

# Metodología de cálculo de las cifras de población censal

En este documento se detalla la metodología empleada para el cálculo de las cifras de población censal obtenidas como resultado de los censos de Población y Viviendas de 2011.

De acuerdo con el Proyecto de los Censos Demográficos de 2011 el Censo de Población y Viviendas 2011 se ha planteado como una operación basada en la combinación de los siguientes elementos:

- Un “**fichero precensal**” realizado a partir de un aprovechamiento máximo de los registros administrativos disponibles, tomando al Padrón como elemento básico de su estructura.
- Un trabajo de campo que incluye dos grandes operaciones:
  - Un **Censo de Edificios** exhaustivo que permite la georreferenciación de todos los edificios.
  - Una **gran encuesta por muestreo**, dirigida a un porcentaje relativamente alto de la población para conocer el resto de características de las personas y las viviendas, la Encuesta de Población.

La combinación del fichero precensal con la información obtenida de la encuesta proporciona toda la información censal. En particular, las cifras de población se obtienen mediante el recuento de los registros contenidos en el fichero precensal, ponderados -cuando sea necesario- con unos *factores de recuento* obtenidos a partir de la encuesta.

---

## La operación de campo: censo de edificios y muestra de viviendas

A finales de septiembre de 2011 se inició el envío de cartas a unos 2,2 millones de hogares solicitando que rellenaran por Internet el cuestionario censal. En noviembre se inició el envío de recordatorios y en diciembre se empezaron a enviar cartas con cuestionarios en papel para aquellos hogares que no habían respondido por Internet.

A finales de noviembre se procedió a la contratación del personal censal, formado por 4.000 agentes, 800 encargados de grupo y 160 encargados comarcales. Las fechas de finalización fueron distintas en cada provincia, pero la media se situó en 17,5 semanas de trabajo (aproximadamente hasta finales de marzo de 2012).

La operación de campo correspondiente al Censo de Edificios se realizó entre diciembre de 2011 y febrero de 2012. De las 28.799 secciones en que estaba previsto realizar este censo se terminaron 27.637 (un 96%). En 7.200 secciones no se consideró necesario realizar este recorrido debido a la calidad de los directorios territoriales de partida y al grado de coincidencia con el Catastro. En esta fase se seleccionó una muestra complementaria de las viviendas que se fueron incorporando como altas.

Finalizado el Censo de Edificios, a lo largo de febrero de 2012, los agentes censales empezaron a entrevistar personalmente a aquellos hogares no

colaboradores hasta el momento, tanto de la muestra inicial como de la muestra complementaria. Se recogieron los cuestionarios mediante el uso de ordenador portátil. También se recogió la información de la muestra de viviendas no principales correspondiente a las 7200 secciones en las que no se realizó el recorrido, y que en las restantes secciones se recogió durante el Censo de Edificios.

La muestra efectivamente recogida asciende a unos **4,2 millones de personas, que se corresponden con 1,65 millones de viviendas principales** y de unas 600.000 viviendas no principales. La recogida por distintos canales para viviendas principales ha sido: **39 % por Internet, 51 % en papel y 10 % por entrevista personal.**

---

## **La obtención de las cifras de población a partir del fichero precensal**

El fichero precensal a 1 de noviembre de 2011, que es la fecha de referencia del Censo, se obtuvo a partir de un aprovechamiento máximo de los registros administrativos disponibles, tomando al Padrón como elemento básico de su estructura.

Las cifras de población se obtienen mediante el recuento de los registros que contiene el *fichero precensal*, ponderados -cuando sea necesario- con unos *factores de recuento* obtenidos a partir de la encuesta.

El objetivo fundamental del fichero precensal era disponer de información suficiente para realizar el recuento de la población y el análisis de su estructura (no solo utilizando las variables más básicas, sino también todas aquellas variables censales que puedan incorporarse a este fichero). Para ello, se partió del Padrón y de la información estadística y de fuentes administrativas (registros del DNI y Tarjetas de Residencia del Ministerio del Interior, nacimientos y defunciones del Movimiento Natural de la Población, Seguridad Social, Agencia Tributaria).

Como consecuencia del cruce del Padrón con estos registros administrativos, para la gran mayoría de las personas que figuran en el fichero precensal se considera que tienen confirmada su residencia en España ya que su identificación se ha conseguido cruzar con estas fuentes y eso significa que son afiliados dados de alta como trabajadores o pensionistas o que figuran en las bases de datos de la Agencia Tributaria. Para estas personas se considera que su residencia en España es segura y se les asigna un factor de recuento igual a 1.

Otros registros tendrán un factor de recuento 0, como los que se hayan encontrado en Defunciones de MNP y que aún figuran en el padrón, generalmente porque se trata de defunciones recientes aún no procesadas.

Asimismo, como consecuencia de este proceso se incorporan al fichero precensal aquellos nacimientos que aún no han sido procesados para su incorporación al padrón pero de los cuales se tiene ya constancia por los boletines del Movimiento Natural de la Población.

Las personas cuya identificación no se ha conseguido encontrar en estos registros (o se ha encontrado pero el motivo de encontrarse no permite asegurar su residencia, por ejemplo si se trata de un inmigrante peticionario de tarjeta sanitaria o de un receptor exclusivamente de rentas del capital o del patrimonio) se consideraron inicialmente dudosas, ya que su residencia en España no podía confirmarse por completo, por lo que su factor de recuento era desconocido.

La realización de la Encuesta de Población ha permitido estimar la proporción de personas dudosas a partir de la información recogida en la Encuesta y esta estimación ha permitido establecer los factores de recuento correspondientes a las personas dudosas.

De esta manera el fichero precensal ponderado por estos factores de recuento, que en su mayoría son 1 pero en ocasiones han resultado inferiores o superiores a la unidad, ha permitido obtener la cifra de población censal.

A este fichero con todos los factores de recuento asignados y que se terminaron de obtener como resultado de la operación muestral, se le denomina Fichero Censal Final Ponderado (FCFP).

---

## **Resultado de los cruces del fichero precensal**

Al cruzar el fichero precensal con los registros administrativos casi un 81% se han encontrado en la Seguridad Social o en la Agencia Tributaria. Hay que considerar que los menores de 14 años en general no cruzan por no figurar en los ficheros externos.

Del conjunto de registros que no se encontraron en otras fuentes no todos se consideraron dudosos, pues algunos no cumplían ningún otro criterio para serlo que el simple hecho de no cruzar (sobre todo los menores). Se definieron una serie de criterios por los cuales un registro, independientemente del resultado de los cruces, se consideraba dudoso. Estos criterios son del tipo: no disponer de identificador válido o confirmado con el Ministerio del Interior (ya sea DNI o tarjeta de residencia), ser un alta padronal posterior al censo de 2001 de un español adulto que no procede del extranjero, ser un extranjero para el que han transcurrido 6 meses desde su última renovación padronal (distinguiendo si no llega a los 2 años o supera este tiempo), etc.

Finalmente del total de registros del fichero precensal un 97,2% tuvo factor de recuento 1, un 0,6% tuvo factor de recuento 0 y un 2,2% tenía factor de recuento a asignar de acuerdo con los resultados de la Encuesta.

En definitiva **el número de registros dudosos** a los que hubo que asignar factor de recuento **fue cercano a 1.040.000 personas, de ellas el 87% extranjeros.**

---

## **Partición del fichero precensal en clases**

Una vez determinados los registros del fichero precensal que se consideraron finalmente dudosos, y a los que había que asignar un factor de recuento, se realizó una partición del fichero completo en *clases* o grupos de población definidos por características sociodemográficas. Estas características que determinan las clases son la edad, según intervalos de edades agrupadas, la nacionalidad y la provincia de residencia.

En cada clase puede haber registros dudosos, con factor de recuento a asignar, y seguros (con factor de recuento 1). La mayoría de las clases contienen población dudosa y ésta se estima a partir de la muestra, pero si una clase no tiene registros dudosos no es necesario realizar estimaciones según la muestra y para esa clase la población coincidirá con la que tenga en el FPC.

Las clases se configuraron de forma que al menos contuvieran 1000 registros dudosos y que todo registro dudoso perteneciera a una clase. Las clases que resultan con menos de 1000 dudosos se agruparon con el objetivo de llegar al menos a 1000 dudosos. La forma de agrupar ha sido, para una misma nacionalidad y provincia, considerar mayores intervalos de edad, si esto fuera insuficiente se ha agrupado para esa misma nacionalidad a nivel de la Comunidad Autónoma y si aún así no se consigue el mínimo se agrupaba para dicha nacionalidad a nivel nacional, por último y residualmente se llegaba a agrupar nacionalidades a nivel nacional. De esta forma el fichero precensal quedó configurado en 724 clases con al menos 1000 dudosos cada una.

Se estableció el mínimo de 1000 dudosos para tener garantías de que se encontrasen representados en la Encuesta ya que esta iba dirigida a un 12% de la población.

Aunque las nacionalidades se han considerado individualizadamente en el caso de los españoles se han distinguido los nacidos en España de los nacidos en el extranjero como si se tratara de nacionalidades diferentes.

---

## **Asignación de la muestra a clases y enlace nominal con el fichero precensal**

Al igual que el fichero precensal se ha dividido en clases, la muestra recogida se asigna a las clases a que pertenezca en función de las características que definen éstas. Además se realiza un enlace nominal entre los registros de la muestra recogida y el fichero precensal para determinar si los registros de la muestra están clasificados como dudosos.

Las características que definen las clases (edad, nacionalidad, provincia de residencia) son características observables al realizar la Encuesta ya que se recogían en el cuestionario. Sin embargo la característica de dudoso no es observable ya que depende fundamentalmente del resultado de los cruces con los registros administrativos. Por este motivo se efectuó un cruce nominal

(nombre, apellidos, sexo, fecha y lugar de nacimiento) entre la muestra recogida y el fichero precensal de manera que se pudieran enlazar los registros muestrales con los del fichero precensal y así poder determinar los que, si bien pudieran figurar en el FPC como dudosos, fueron encontrados en la realidad ya que se recogieron en la muestra. Este enlace nominal resultó del 96% (del 88% si solo se consideran extranjeros), por lo que para la gran mayoría de registros de la muestra se conoce si figuran como dudosos en el FPC.

Se asociaron los registros de la muestra a las clases predefinidas por edad-nacionalidad-provincia y se estimó la proporción de dudosos de cada clase con respecto al total de la clase. A efectos de determinar la situación de los no enlazados se efectuó un reparto proporcional entre seguros y dudosos aplicándoles un factor de corrección calculado como la inversa del porcentaje de enlace.

Para mejorar las estimaciones en lugar de utilizar el factor general de corrección derivado de la calidad del enlace, se utilizaron los factores de corrección correspondientes a cada clase ya que los enlaces son de mejor o peor calidad en función de las características de la clase (edad, nacionalidad y provincia).

---

## **Estimación de la proporción de dudosos por clases y determinación del factor de recuento**

Utilizando la información recogida en la Encuesta de Población se ha estimado, para cada clase separadamente, la proporción de registros dudosos que se encuentran en la realidad, permitiendo calcular un factor de recuento para los registros dudosos de la clase, factor que puede ser mayor, menor o igual a 1.

Una vez estimada la proporción "p" de dudosos de una clase, como se conoce la población segura "S" (los que tienen factor de recuento 1) y la población dudosa "D" (los que tienen factor de recuento a asignar), se calcula el factor de recuento FR como  $(S \cdot p / (1 - p)) / D$ .

De esta manera el fichero precensal se convierte en un *Fichero Censal Final Ponderado* incorporando esos factores de recuento y permite obtener las cifras de población censales.

Este factor de recuento, obtenido para cada clase, es el que se aplica a los registros del Fichero Precensal con factor de recuento no conocido pertenecientes a la clase. Al final de este proceso todos los registros tienen asignado un factor de recuento en el Fichero Censal Final Ponderado (FCFP).

Este procedimiento justifica, desde un punto de vista estadístico, la cifra final de población censal, pero como es natural no puede ser utilizado para justificar la misma desde un punto de vista administrativo, ya que no será posible determinar a que ciudadanos concretos se está contando y a cuales no. La población de una determinada área geográfica se obtiene como suma de los factores de recuento de los registros del FCFP en dicha área.

## Cuantificación de la población dudosa

Como resultado de los cruces del fichero precensal con los registros administrativos se han obtenido algo más de 1.040.000 registros dudosos cuyo factor de recuento se ha determinado en función de los resultados de la fase de encuesta de población.

**El factor de recuento medio ha resultado ser 0,424, lo que implica que los 1.040.000 dudosos se contabilizan como una población de 440.000 personas aproximadamente.**

El 68,4 % de los dudosos se consigue estimar definiendo clases a nivel de nacionalidad y provincia. Si se consideran también las clases a nivel de nacionalidad y CCAA llegan a englobar el 80% de los dudosos.

Las provincias que presentan una mayor proporción de dudosos son Alicante con un 7%, Málaga con un 4,7%, Almería con un 4,2%, Baleares con 4% y Tenerife con un 3,9%, siendo la media nacional del 2,2%.

Entre las nacionalidades con mayor representación en el conjunto de la población las que presentan un mayor porcentaje de dudosos son Paraguay con el 35´4%, Alemania con el 30%, Reino Unido con el 28´4%, Brasil con el 27% y Francia con el 22,8%.

La clase más destacada, en cuanto a proporción de población dudosa, es la de nacionales de Alemania mayores de 80 años en la provincia de Alicante, formada por cerca de 3.000 personas de las cuales más del 70% se consideran dudosos. La más numerosa, en cuanto a dudosos, es la de ciudadanos del Reino Unido también en Alicante, con edades entre 60 y 65 años, que suponen casi 6.000 dudosos.

A continuación se presenta una tabla resumen con los factores de recuento medio obtenidos:

| Factor de recuento | Nº de clases | FR medio |
|--------------------|--------------|----------|
| FR<=0,7            | 619          | 0,34     |
| 0,7<FR<=0,9        | 65           | 0,78     |
| 0,9<FR=<1          | 18           | 0,95     |
| FR>1               | 22           | 1,38     |
| Total de clases    | 724          | 0,42     |

Entre las clases que obtienen un factor de recuento superior a 1 se encuentra la correspondiente a los nacionales de Pakistán entre 25 y 30 años, con un factor de recuento de 1,67.