

Documentación y metainformación requerida en la entrega de proyectos para su depósito en ARCES

La documentación aportada del proyecto debe incluir los siguientes contenidos: (como modelo, puede consultar las encuestas del CIS en www.cis.es o www.analisis.cis.es).

(Los campos entre corchetes [] son opcionales)

- Nombre y datos de contacto del(a) **Investigador(a) Principal** y, opcionalmente, del equipo completo
- Título del proyecto y código asignado al mismo
- Informe descriptivo del proyecto (indicando objeto, objetivos, antecedentes, contexto, fases del estudio, metodología, incidencias etc.)
- Informe de resultados (incluyendo por ejemplo: tablas de distribución de frecuencias o marginales, conclusiones, etc.)
- Ficha técnica con las principales características técnicas de la encuesta:
 - universo poblacional
 - ámbito (localización física o territorial del universo)
 - afijación
 - ponderación
 - criterios del diseño muestral:
 - tamaño de la muestra diseñada y realizada
 - método de muestreo
 - [puntos de muestreo]
 - [error muestral]
 - modo de administración de la encuesta (PAPI, CAPI, CATI, etc.)
 - instrucciones para el trabajo de campo (sustituciones de cuotas, rutas, etc.)
 - fechas de realización del trabajo de campo con inicio y fin
- Cuestionarios empleados (en los idiomas utilizados), así como cualquier material de apoyo y otros documentos relevantes de recogida de datos cuando los haya:
 - Tarjetas (se muestran a la persona entrevistada con las posibles respuestas sobre las que elegir)
 - instrucciones para la administración de las preguntas (forma de preguntar, filtros, etc.)
- Fichero.sav bien construido (ANEXO I) o en su defecto:
 - Fichero de microdatos en ASCII, **anonimizados** de forma que se asegure la imposibilidad e irreversibilidad de la identificación directa de personas físicas o jurídicas que hayan participado como encuestadoras o encuestadas en el estudio. (ANEXO II)
 - Fichero de sintaxis en formato SPSS, o compatible. (ANEXO II)
 - Libro de códigos y estándares de codificación seguidos para todas las variables cuando proceda.
- Se detallarán las variables anonimizadas. En caso de encontrarse un error en esta anonimización, es responsabilidad del(a) Investigador(a) Principal y del Equipo Adjudicatario del Proyecto la obtención del nuevo fichero corregido y su envío por el método más rápido al CIS.
- [Fichero de metadatos], en su caso, en XML según el estándar DDI.

Previamente a su aceptación para el depósito en el archivo ARCES, el CIS podrá requerirle aclaraciones o información adicional que permita interpretar la documentación aportada.

FICHERO.SAV BIEN DEFINIDO¹

El formato adecuado para el depósito de los estudios en nuestro Banco de Datos es el formato ASCII (ver Anexo II).

En su defecto, se aceptarán ficheros de datos con extensión .sav siempre y cuando tengan cumplimentados todos sus atributos de forma exhaustiva.

Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
--------	------	---------	-----------	----------	---------	----------	----------	------------	--------	-----

Los requisitos que deberán cumplir son:

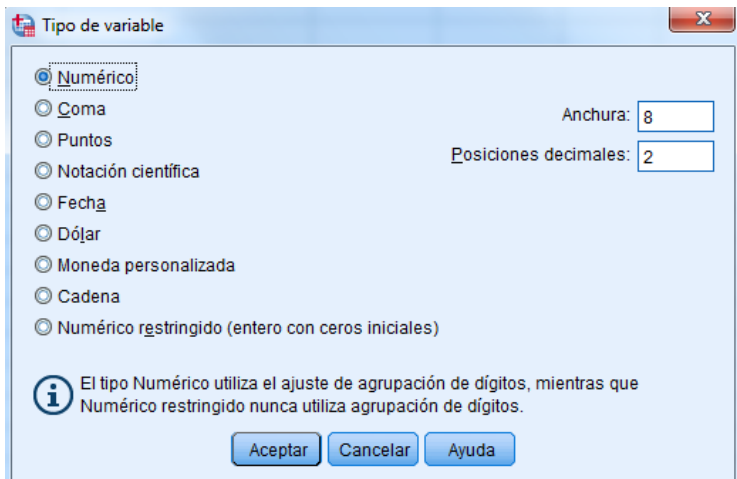
Nombre de variables

Para los nombres de variable en SPSS se aplican las siguientes normas, entre otras:

- Cada nombre de variable debe ser exclusivo; no se permiten duplicados.
- Los nombres de variable pueden tener una longitud de hasta 64 caracteres y el primer carácter debe ser una letra. Los caracteres posteriores pueden ser números.
- Las variables no pueden contener espacios, acentos ni caracteres especiales (ñ, ç, ü, ª, etc.).
- El punto, el guión bajo y los caracteres \$, # y @ se pueden utilizar dentro de los nombres de variable.
- Se deben evitar los nombres de variable que terminan con un punto, ya que el punto puede interpretarse como un terminador de comando.
- Las palabras reservadas no se pueden utilizar como nombres de variable: ALL, AND, BY, EQ, GE, GT, LE, LT, NE, NOT, OR, TO y WITH.

Tipo

Según se graben los datos en 'Tipo' habrá que escoger entre estas opciones:



¹ Estas instrucciones son de aplicación cuando se trabaja en SPSS por medio del interfaz de usuario

Anchura y Decimales

En 'Anchura' hay que poner el número máximo de espacios que se necesita para grabar los datos de esa variable. Por ejemplo, si la variable numérica puede tener como máximo dos números enteros y dos decimales, en 'Anchura' se pondrá el total, 4 y en 'Decimales': 2.

Anchura	Decimales
4	2

Si una variable categórica se graba con códigos numéricos, por ejemplo, código 1 para *hombre* y 2 para *mujer*, solo es necesario un entero y ningún decimal:

Anchura	Decimales
1	0

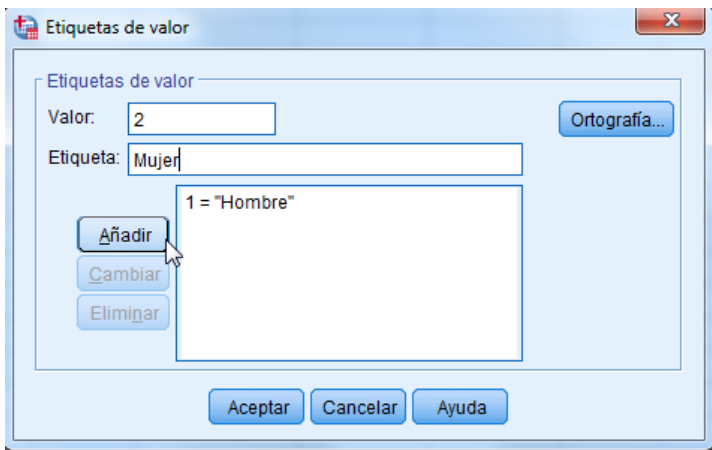
IMPORTANTE: No se debe dejar ancho mayor al que realmente necesita la variable.

Etiqueta

En esta columna se graban las etiquetas de las distintas variables (títulos). Se puede asignar etiquetas de variable descriptivas de hasta 256 caracteres de longitud. Las etiquetas de variable pueden contener espacios y caracteres reservados que no se admiten en los nombres de variable y van a representar de forma más exhaustiva que el nombre el significado de esa variable.

Valores

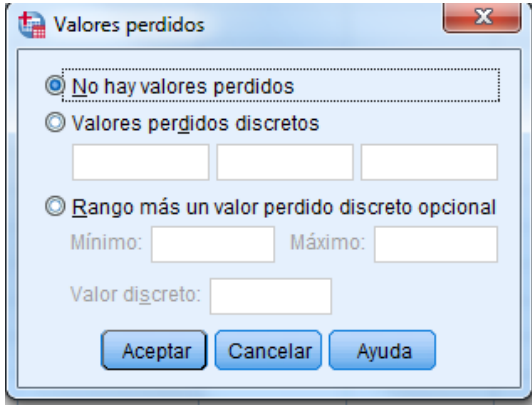
Aquí se pondrán las etiquetas de los distintos valores que puede obtener una variable cuando se utilizan códigos numéricos para representar categorías que no son numéricas (por ejemplo, códigos 1 y 2 para *hombre* y *mujer*, respectivamente):



Las etiquetas en este caso pueden tener una longitud máxima de 120 caracteres para cada valor.

Perdidos

Si se considera que algún valor de la variable tiene que mantenerse al margen de los cálculos, se puede definir, personalmente, como valor perdido en el cuadro que aparece en esta columna:



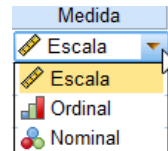
Esos valores, definidos como **perdidos por el usuario**, se excluyen de la mayoría de los cálculos. Son por ejemplo, el código con el que se graban los casos filtrados en una variable o valores no válidos en esa variable y que, de aparecer, se deben a un error de grabación o de otro tipo, por ejemplo, si se graba un 3 en la variable de sexo que hemos codificado como 1 y 2 para *hombre* y *mujer*.

Columnas y Alineación

Estas columnas tienen menor relevancia ya que, en el primer caso, solo afecta a la forma de visualizarse las distintas columnas de las variables en el Editor de datos de SPSS y en el segundo caso, se pone automáticamente cuando se selecciona al principio el tipo de variable y aunque se pueda cambiar, conviene dejarlo por defecto.

Medida

El nivel de medida de cada variable se define en esta columna:



Rol

En esta columna no es necesaria ninguna modificación, todas pueden quedar como variable de Entrada.

En conclusión, hay que prestar especial atención a los siguientes atributos:

- Nombre
- Tipo
- Anchura
- Decimales
- Etiqueta
- Valores
- Perdidos
- Medida

INSTRUCCIONES PARA GENERAR FICHEROS ASCII A PARTIR DE UN FICHERO .sav DE SPSS

Los ficheros a crear son: Datos_ID
Sintaxis_ID

ADVERTENCIA: La conversión de ficheros .sav a ficheros de datos en ASCII (Datos_ID) puede ser automática pero no lo es la conversión a ficheros de sintaxis (Sintaxis_ID). Lo que aquí proponemos es un método que puede facilitar esta tarea a quienes no estén familiarizados con la construcción de sintaxis de SPSS.

Generación del fichero de datos (Datos_ID)

La generación del fichero de microdatos (Datos_ID) en ASCII se realiza mediante sintaxis de SPSS con el comando WRITE OUTFILE.

También se puede generar el fichero a través del menú de la ventana del editor de datos (.sav) con la opción 'Archivo>Guardar como' poniendo el nombre del fichero 'Datos_ID' y el tipo: ASCII en formato fijo (.dat).

Generación del fichero de sintaxis (Sintaxis_ID)

Para leer este fichero de datos generado hay que crear un fichero de sintaxis (Sintaxis_ID). Los comandos básicos que tienen que aparecer son:

DATA LIST
VARIABLE LABELS
VALUE LABELS
WEIGHT (si hay ponderación)

Si se conocen los comandos se hará directamente en la ventana de sintaxis. También se puede utilizar la información que aparece cuando está abierto el fichero .sav en 'Archivo>Mostrar información del archivo de datos>Archivo de trabajo' para facilitar su creación.

De las tablas que nos proporciona podemos extraer la información necesaria para construir la sintaxis. Para transformarlas resulta más cómodo copiarlas en un procesador de texto. Por ejemplo:

Información sobre las variables

Variable	Ubicación	Etiqueta	Nivel de medida	Papel	Ancho de columna	Alineación	Formato de impresión	Formato de escritura	Valores perdidos
ESTU	1	Estudio	Nominal	Entrada	6	Derecha	F4	F4	
CUES	2	CUES	Escala	Entrada	7	Derecha	F5	F5	
CCAA	3	Comunidad autónoma	Nominal	Entrada	6	Derecha	F2	F2	
PROV	4	Provincia	Escala	Entrada	6	Derecha	F2	F2	
MUN	5	Municipio	Escala	Entrada	5	Derecha	F3	F3	
P1	13	Valoración de	Nominal	Entrada	4	Derecha	F1	F1	
P2	14	Valoración retrospectiva de ...	Nominal	Entrada	4	Derecha	F1	F1	
P3	90	<ninguno>	Nominal	Entrada	7	Izquierda	A3	A3	
PESO	144	Ponderación	Escala	Entrada	8	Derecha	F6.3	F6.3	

Esta primera tabla va a ayudar a construir el comando de lectura de los datos (DATA LIST) y el comando que etiqueta las variables (VARIABLE LABELS). Pasos:

- En el procesador de texto eliminar todas las columnas no señaladas en amarillo y duplicar la tabla para crear un comando con cada una.
- Eliminar en la primera tabla la columna de 'Etiquetas' que en este caso no va a hacer falta y pasar la tabla a texto escogiendo como separador un espacio:

```
ESTU F4  
CUES F5  
CCAA F2  
PROV F2  
MUN F3  
P1 F1  
P2 F1  
P3 A3  
PESO F6.3
```

- Sustituir el formato de cada variable por las columnas en las que se han grabado. En la primera se cambia F4 por 1-4 porque F4 significa que tiene un formato numérico de 4 columnas. La siguiente variable como es F5 se cambiaría por las 5 siguientes columnas, es decir 5-9 y en la siguiente se cambiaría por 10-11. Así sucesivamente.

En el caso de la P3 aparece un formato A3, lo que indica que el formato es alfanumérico con 3 posiciones y se tendría que sustituir por las columnas que le corresponde con indicación del formato: 19-21(A).

La última variable tiene un formato numérico con decimales porque pone F6.3 lo que indica que es un formato de 6 columnas de las cuales las 3 últimas son decimales, por lo que se sustituiría por 22-27(3). El valor entre paréntesis indica el número de decimales.

- Una vez cambiados los formatos por sus columnas se construye el **DATA LIST**:
Tenemos que saber la dirección de la carpeta en la que se ha guardado el fichero **Datos_ID.dat**, por ejemplo: en la unidad C:\ en una carpeta que se llama estudios:

```
DATA LIST FILE='C:\estudios\Datos_ID.dat'  
/ ESTU 1-4 CUES 5-9 CCAA 10-11 PROV 12-13 MUN 14-16 P1 17 P2 18  
P3 19-21 (A) PESO 22-27 (3).
```

Si todas las variables con las que se trabaja tienen el mismo formato se puede simplificar este comando. Por ejemplo, si tenemos 100 variables con formato F2 quedaría simplificado a:

```
DATA LIST FILE='C:\estudios\Datos_ID.dat'  
/ VAR0001 TO VAR0100 1-200.
```

También se puede simplificar agrupando variables consecutivas con el mismo formato. Por ejemplo, tenemos 20 variables seguidas con F1, 20 más con F2 y a continuación otras 10 con F1:

```
DATA LIST FILE='C:\estudios\Datos_ID.dat'  
/ VAR0001 TO VAR0020 1-20 VAR0021 TO VAR0040 21-60 VAR0041 TO VAR0050 61-70.
```

- Eliminar en la segunda tabla la columna de 'formato' y pasar a texto con tabulación de separadores.

```
ESTU    Estudio  
CUES    Cuestionario  
CCAA    Comunidad autónoma  
PROV    Provincia  
MUN     Municipio  
P1      Valoración de la situación general de España  
PESO    Ponderación
```

- Sustituir tabulación por espacio y comilla simple y añadir comilla simple al final de cada etiqueta, y ya tenemos el segundo comando (**VARIABLE LABELS**):

```
VARIABLE LABELS ESTU 'Estudio'
CUES 'Cuestionario'
CCAA 'Comunidad autónoma'
PROV 'Provincia'
MUN 'Municipio'
P1 'Valoración de la situación general de España'
PESO 'Ponderación'
```

Con la siguiente tabla que nos ha facilitado SPSS creamos el comando que pone etiquetas a los valores de cada variable*, excepto a variables continuas (**VALUE LABELS**):

Valores de las variables

Valor	Etiqueta	
CCAA	1	Andalucía
	2	Aragón
	3	Asturias (Principado de)
	4	Balears (Illes)
	5	Canarias
	6	Cantabria
	7	Castilla La Mancha
	8	Castilla y León
	9	Cataluña
	10	Comunitat Valenciana
	11	Extremadura
	12	Galicia
	13	Madrid (Comunidad de)
	14	Murcia (Región de)
	15	Navarra (Comunidad Foral de)
	16	Pais Vasco
	17	Rioja (La)
	18	Ceuta (Ciudad Autónoma de)
	19	Melilla (Ciudad Autónoma de)
P1	1	Muy buena
	2	Buena
	3	Regular
	4	Mala
	5	Muy mala
	8	N.S.
9	N.C.	

* No se incluyen etiquetas de provincias ni municipios, en el ejemplo, para no alargar demasiado la tabla.

Pasar la tabla a texto y tratar el texto para que cumpla los siguientes requisitos:

- Delante de cada variable, excepto en la primera, hay que poner una barra inclinada /
- Poner comillas simples al principio y al final de las etiquetas de cada valor

De esta manera se construye el siguiente comando:

```
VALUE LABELS CCAA 1 'Andalucía'
2 'Aragón'
3 'Asturias (Principado de)'
4 'Balears (Illes)'
5 'Canarias'
6 'Cantabria'
7 'Castilla La Mancha'
8 'Castilla y León'
9 'Cataluña'
10 'Comunitat Valenciana'
11 'Extremadura'
12 'Galicia'
13 'Madrid (Comunidad de)'
14 'Murcia (Región de)'
15 'Navarra (Comunidad Foral de)'
16 'Pais Vasco'
```

17 'Rioja (La)'
18 'Ceuta (Ciudad Autónoma de)'
19 'Melilla (Ciudad Autónoma de)'
/P1 1 'Muy buena'
2 'Buena'
3 'Regular'
4 'Mala'
5 'Muy mala'
8 'N.S.'
9 'N.C.'

Con estos tres comandos se construye el programa de sintaxis básico. El fichero se guardará con el nombre **Sintaxis_ID**. Cada comando tiene que terminar con un punto. Además, si existen una o más variables por las que se tienen que ponderar los datos, hay que añadir el comando **WEIGHT**.

```
DATA LIST FILE='C:\estudios\Datos_ID.dat'  
  / ESTU 1-4 CUES 5-9 CCAA 10-11 PROV 12-13 MUN 14-16 P1 17 P2 18  
  P3 19-21 (A) PESO 22-27 (3).  
VARIABLE LABELS  ESTU 'Estudio'  
  CUES 'Cuestionario'  
  CCAA 'Comunidad autónoma'  
  PROV 'Provincia'  
  MUN 'Municipio'  
  P1 'Valoración de la situación general de España'  
  P2 'Valoración retrospectiva de la situación general de España (1 año)'  
  PESO 'Ponderación'.  
VALUE LABELS  CCAA 1 'Andalucía'  
  2 'Aragón'  
  3 'Asturias (Principado de)'  
  4 'Balears (Illes)'  
  5 'Canarias'  
  6 'Cantabria'  
  7 'Castilla La Mancha'  
  8 'Castilla y León'  
  9 'Cataluña'  
 10 'Comunitat Valenciana'  
 11 'Extremadura'  
 12 'Galicia'  
 13 'Madrid (Comunidad de)'  
 14 'Murcia (Región de)'  
 15 'Navarra (Comunidad Foral de)'  
 16 'País Vasco'  
 17 'Rioja (La)'  
 18 'Ceuta (Ciudad Autónoma de)'  
 19 'Melilla (Ciudad Autónoma de)'  
/P1 1 'Muy buena'  
  2 'Buena'  
  3 'Regular'  
  4 'Mala'  
  5 'Muy mala'  
  8 'N.S.'  
  9 'N.C.'  
/P2 1 'Mejor'  
  2 'Igual'  
  3 'Peor'  
  8 'N.S.'  
  9 'N.C.'.  
WEIGHT BY (nombre de la variable con los coeficientes de ponderación).
```