

Nota informativa relativa a los ficheros microdato de la Encuesta de Presupuestos Familiares 80-81 (EPF 80-81)

En la EPF 80-81 se puso a disposición de los usuarios el fichero de uso interno de la encuesta debidamente anonimizado. En la actualidad se utilizan unos ficheros finales que contienen toda la información de la encuesta a partir de los cuales se obtienen tanto las tablas de resultados detallados como los ficheros microdato que se ponen a disposición de los usuarios y en los que se elimina la información de uso interno, se anonimiza a los hogares colaboradores y se agregan las categorías de algunas variables para garantizar la confidencialidad de la información.

Sin embargo, el fichero que está colgado en INEBASE relativo a la EPF 80-81, aunque sí que está debidamente anonimizado y garantiza la confidencialidad incorpora, por un lado, información de gasto desagregada para uso interno y, por otro, información más agregada a nivel de hogar para el usuario.

Así, para utilizar **información agregada por grupo de gasto**, se deben utilizar los totalizadores de los grupos 1 a 9 (posiciones 7707 en adelante), teniendo en cuenta que estos datos están elevados temporalmente pero no espacialmente. Por tanto, para obtener datos anuales y poblacionales, deberemos utilizar el producto del factor de elevación anual (FACELE1, posiciones 7691 a 7698 con cuatro decimales) y los dos multiplicadores (MULT1 y MULT2, posiciones 7645-7646 y 7647-7648 respectivamente). A partir de esta información, se obtendrá, por tanto, información a nivel de hogar por grupo de gasto elevada al año y elevada también a la población.

Si lo que se pretende es trabajar con **información más desagregada por código de gasto** se debe recurrir a la información interna y se precisa además información adicional que permita relacionar los códigos internos de campo del fichero en cuestión con los códigos a 6 dígitos de la clasificación final de gasto que aparecen en las publicaciones en papel.

En el diseño de registro que acompaña a estos ficheros de usuario los códigos internos se leen en cuatro pasos:

- Primeros 99 elementos del grupo 1.
- Restantes 99 elementos del grupo 1.
- Otros bienes de referencia semanal (exceptuando los del grupo 1).
- Información con bienes de los grupos 2 al 9, exceptuando los de referencia semanal.

Esta información a nivel de código está recogida como anotación del hogar en el determinado bien, omitiéndose su periodo de referencia (semanal, mensual,..), y por tanto también su correspondiente factor de elevación temporal.

Por todo lo anterior y para poder trabajar con esta información a un nivel más desagregado se facilita un fichero con la correspondencia entre los códigos de recogida en campo y los códigos finales de la encuesta. Este fichero contiene las siguientes variables:

CODIGO: Código de la clasificación final de gasto.

COD CAMPO: Código interno de recogida en campo. Es el que se encuentra en el fichero microdato.

FACTOR TEMPORAL: Es el valor por el que hay que multiplicar el importe de cada anotación para elevarlo en el tiempo.

PERIODO: Período de referencia del código.

En esta encuesta existen cuatro periodos de referencia: semanal, mensual, bimensual y anual. Para cada uno de éstos existe su correspondiente factor de elevación temporal, los cuales se detallan a continuación:

- Semanal: Factor de elevación 52.
- Mensual: Factor de elevación 12.
- Bimensual: Factor de elevación 6.
- Anual: Factor de elevación 1.

Nota: existe un código de recogida en campo con valor 0 y con un importe de 336 pesetas en el hogar 18036, que debe ser eliminado, ya que es un código no válido que ya ha sido asignado anteriormente al código que le corresponde.

Por último, se debe tener en cuenta que en la utilización de los ficheros microdato, el usuario se ve obligado a guardar determinados compromisos en cuanto a la calidad de la información y la anonimización de los datos. En concreto el punto 4 de las condiciones de compromiso dice lo siguiente:

4. En toda publicación o documento conteniendo resultados obtenidos a partir de estos ficheros se aplicarán las siguientes restricciones en las celdas de clasificación de datos.

No pueden ser publicados resultados relativos a menos de 20 observaciones muestrales.

Entre 20 y 49 observaciones las respectivas celdas deben ser identificadas (marcadas).

Por razones de confidencialidad, en estos casos y en aquellas publicaciones en que se incluyen los tamaños muestrales se mencionará menos de 20 observaciones y entre 20 y 49 observaciones sin especificar el número exacto de observaciones.

Madrid, 16 de septiembre de 2015