

Os arquivos de microdatos (matrices de datos) das enquisas do CIS, xunto coa documentación dos estudos, descárganse nun arquivo comprimido (MDxxxx.zip), que se deberá de descomprimir.

Neste arquivo encóntranse os seguintes documentos, necesarios para a explotación do estudo:

- **Arquivo de microdatos anonimizado** (DAxxxx) En formato ASCII. (*Ler a alínea 'Formato do arquivo de datos'*)
- **Cuestionario** (cuesxxxx). En formato .pdf.
- **Ficha técnica** (Ftxxxx). En formato .pdf
- **Libro de códigos** (codigoxxxx). En formato .pdf.
- **Tarjetas** (tarjetasxxxx). En formato .pdf. (Só para enquisas realizadas mediante entrevista persoal)
- **Programa de sintaxe** de lectura do arquivo de microdatos en **SPSS** (ESxxxx). En formato ASCII. (*Ler a alínea 'Formato do arquivo de sintaxe-SPSS'*)
- **Programa de sintaxe** de lectura do arquivo de microdatos en **SAS** (Sasxxxx). En formato ASCII. (*Ler a alínea 'Formato do arquivo de sintaxe-SAS'*)

Cando o estudo ten varios arquivos dalgún tipo (por exemplo máis de 1 cuestionario) inclúense todos no mesmo arquivo comprimido. Á súa vez, os microdatos, segundo o seu tamaño en bytes, tamén se poden presentar comprimidos.

Ante calquera problema cos arquivos ou o seu contido pode dirixirse a:

Centro de Investigaciones Sociológicas
Departamento de Banco de Datos
Tel.: +34 91 580 76 17 / 22
Fax: +34 91 580 76 19
Correo electrónico: bancodedatos@cis.es
Horario: de luns a venres de 9.30 a 14.00 horas.

FORMATO DO ARQUIVO DE DATOS

Os arquivos de datos do CIS grávanse en formato ASCII (formato de texto), podendo ser tratados con calquera programa estatístico SPSS, SAS, etc.

A aparencia dun arquivo de microdatos é a dunha matriz de números onde as filas ou rexistros se corresponden cos entrevistados e as columnas coas variables. As columnas que ocupa cada variable veñen especificadas no cuestionario por un número entre paréntese ao lado dereito das categorías de resposta e no libro de códigos.

Antigamente o número de columnas da matriz estaba limitada a 80 polo que se utilizaban tantas filas como fose necesario para completar a información de cada entrevistado. Cada liña de 80 columnas correspóndese cunha TARXETA (RECORD). O cambio dunha tarxeta a outra indícase no cuestionario. A partir do estudo Núm. 2.083 (febreiro 1994), os cuestionarios do CIS deixan de estar estruturados por tarxetas e grávase a información de cada entrevistado nunha única liña.

Os microdatos están anonimizados para os efectos de imposibilitar a identificación individual dos entrevistados. As variables que se ven afectadas neste proceso dependen do estudo en concreto, pero polo xeral son: municipio, comarca, sección, distrito e entrevistador.

FORMATO DO ARQUIVO DE SINTAXE - SPSS

O arquivo de sintaxe preséntase en formato ASCII pero está deseñado para traballar co programa de tratamento estatístico SPSS. Contén as instrucións mínimas necesarias para ler correctamente os microdatos: a definición de variables (DATA LIST) e a ponderación (WEIGHT), cando o deseño mostral así o requira. Os coeficientes de ponderación detállanse na ficha técnica.

Outras instrucións de SPSS non se inclúen por defecto. É o caso das etiquetas das variables (VARIABLE LABELS) e das categorías de resposta (VALUE LABELS), non sempre dispoñibles nin completas, e comandos como RECODE, MISSING VALUES, FREQUENCIES, etc.

Para as enquisas realizadas no período 1992-1997, o programa de sintaxe está escrito de acordo a versións antigas de SPSS, en concreto de UNIX, cuxo tratamento é algo diferente ao de versións máis modernas. Nestes casos haberá que eliminar ou, polo menos, non executar, as instrucións que aparecen por enriba do DATA LIST e poñer os finalizadores de comando que poden ser un punto (.) ou unha liña en branco separando as distintas instrucións.

FORMATO DO ARQUIVO DE SINTAXE - SAS

O arquivo de sintaxe preséntase en formato ASCII pero está deseñado para traballar co programa de tratamento estatístico SAS.

Como mínimo contén as sentenzas necesarias para a correcta lectura dos microdatos. Outras que aparecen no programa son as que etiquetan as variables e as categorías de resposta.