2203 131-132
12 P2204 133-134 P2205 135-136 P2206 137-138 P2207 139-140 P2208 141-142 P2209 143-144 P2210 145-146
13 P2211 147-148 P2212 149-150 P2213 151-152 P23 153 P24 154 P25 155-156 P26 157-158 P27 159-160
14 P28 161 P28A 162-163 P29 164 P30 165-166 P31 167 P31A 168-169 P32 170 P32A 171 P33 172 P34 173
15 P35 174 P36 175 P37 176 P37A 177 P37B 178 P38 179-181 P39 182 P39A 183 P40 184-185 P41 186 P41A 187
16 P41B 188 P42 189 P4301 190 P4302 191 P4303 192 P4304 193 P4401 194-196 P4402 197-199 P4403 200-202
17 P4404 203-205 P4405 206-208 P4501 209-211 P4502 212-214 P4503 215-217 P4504 218-220 P4505 221-223
18 P46 224 P47 225 P48 226 P49 227 11 228-230 12 231-233 13 234-236 14 237-239 15 240-242 16 243-245
19 I7 246-248 18 249-251 19 252-254 E101 255-256 E102 257-258 E103 259-260 E2 261 E3 262-264 E4 265
20 C1 266 C1A 267-268 C2 269 C2A 270 C2B 271-272 C3 273 C4 274-275 P26A 276-277 VOTOSIM 278-279
21 RECUERDO 280-281 ESTUDIOS 282 OCUPA 283-284 CONDICION 285-286 ESTATUS 287 .
22

```

La expresión **DATA LIST FILE='K:\CIS\DA2927.TXT'** nos indica el lugar donde se alojan nuestros datos. Obsérvese que el *path* va indicado entre comillas. (Se pueden utilizar las dobles comillas " " o también como es habitual en inglés, las simples.)

La barra "/" es el comienzo de la segunda parte del comando que está destinado a nombrar las variables e indicar su posición. Si se consulta el fichero **cues2927.pdf** que contiene el cuestionario podemos interpretar fácilmente las siguientes líneas. Recuérdese que en el cuestionario se indican las columnas de las variables entre paréntesis.

La expresión **ESTU 1-4**, indica que las cuatro primeras columnas del fichero de datos se refieren al número de estudio. **CUES 5-9** significa que el número de cuestionario, lo vamos a llamar CUES y dicho dato ocupan las columnas 5 a 9. (Tiene una longitud de 5 dígitos)



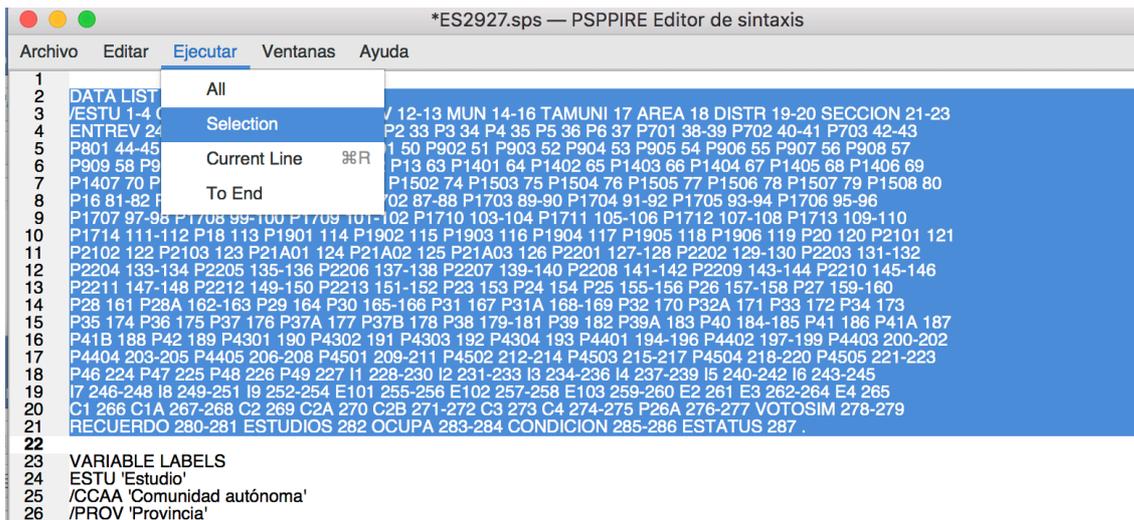
CIS
Centro de Investigaciones Sociológicas

Departamento de Investigación

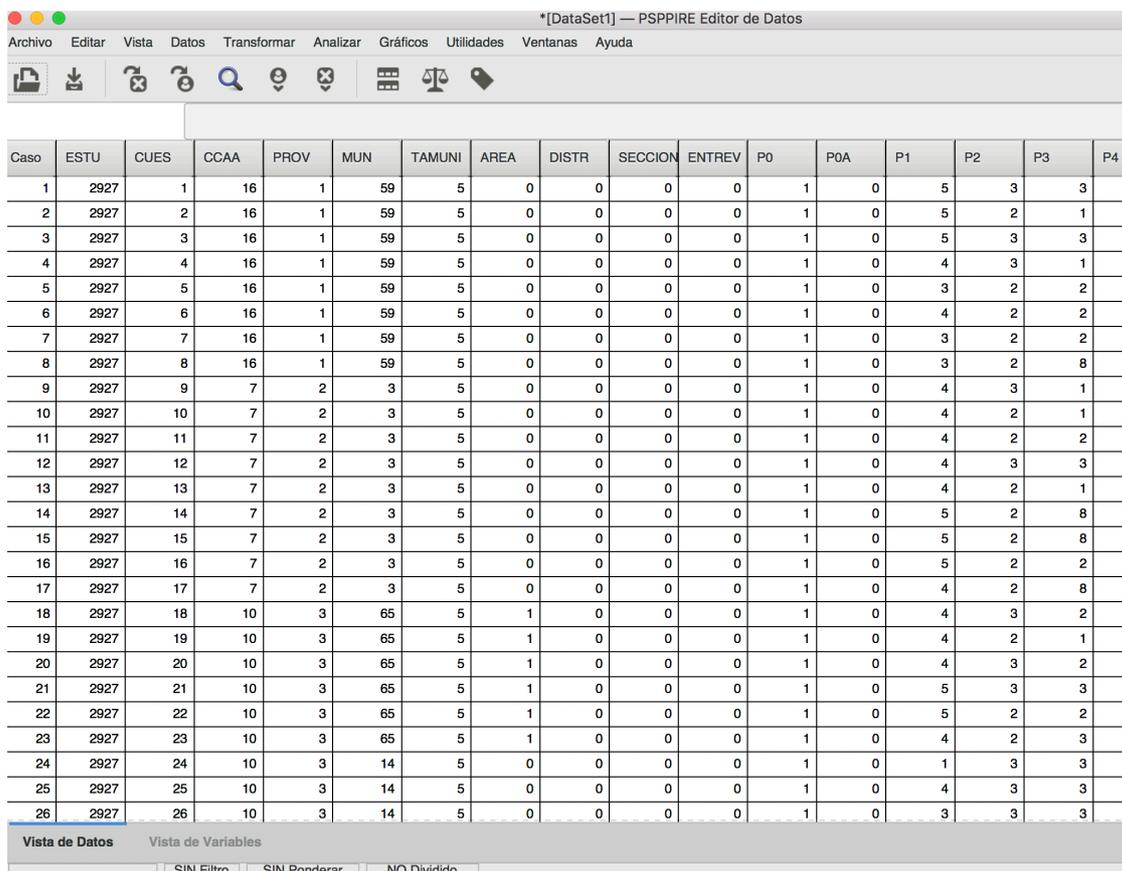
ENERO 2012

Comunidad Autónoma _____	<input type="text"/>	(10)(11)	Nº ESTUDIO	Nº CUESTIONARIO
Provincia _____	<input type="text"/>	(12)(13)	<input type="text" value="2 9 2 7"/>	<input type="text"/>
Municipio _____ (nombre municipio)	<input type="text"/>	(14)(15)(16)	(1)(2)(3)(4)	(5)(6)(7)(8)(9)
Tamaño de hábitat _____	<input type="text"/>	(17)(18)		
Distrito _____	<input type="text"/>	(19)(20)		
Sección _____	<input type="text"/>	(21)(22)(23)		
Entrevistador/a _____	<input type="text"/>	(24)(25)(26)(27)		

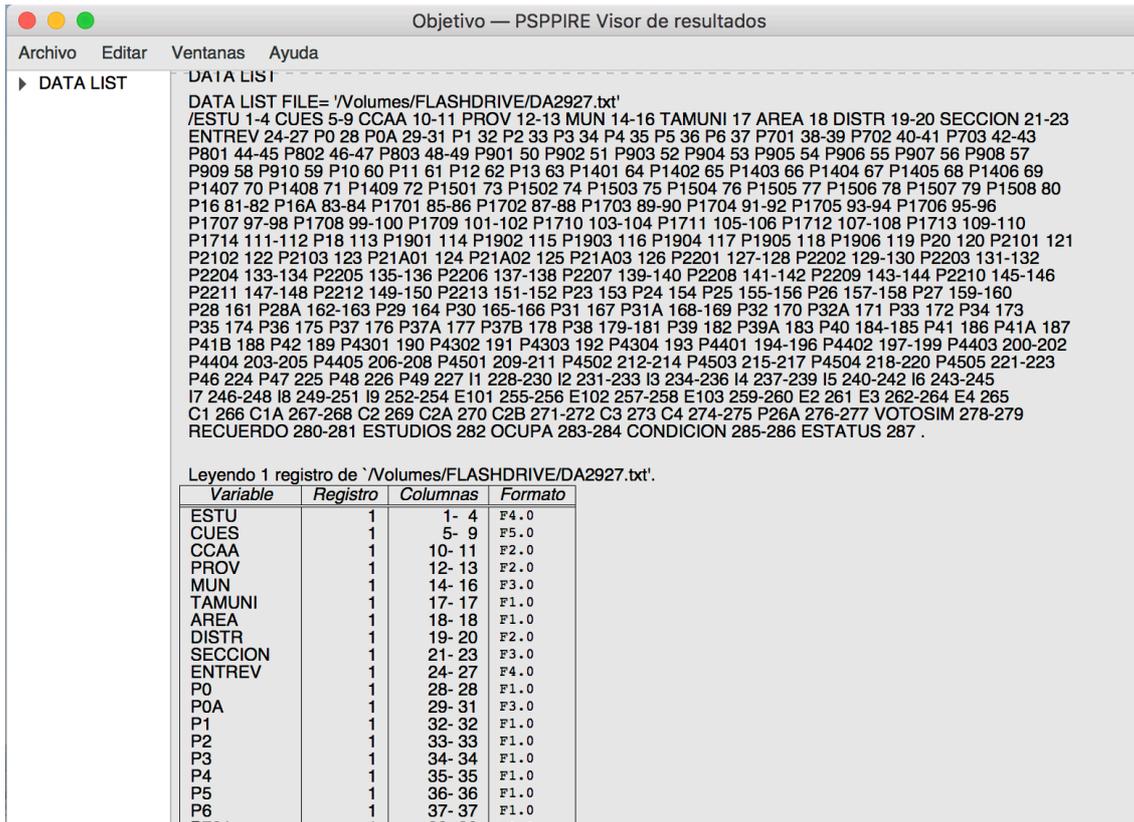
La expresión **CCAA 10-11 PROV 12-13** define dos variables: CCAA indica la Comunidad Autónoma y PROV la provincia. Los datos de CCAA están en las columnas 10 y 11, los de PROV en 12 y 13.



La ejecución del comando abrirá una ventana con la matriz de datos:



Y también una tercera ventana de resultados, en la que se van registrando las ordenes de la sesión y los resultados de la misma. En este caso ofrece una tabla con el nombre de las variables y las columnas que ocupan así como el formato. Por ejemplo F4.0 señala que es una variable numérica con cinco dígitos y sin decimales.



Si vamos a la pantalla de la rejilla de datos podemos ver que el primer cuestionario del fichero es el 1, pertenece a la Comunidad Autónoma 16, provincia 1. Si consultamos el libro de códigos, fichero **codigo2927.pdf**, podemos comprobar que es una entrevista realizada en Álava. A la pregunta número 1 responde con 5, es decir que considera la situación económica de España como Muy Mala.

Resulta evidente, que la lectura de los datos mediante los códigos resulta engorrosa. Para ello tenemos dos comandos que nos facilitan dicha labor. Si volvemos al fichero de sintaxis, podemos observar otros comandos (Variable Label y Value Label):

```
VARIABLE LABEL
  ESTU 'Estudio'
  /CCAA 'Comunidad autónoma'
  /PROV 'Provincia'
  /TAMUNI 'Tamaño de hábitat'
  /AREA 'Área metropolitana'
  /MUN 'Municipio'
  /p901 'El paro'
  /p902 'La inseguridad ciudadana'
  /p903 'El terrorismo'
  /p904 'Las drogas'
  /p905 'El medio ambiente'
  /p906 'La asistencia sanitaria'
  /p907 'La vivienda'
  /p908 'La inmigración'
  /p909 'La educación'

...

/RECUERDO 'Recuerdo de voto'
/ESTUDIOS 'Estudios del entrevistado'
/OCUPA 'Ocupación del entrevistado'
/CONDICION 'Condición socioeconómica'
/ESTATUS 'Estatus socioeconómico'.
```

VARIABLE LABEL, que nos rotula a las variables. Por ejemplo, en los resultados en vez de que CCAA aparecerá “Comunidad Autónoma” o “Provincia en vez de PROV. La sintaxis de dicho comando es simple. Nombre de la variable y entre comillas el rótulo, las variables se separan con la barra “/”. (Nótese como es habitual que el comando termina con un punto).

VALUE LABEL CCAA

01'Andalucía' 02'Aragón' 03'Asturias (Principado de)'
04'Baleares (Islas)' 05'Canarias' 06'Castabria'
07'Castilla La Mancha' 08'Castilla y León'
09'Cataluña' 10'Comunidad Valenciana' 11'Extremadura'
12'Galicia' 13'Madrid (Comunidad de)' 14'Murcia (Región de)'
15'Navarra (Comunidad Foral de)' 16'País Vasco' 17'Rioja (La)'
/PROV

01'Álava' 02'Albacete' 03'Alicante' 04'Almería' 33'Asturias'
05'Ávila' 06'Badajoz' 07'Baleares' 08'Barcelona' 09'Burgos'
10'Cáceres' 11'Cádiz' 39'Castabria' 12'Castellón de la Plana'
13'Ciudad Real' 14'Córdoba' 15'Coruña (A)' 16'Cuenca'
17'Girona' 18'Granada' 19'Guadalajara' 20'Guipúzcoa'
21'Huelva' 22'Huesca' 23'Jaén' 24'León' 25'Lleida' 27'Lugo' 28'Madrid'
29'Málaga' 30'Murcia' 31'Navarra' 32'Ourense' 34'Palencia'
35'Palmas (Las)' 36'Pontevedra' 26'Rioja (La)' 37'Salamanca'
38'Santa Cruz de Tenerife' 40'Segovia' 41'Sevilla' 42'Soria'
43'Tarragona' 44'Teruel' 45'Toledo' 46'Valencia' 47'Valladolid'
48'Vizcaya' 49'Zamora' 50'Zaragoza'

...

/CONDICION

1 'Directores y profesionales'
2 'Técnicos y cuadro medios'
3 'Pequeños empresarios'
4 'Agricultores'
5 'Empleados de oficinas y servicios'
6 'Obreros cualificados'
7 'Obreros no cualificados'
8 'Jubilados y pensionistas'
9 'Parados'
10 'Estudiantes'
11 'Trabajo doméstico no remunerado'
12 'No clasificables'

/ESTATUS

1 'Clase alta/ media-alta'
2 'Nuevas clases medias'
3 'Viejas clases medias'
4 'Obreros cualificados'
5 'Obreros no cualificados'
9 'No consta'.

EXECUTE.

VALUE LABEL, que nos rotula los valores de las variables. La sintaxis es también sencilla, después de cada nombre de variable se van colocando los distintos valores con sus rótulos entre comillas.

Después de la ejecución de ambos comandos –marcaremos con el ratón las líneas de cada uno de los comandos y en **Ejecutar** pulsaremos la opción **Selección**, nos encontramos con la siguiente rejilla de datos, que en vez de códigos muestra las etiquetas.

En la selección de los comandos a ejecutar (Variable y Value Labels) podemos incluir también el Comando:

EXECUTE.

Dicho comando, indica a PSPP que ejecute inmediatamente la orden. (Si no lo incluimos, las órdenes quedan en espera hasta que indiquemos “Ejecutar transformaciones pendientes”: En matriz de datos: Transformar-Ejecutar transformaciones pendientes).

Una vez procesadas las órdenes veremos en la barra superior de la matriz de datos el botón de

Etiquetas de Valor que nos permitirá mostrar bien los códigos, bien las etiquetas de valor.



Caso	ESTU	CUES	CCAA	PROV	MUN	TAMUNI	AREA	DISTR	SECCION	ENTREV	P0	P0A	P1	P2	P3	P4
1	2927	1	País Vasco	Álava	Burgos/Vit	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Muy mala	Peor	Peor	Muy mala
2	2927	2	País Vasco	Álava	Burgos/Vit	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Muy mala	Igual	Mejor	Mala
3	2927	3	País Vasco	Álava	Burgos/Vit	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Muy mala	Peor	Peor	Mala
4	2927	4	País Vasco	Álava	Burgos/Vit	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Mala	Peor	Mejor	Regular
5	2927	5	País Vasco	Álava	Burgos/Vit	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Regular	Igual	Igual	Mala
6	2927	6	País Vasco	Álava	Burgos/Vit	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Mala	Igual	Igual	Mala
7	2927	7	País Vasco	Álava	Burgos/Vit	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Regular	Igual	Igual	Mala
8	2927	8	País Vasco	Álava	Burgos/Vit	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Regular	Igual	N.S.	Regular
9	2927	9	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Mala	Peor	Mejor	Buena
10	2927	10	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Mala	Igual	Mejor	Mala
11	2927	11	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Muy mala	Igual	Igual	Mala
12	2927	12	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Mala	Peor	Peor	Muy mala
13	2927	13	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Mala	Igual	Mejor	Regular
14	2927	14	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Muy mala	Igual	N.S.	Muy mala
15	2927	15	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Muy mala	Igual	N.S.	Mala
16	2927	16	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Muy mala	Igual	Igual	Mala
17	2927	17	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Mala	Igual	N.S.	Mala

Recuérdese que también pueden cambiarse las etiquetas de valor y de variable de forma manual, accediendo por la pestaña de **Vista de Variables**. Sin embargo para grandes conjuntos de datos, resulta más cómodo el uso de ficheros de sintaxis con las etiquetas.

12	2927	12	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza
13	2927	13	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza
14	2927	14	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza
15	2927	15	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza
16	2927	16	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza
17	2927	17	Castilla-La	Albacete	Albacete	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza
18	2927	18	Comunida	Alicante	Elche/Get	100.001 a	Pertenece	Anonimiza	Anonimiza	Anonimiza
19	2927	19	Comunida	Alicante	Elche/Get	100.001 a	Pertenece	Anonimiza	Anonimiza	Anonimiza
20	2927	20	Comunida	Alicante	Elche/Get	100.001 a	Pertenece	Anonimiza	Anonimiza	Anonimiza
21	2927	21	Comunida	Alicante	Elche/Get	100.001 a	Pertenece	Anonimiza	Anonimiza	Anonimiza
22	2927	22	Comunida	Alicante	Elche/Get	100.001 a	Pertenece	Anonimiza	Anonimiza	Anonimiza
23	2927	23	Comunida	Alicante	Elche/Get	100.001 a	Pertenece	Anonimiza	Anonimiza	Anonimiza
24	2927	24	Comunida	Alicante	Alicante	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza
25	2927	25	Comunida	Alicante	Alicante	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza
26	2927	26	Comunida	Alicante	Alicante	100.001 a	No pertene	Anonimiza	Anonimiza	Anonimiza

Vista de Datos **Vista de Variables**

SIN Filtro SIN Ponderar NO Dividido

Una vez que hemos ejecutado los comandos DATA LIST, VARIABLE LABEL y VALUE LABEL podemos generar nuestro fichero activo, es decir el fichero sobre el que pediremos los resultados. Para ello pulsamos sobre **Archivo** y seleccionamos la opción **Guardar como**.

*[DataSet1] — PSPPIRE Editor de Datos

Archivo Editar Vista Datos Transformar Analizar Gráficos Utilidades Ventanas Ayuda

- New
- Open
- Import Data...
- Guardar...
- Guardar como...**
- Renombrar Archivo de Datos...
- Display Data File Information
- Recently Used Data
- Archivos utilizados recientemente
- Quit

	MUN	TAMUNI	AREA	DISTR	SECCION	ENTREV	P0	P0A
	Burgos/Vit	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No pi
	Burgos/Vit	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No pi
	Burgos/Vit	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No pi
	Burgos/Vit	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No pi
	Burgos/Vit	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No pi
7	2927	7	País Vasco	Álava	Burgos/Vit	100.001 a	No pertenc	Anonimiza
8	2927	8	País Vasco	Álava	Burgos/Vit	100.001 a	No pertenc	Anonimiza
9	2927	9	Castilla-La	Albacete	Albacete	100.001 a	No pertenc	Anonimiza
10	2927	10	Castilla-La	Albacete	Albacete	100.001 a	No pertenc	Anonimiza
11	2927	11	Castilla-La	Albacete	Albacete	100.001 a	No pertenc	Anonimiza

En la ventana que se abrirá seleccionamos la carpeta en la que estamos trabajando. Recordemos que estamos guardando nuestros datos en la carpeta CIS en la unidad [K:]

Al fichero activo lo vamos a llamar CIS2927, tendrá como extensión por defecto "SAV". Después de pulsar la opción **Guardar** habremos generado el fichero activo. En próximas sesiones simplemente abriendo [K:\CIS\CIS2927.SAV] accederemos a la matriz ya con datos y etiquetas.

Segunda Parte: Obtención de estadísticos de resumen

Desde la parrilla del fichero activo podemos comenzar a solicitar datos sobre la encuesta. En concreto estamos interesados en conocer las valoraciones dadas a diferentes políticos. En concreto la pregunta 17, en la que se valoran desde 0 a 10 a un total de 14 políticos.

P.17 Le agradecería que me indicara si conoce a cada uno de los siguientes líderes políticos y qué valoración le merece su actuación política. Puntúelos de 0 a 10, sabiendo que el 0 significa que lo valora "muy mal" y el 10 que lo valora "muy bien".

	No conoce	Valoración	NS	NC	
	97	00-10	98	99	
- Enrique Álvarez Sostres.....	97	_____	98	99	(85)(86)
- Iñaki Antigüedad	97	_____	98	99	(87)(88)
- Joan Baldovi Roda	97	_____	98	99	(89)(90)
- Uxue Barkos	97	_____	98	99	(91)(92)
- Alfred Bosch.....	97	_____	98	99	(93)(94)
- Rosa Díez	97	_____	98	99	(95)(96)
- Josep A. Durán i Lleida	97	_____	98	99	(97)(98)
- Josu Erkoreka	97	_____	98	99	(99)(100)
- Francisco Jorquera.....	97	_____	98	99	(101)(102)
- Cayo Lara.....	97	_____	98	99	(103)(104)
- Ana María Oramas	97	_____	98	99	(105)(106)
- Alfredo Pérez Rubalcaba	97	_____	98	99	(107)(108)
- Mariano Rajoy.....	97	_____	98	99	(109)(110)
- Carlos Salvador	97	_____	98	99	(111)(112)

Antes de solicitar cualquier estadístico conviene observar los datos, para ello vamos a solicitar la distribución de frecuencias. Como forma expositiva, vamos a analizar los datos de Mariano Rajoy y de Alfredo Pérez Rubalcaba. Conviene que el alumno, realice la práctica con la valoración de otros dos políticos.

Para la obtención de frecuencias nos iremos –desde la ventana del fichero activo- a **Analizar**, opción **Estadística Descriptiva**, para seleccionar **Frecuencias**

Caso	ESTU	CUES	CCAA	PROV										
1	2927	1	País Vasco	Álava	Burgos/Vit	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Muy mala	
2	2927	2	País Vasco	Álava	Burgos/Vit	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Muy mala	
3	2927	3	País Vasco	Álava	Burgos/Vit	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Muy mala	
4	2927	4	País Vasco	Álava	Burgos/Vit	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Mala	
5	2927	5	País Vasco	Álava	Burgos/Vit	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Regular	
6	2927	6	País Vasco	Álava	Burgos/Vit	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Mala	
7	2927	7	País Vasco	Álava	Burgos/Vit	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Regular	
8	2927	8	País Vasco	Álava	Burgos/Vit	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Regular	
9	2927	9	Castilla-La	Albacete	Albacete	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Mala	
10	2927	10	Castilla-La	Albacete	Albacete	100.001 a	No pertenc	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Mala	

En la ventana que se abre seleccionamos a ambos políticos y los enviamos a la caja de **Variable(s)**, en dicha caja aparecerán con el nombre de variable P1713 que corresponde a Mariano Rajoy (pregunta 17, opción 13) y P1712 que corresponde a Alfredo Pérez Rubalcaba (pregunta 17 opción 12).

The screenshot shows the PSPPIRE Editor de Datos interface. A 'Frecuencias' dialog box is open, allowing the user to select variables for analysis. The 'Variable(s):' field contains 'P1712' and 'P1713'. The 'Estadísticos:' section has the following options checked: 'Media', 'Desviación Estándar', 'Mínimo', and 'Máximo'. Other options like 'Error estándar en la media', 'Varianza', 'Asimetría', 'Error estándar de la asimetría', 'Intervalo', 'Modo', 'Curtosis', 'Error estándar en la curtosis', 'Mediana', and 'Suma' are unchecked. The 'Incluir valores perdidos' option is also unchecked. The 'OK' button is highlighted, indicating the user's intention to proceed with the analysis.

Caso	P1706	P1707	P1708	P1709	P1710	P1711	P1714	P18	P1901	P1902	P1903	P1904	P1905	P1906	P20	P2101	P2102	P2103	P21A01	P21A02	P21A03	P2201	P2202	P2203	P2204	P2205	P2206	P2207	P2208	P2209	P2210	P2211	P2212	P2213	P23	P24	P25	P26	P27				
1																																											
2																																											
3																																											
4																																											
5																																											
6																																											
7																																											
8																																											
9																																											
10																																											
11																																											
12																																											
13																																											
14																																											
15																																											
16																																											
17																																											
18																																											
19																																											
20																																											
21																																											
22																																											
23																																											
24																																											
25																																											
26																																											
27																																											
28																																											
29																																											
30																																											
31	2927	31	Comunida	Alicante	Alicante	100.001 a	No perten	Anonimiza	Anonimiza	Anonimiza	Española	No proced	Muy mala	Peor																													

Pulsaremos al botón **Aceptar** y los resultados aparecerán en la ventana de resultados.

Con la opción **Aceptar** obtenemos inmediatamente los resultados. Si utilizamos el botón **Pegar** en vez de obtener resultados, el programa nos añadirá en el fichero de sintaxis un comando, que deberíamos ejecutar desde dicho fichero. Esto será útil en algunas ocasiones.

FRE COMANDO FREQUENCIES

En el fichero de sintaxis ES2927.SPS ya venía en la última línea un comando para el cálculo de frecuencias:

FREQUENCIES VARIABLES=ALL.

Dicho comando pide las distribuciones de frecuencias de todas las variables del estudio.

La sintaxis del comando FREQUENCIES resulta sencilla. Dicho comando se puede abreviar escribiendo únicamente las tres primeras letras FRE y especificando en el subcomando VAR= la lista de variables.

En el fichero ES2927.SPS resulta conveniente borrar dicha instrucción, no tiene mucho sentido un listado largo con todas las frecuencias. En el caso que nos ocupa podemos escribir mediante sintaxis la demanda de frecuencias para las valoraciones de ambos políticos

FRE VAR=P1712 P1713.

(No hay que olvidar el punto final)

Los resultados que vamos a obtener son:

Alfredo Pérez Rubalcaba

<i>Etiqueta de Valor</i>	<i>Valor</i>	<i>Frecuencia</i>	<i>Porcentaje</i>	<i>Valid Percent</i>	<i>Cum Percent</i>
	0	379	15,28	15,28	15,28
	1	84	3,39	3,39	18,67
	2	160	6,45	6,45	25,12
	3	175	7,06	7,06	32,18
	4	240	9,68	9,68	41,85
	5	465	18,75	18,75	60,60
	6	226	9,11	9,11	69,72
	7	220	8,87	8,87	78,59
	8	153	6,17	6,17	84,76
	9	62	2,50	2,50	87,26
	10	54	2,18	2,18	89,44
No conoce	97	24	,97	,97	90,40
N.S.	98	188	7,58	7,58	97,98
N.C	99	50	2,02	2,02	100,00
<i>Total</i>		2480	100,0	100,0	

Alfredo Pérez Rubalcaba

<i>N</i>	<i>Válido</i>	2480
	<i>Perdidos</i>	0
<i>Media</i>		14,17
<i>Desv Std</i>		28,97
<i>Mínimo</i>		,00
<i>Máximo</i>		99,00

Mariano Rajoy

<i>Etiqueta de Valor</i>	<i>Valor</i>	<i>Frecuencia</i>	<i>Porcentaje</i>	<i>Valid Percent</i>	<i>Cum Percent</i>
	0	350	14,11	14,11	14,11
	1	82	3,31	3,31	17,42
	2	126	5,08	5,08	22,50
	3	182	7,34	7,34	29,84
	4	200	8,06	8,06	37,90
	5	456	18,39	18,39	56,29
	6	241	9,72	9,72	66,01
	7	214	8,63	8,63	74,64
	8	190	7,66	7,66	82,30
	9	76	3,06	3,06	85,36
	10	96	3,87	3,87	89,23
No conoce	97	10	,40	,40	89,64
N.S.	98	205	8,27	8,27	97,90
N.C	99	52	2,10	2,10	100,00
<i>Total</i>		2480	100,0	100,0	

Mariano Rajoy

<i>N</i>	<i>Válido</i>	2480
	<i>Perdidos</i>	0
<i>Media</i>		14,63
<i>Desv Std</i>		29,15
<i>Mínimo</i>		,00
<i>Máximo</i>		99,00

Las tablas anteriores nos muestran la distribución de frecuencias para cada uno de los valores desde 0 a 10 más los códigos No conoce (97) No sabe (98) No contesta (99). Además de las frecuencias absolutas, tenemos el porcentaje, los porcentajes válidos –que explicaremos más adelante su significado- y la distribución porcentual acumulada basada en los porcentajes válidos.

Después de la tabla de frecuencias aparece una tabla con diferentes estadísticos. Nos dice que N, el número de casos, ha sido de 2480 para las dos variables, los 2480 han sido válidos porque no hay ningún caso perdido. Es decir, para todos los cuestionarios se ha rellenado dicha variable, bien con una valoración, bien con un código de conocimiento o de no respuesta. Sin embargo, hay situaciones en que puede que no exista para algunos casos valor en dicha variable, por ejemplo porque no es procedente la respuesta o no está definida la variable (por ejemplo lugar de trabajo para personas en paro).

A continuación nos ofrece las medias y las desviaciones típicas, así como los mínimos y máximos.

Probablemente el lector haya pensado que algo falla. Por ejemplo M. Rajoy tiene una valoración media de 14,63 y A. Pérez de 14,17. En ambos casos la media de la puntuación entre 0 y 10 es mayor de 10. Resulta evidente que no tiene sentido dicho resultado. El problema está en que los códigos 97,98 y 99 se han computado también como valores.

Para calcular los estadísticos de la valoración deberemos seleccionar únicamente a aquéllos entrevistados que han valorado a los políticos. Para realizar dicha operación lo haremos en dos fases; primero haremos una copia de la variable, y segundo sobre la variable copiada, declararemos inválidos los valores 97,98 y 99.

Dicho conjunto de operaciones lo vamos a realizar con comandos de sintaxis. Los siguientes comandos los escribiremos al final de nuestro fichero de sintaxis (ES2927.SPS)

En primer lugar para copiar las variables utilizamos el comando **COMPUTE**

COMPUTE P1713V=P1713.
COMPUTE P1712V=P1712.

La sintaxis del comando **COMPUTE** es sencilla, la nueva variable = vieja variable. (Observe que cada comando como es habitual termina con un punto). En una columna nueva hemos copiado los valores de la variable P1713 y a esa nueva variable la hemos denominado P1713V. (Hemos elegido el mismo nombre de variable añadiendo la V para indicar que dicha variable será la que contenga los valores).

Si ejecutamos ambos comandos, veremos que al final de la rejilla del fichero activo se han añadido dos nuevas columnas después de la variable STATUS que era la última.

	E3	E4	C1	C1A	C2	C2A	C2B	C3	C4	P26A	VOTOSIM	RECUERE	ESTUDIO	OCUPA	CONDIC	ESTATUS	P1713V	P1712V
nes	12	La mañan	Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	No proced	PNV	Otros parti	Sin estud	Artesanos	Jubilados	Obreros ci	.00	6.00
nes	15	La mañan	Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	Ninguno	Ninguno	No votó	Sin estud	Trabajado	Jubilados	Obreros n	98.00	98.00
nes	12	La mañan	Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	No proced	Amalur	Otros parti	Superiores	Técnicos y	Empleado	Nuevas ci	.00	4.00
nes	16	Mediodía (Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	No proced	PP	N.C.	Primaria	Trabajado	Trabajo dc	Obreros n	6.00	4.00
nes	12	Mediodía (Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	No proced	PNV	Otros parti	Primaria	Artesanos	Jubilados	Viejas clas	3.00	2.00
nes	15	Mediodía (Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	N.S.	N.S.	N.C.	F.P.	Operadore	Jubilados	Viejas clas	2.00	6.00
nes	15	Mediodía (Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	PNV	PNV	Otros parti	Primaria	Trabajado	Obreros n	Obreros n	3.00	3.00
nes	15	Mediodía (Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	No proced	PP	PP	F.P.	Operadore	Pequeños	Obreros ci	5.00	5.00
les	15	La mañan	Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	No proced	PP	PP	Secundari	Técnicos y	Parados	Nuevas ci	9.00	.00
les	15	Mediodía (Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	Ninguno	Ninguno	PP	Primaria	Trabajado	Obreros n	Obreros n	.00	5.00
les	20	Mediodía (Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	No proced	IU (ICV en	PP	Sin estud	Trabajado	Jubilados	Viejas clas	98.00	98.00
les	20	Mediodía (Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	No proced	PSOE	PSOE	Sin estud	Artesanos	Jubilados	Obreros ci	.00	9.00
les	15	Mediodía (Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	Ninguno	Ninguno	No votó	Primaria	Artesanos	Parados	Obreros ci	98.00	98.00
les	15	Mediodía (Correcto	.	Entrevista	.	.	Entrevista	Anonimiza	No proced	Otros parti	Otros parti	Primaria	Trabajado	Empleado	Nuevas ci	3.00	5.00

(Si no aparecen los valores aún, vaya a Transformar+Ejecutar transformaciones pendientes)

Veremos que ambas columnas contienen una copia de los valores de las variables originales. (P1713 y P1712)

Una vez copiadas las variables, vamos a eliminar los valores 97,98 y 99. Para ello utilizaremos el comando **MISSING VALUES**

MISSING VALUES P1713V P1712V (97,98,99).

El comando MISSING VALUES especifica las variables y entre paréntesis el conjunto de valores que se declaran perdidos separados por comas. Como todos los comandos termina con un punto.

En el fichero de sintaxis (ES2917.SPS) en la parte final verá un comando MISSING VALUE.

MISSING VALUES

POA (0) P21A01 (0) P21A02 (0) P21A03 (0) P28A (0) P31A (0) P32A (0)
P37A (0) P37B (0) P39A (0) P41A (0) P41B (0) P26A (0).

La ejecución de dicho comando declara como valor perdido el código 0 para el conjunto de variables especificado.

Para terminar nuestra operación podemos añadir un comando variable label para etiquetar las nuevas variables. A continuación escribimos la secuencia de sintaxis. Añadiremos también el comando *. Este comando significa comentario y termina también con un punto. Desde el asterisco hasta el punto podemos introducir el texto y anotaciones que queramos en el fichero de sintaxis. (Habitualmente para que sea más visible en vez de un asterisco se utilizan más).

****Sintaxis para la transformación de las variables de valoración de políticos para la obtención de puntuaciones medias. Primero con COMPUTE se hace una copia de las variables. Después sobre las nuevas variables se declaran los códigos perdidos. Finalmente se etiquetan las nuevas variables.

```
COMPUTE P1713V=P1713.
```

```
COMPUTE P1712V=P1712.
```

```
MISSING VALUES P1713V P1712V (97,98,99).
```

Ahora simplemente es cuestión de solicitar la distribución de frecuencias de P1713V y de P1712V. Recordemos que desde la parrilla de datos, seleccionamos **Analizar: Estadística Descriptiva: Frecuencias**, e incluimos las dos variables, que las encontraremos al final del listado. Los resultados ahora serán:

Valoración de Mariano Rajoy

<i>Etiqueta de Valor</i>	<i>Valor</i>	<i>Frecuencia</i>	<i>Porcentaje</i>	<i>Valid Percent</i>	<i>Cum Percent</i>
	,00	350	14,11	15,82	15,82
	1,00	82	3,31	3,71	19,52
	2,00	126	5,08	5,69	25,21
	3,00	182	7,34	8,22	33,44
	4,00	200	8,06	9,04	42,48
	5,00	456	18,39	20,61	63,08
	6,00	241	9,72	10,89	73,97
	7,00	214	8,63	9,67	83,64
	8,00	190	7,66	8,59	92,23
	9,00	76	3,06	3,43	95,66
	10,00	96	3,87	4,34	100,00
	97,00	10	,40	Perdidos	
	98,00	205	8,27	Perdidos	
	99,00	52	2,10	Perdidos	
<i>Total</i>		2480	100,0	100,0	

Valoración de Mariano Rajoy

<i>N</i>	<i>Válido</i>	2213
	<i>Perdidos</i>	267
<i>Media</i>		4,55
<i>Desv Std</i>		2,86
<i>Mínimo</i>		,00
<i>Máximo</i>		10,00

Valoración de Alfredo Pérez Rubalcaba

<i>Etiqueta de Valor</i>	<i>Valor</i>	<i>Frecuencia</i>	<i>Porcentaje</i>	<i>Valid Percent</i>	<i>Cum Percent</i>
	,00	379	15,28	17,09	17,09
	1,00	84	3,39	3,79	20,87
	2,00	160	6,45	7,21	28,09
	3,00	175	7,06	7,89	35,98
	4,00	240	9,68	10,82	46,80
	5,00	465	18,75	20,96	67,76
	6,00	226	9,11	10,19	77,95
	7,00	220	8,87	9,92	87,87
	8,00	153	6,17	6,90	94,77
	9,00	62	2,50	2,80	97,57
	10,00	54	2,18	2,43	100,00
	97,00	24	,97	Perdidos	
	98,00	188	7,58	Perdidos	
	99,00	50	2,02	Perdidos	
<i>Total</i>		2480	100,0	100,0	

Valoración de Alfredo Pérez Rubalcaba

<i>N</i>	<i>Válido</i>	2218
	<i>Perdidos</i>	262
<i>Media</i>		4,25
<i>Desv Std</i>		2,75
<i>Mínimo</i>		,00
<i>Máximo</i>		10,00

Ahora podemos comprender la diferencia entre la distribución de porcentajes y la de "Porcentajes Válidos". Por ejemplo los 456 entrevistados que han valorado con un 5 a Mariano Rajoy, son el 18,39% de la muestra (456/2480) pero suponen el 20,61% (456/2213) de quienes han valorado a dicho político que han sido 2213 entrevistados.

Los porcentajes acumulados, están basados en los porcentajes de casos válidos. Podemos observar por ejemplo que un 83,64% han valorado a Rajoy con 7 o menos, mientras que dicha cifra es mayor para Rubalcaba 87,87%. La distribución señala que en el caso de Rajoy hay relativamente más puntuaciones altas que en el caso de Rubalcaba.

Ahora las medias son reales, están hechas sobre el número de valoraciones. Podemos observar que desde 0 (como valor mínimo) a 10 (como valor máximo). La valoración de Rajoy es 4,55 ligeramente superior al 4,25 obtenido por Rubalcaba. La tabla también nos ofrece la desviación típica para ambas distribuciones.

En la tabla siguiente encontrará los resultados de valoración al conjunto de políticos incluidos en la pregunta 17. Seleccione un par de ello y reproduzca el procedimiento anterior para encontrar estos resultados.

	Media	Desviación típica	(N)
Enrique Álvarez Sostres	3.71	2.35	(93)
Iñaki Antigüedad	3.08	2.63	(124)
Joan Baldoví Roda	3.54	2.51	(69)
Uxue Barkos	4.41	2.48	(276)
Alfred Bosch	4.02	2.49	(116)
Rosa Díez	4.75	2.45	(1483)
Josep A. Durán i Lleida	4.13	2.44	(1296)
Josu Erkoreka	3.38	2.47	(421)
Francisco Jorquera	3.89	2.42	(121)
Cayo Lara	4.17	2.44	(1146)
Ana María Oramas	4.09	2.61	(136)
Alfredo Pérez Rubalcaba	4.25	2.75	(2218)
Mariano Rajoy	4.55	2.86	(2213)
Carlos Salvador	3.71	2.52	(84)

Para terminar la práctica nos centraremos en el cálculo de otras medidas de tendencia central como son la moda y la mediana. Para ello utilizaremos la pregunta 30, que contiene la edad de los entrevistados. Observemos como es habitual que el código 99 está reservado a los No Contesta. Por lo tanto, copiamos la variable en una nueva –por ejemplo EDAD- y declaramos *missing* el valor 99.

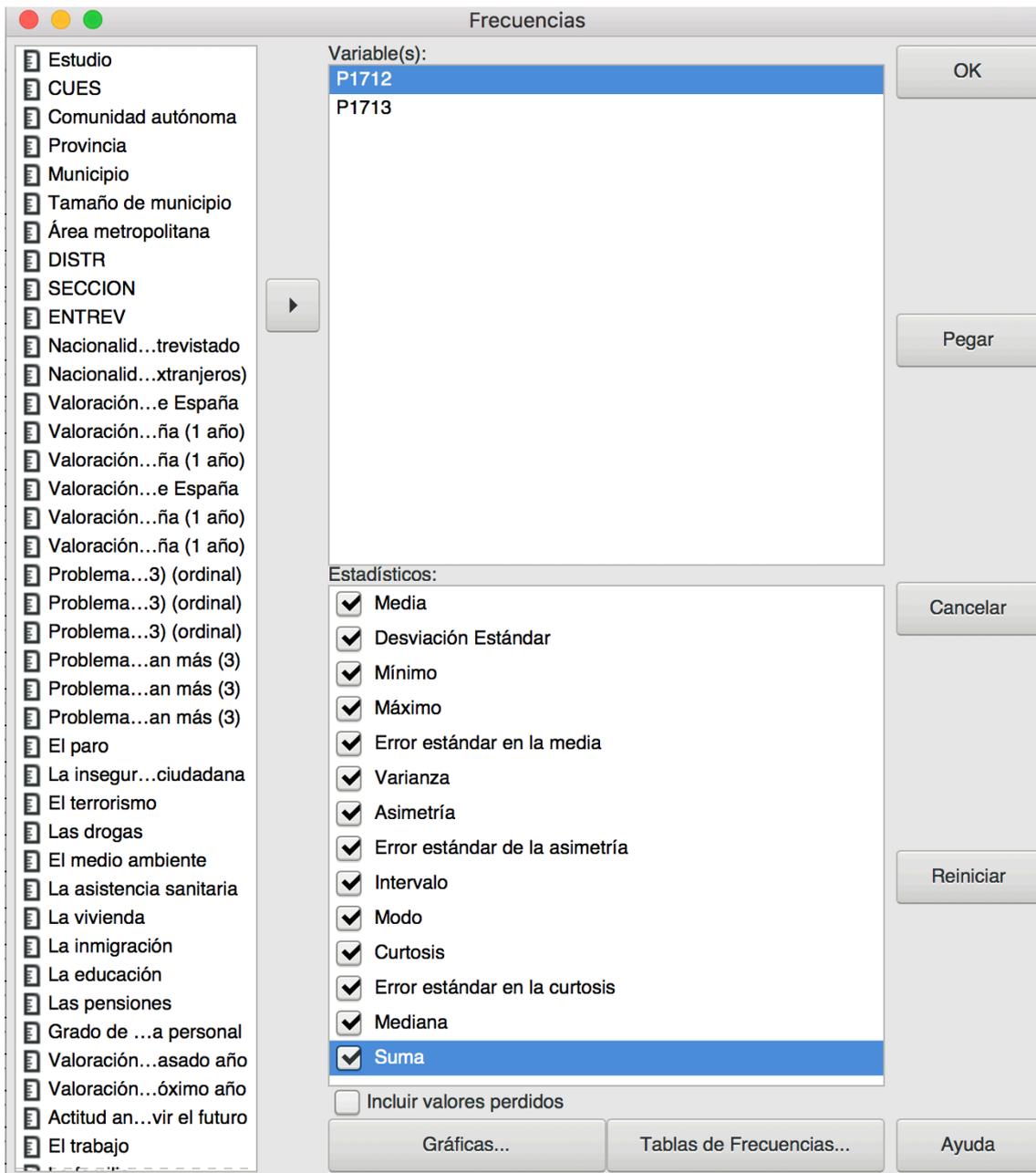
****Instrucciones para generar la variable EDAD.

COMPUTE EDAD=P30.
MISSING VALUE EDAD (99).

Una vez ejecutados los comandos anteriores pedimos los resultados **Analizar: Estadística Descriptiva: Frecuencias**, como variable incluimos EDAD y marcamos todos los estadísticos que hemos estudiado:

Media
Desviación Estándar
Mínimo
Máximo
Varianza
Intervalo
Modo²
Mediana
Suma

² Es un error en la tipografía de PSPP quiere decir Moda.



Obtenemos la tabla de frecuencias y a continuación los estadísticos:

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	18,00	28	1,1	1,1	1,1
	19,00	36	1,5	1,5	2,6
	20,00	24	1,0	1,0	3,6
	21,00	41	1,7	1,7	5,2
	22,00	27	1,1	1,1	6,3
	23,00	35	1,4	1,4	7,7
	24,00	30	1,2	1,2	8,9
	25,00	61	2,5	2,5	11,4
	26,00	43	1,7	1,7	13,1
	27,00	34	1,4	1,4	14,5
	28,00	45	1,8	1,8	16,3
	29,00	26	1,0	1,0	17,4
	30,00	41	1,7	1,7	19,0
	31,00	54	2,2	2,2	21,2
	32,00	52	2,1	2,1	23,3
	33,00	55	2,2	2,2	25,5
	34,00	73	2,9	2,9	28,5
	35,00	48	1,9	1,9	30,4
	36,00	48	1,9	1,9	32,3
	37,00	44	1,8	1,8	34,1
	38,00	43	1,7	1,7	35,8
	39,00	43	1,7	1,7	37,6
	40,00	57	2,3	2,3	39,9
	41,00	44	1,8	1,8	41,6
	42,00	46	1,9	1,9	43,5
	43,00	56	2,3	2,3	45,8
	44,00	76	3,1	3,1	48,8
	45,00	56	2,3	2,3	51,1
	46,00	54	2,2	2,2	53,3
	47,00	47	1,9	1,9	55,2
	48,00	31	1,3	1,3	56,4
	49,00	36	1,5	1,5	57,9
	50,00	45	1,8	1,8	59,7
	51,00	39	1,6	1,6	61,3
	52,00	38	1,5	1,5	62,8
	53,00	50	2,0	2,0	64,8
	54,00	39	1,6	1,6	66,4
	55,00	30	1,2	1,2	67,6
	56,00	28	1,1	1,1	68,7
	57,00	31	1,3	1,3	70,0
	58,00	31	1,3	1,3	71,2
	59,00	34	1,4	1,4	72,6
	60,00	39	1,6	1,6	74,2
	61,00	26	1,0	1,0	75,2
	62,00	33	1,3	1,3	76,6
	63,00	36	1,5	1,5	78,0
	64,00	42	1,7	1,7	79,7
	65,00	41	1,7	1,7	81,4
	66,00	26	1,0	1,0	82,4
	67,00	30	1,2	1,2	83,6
	68,00	30	1,2	1,2	84,8
	69,00	23	,9	,9	85,8
	70,00	40	1,6	1,6	87,4
	71,00	28	1,1	1,1	88,5
	72,00	16	,6	,6	89,1
	73,00	19	,8	,8	89,9
	74,00	25	1,0	1,0	90,9
	75,00	24	1,0	1,0	91,9
	76,00	21	,8	,8	92,7
	77,00	27	1,1	1,1	93,8
	78,00	18	,7	,7	94,6
	79,00	16	,6	,6	95,2
	80,00	17	,7	,7	95,9
	81,00	20	,8	,8	96,7
	82,00	17	,7	,7	97,4
	83,00	9	,4	,4	97,7
	84,00	13	,5	,5	98,3
	85,00	12	,5	,5	98,7
	86,00	8	,3	,3	99,1
	87,00	5	,2	,2	99,3
	88,00	7	,3	,3	99,6
	89,00	5	,2	,2	99,8
	90,00	2	,1	,1	99,8
	91,00	1	,0	,0	99,9
	92,00	1	,0	,0	99,9
	94,00	2	,1	,1	100,0
	Total	2478	99,9	100,0	
Perdidos	99,00	2	,1		
Total		2480	100,0		

EDAD

N	Válido	2478
	Perdidos	2
Media		47,48
Modo		44,00
Desv Std		17,79
Varianza		316,31
Err.Est.Curt.		,10
Err.Est.Asim.		,05
Intervalo		76,00
Mínimo		18,00
Máximo		94,00
Suma		117662,0
Percentiles	50 (Mediana)	45,50

En total hay dos casos perdidos, siendo 2478 los casos válidos –de los que tenemos información- que intervienen en el cálculo de los estadísticos.

- Media: Nos dice que la edad media de la muestra es de 47,48 años (Téngase en cuenta que el universo está compuesto por mayores de 18 años.) Es el resultado de la Suma (117662) entre N Válido (2478).
- Moda: El mayor número de entrevistados tienen 44 años. En total hay 76.
- Desviación Estándar: Aunque el programa la denomina así se trata de la desviación típica muestral o cuasidesviación. (Consulte el tema VIII del texto *Estadística para la Investigación Social*). Para estos tamaños muestrales dado que la diferencia de cálculo es en el denominador, que utiliza N-1 en vez de N, no hay diferencias en el resultado.
- Varianza: Es el cuadrado de la Desviación Estándar.
- Intervalo: Diferencia entre el mayor valor 94 y el menor 18. Es el Recorrido de la variable.
- Suma: Es la suma de todos los valores –edades- de la muestra. En total suman 117662 años.
- Percentil 50 o Mediana: 45,5 años es el valor que separa la muestra en dos grupos iguales. Los diferentes programas utilizan algoritmos distintos para obtener la mediana u otros Cuantiles³.

³ Rob J. Hyndman and Yanan Fan en su artículo “Sample Quantiles in Statistical Packages” publicado en la revista *The American Statistician*, Vol. 50, No. 4 (Nov., 1996), pp. 361-365, repasan los algoritmos utilizados en los principales paquetes estadísticos (BMDP, GLIM, SAS, Splus, SPSS, etc.) para calcular la mediana o los cuantiles y encuentran hasta 9 sistemas diferentes, en gran parte de los cuales no se hace público su sistema de cálculo. Los autores denuncian la falta de estandarización de algoritmos que provocan confusión en los usuarios.