

INSTITUTO NACIONAL DE ESTADÍSTICA



Censos de Población y Viviendas 2021

Metodología.

Versión definitiva

Subdirección General de
Estadísticas Demográficas

Junio 2023

Índice

1	Nota sobre la versión provisional de la metodología	4
2	Origen de la demanda y justificación de su necesidad	4
3	Objetivos del proyecto	5
4	Un censo basado en registros administrativos	6
4.1	Antecedentes: El censo de 2011 y el contexto internacional previo al censo 2021	6
4.2	Los trabajos previos para el censo de población y viviendas 2021	10
4.2.1	La elaboración progresiva de ficheros precensales	12
4.2.2	Elementos estratégicos del censo de 2021	12
5	El censo de población: personas y hogares	14
5.1	El primer objetivo: contar la población	14
5.1.1	Los datos sobre población en España desde 1996	15
5.1.2	La cifra de población del censo 2011	17
5.1.3	Las cifras de población desde 2012	18
5.2	La cifra censal y el método de signos de vida	18
5.3	Variables relativas a la población	26
5.3.1	Variables demográficas (sexo, edad, nacionalidad, lugar de nacimiento)	28
5.3.2	Variables relacionadas con las migraciones	28
5.3.3	Estudios en curso y nivel de estudios alcanzado	35
5.3.4	Lugar de estudios	45
5.3.5	Relación con la actividad económica	47
5.3.6	Variables relativas a los ocupados	52
6	Composición de hogares / relaciones de parentesco	66
7	Censo de viviendas	73
7.1	Aspectos generales del censo de viviendas	73

7.2	Proceso de obtención del directorio de viviendas	75
7.3	Asignación de sección censal a cada una de las viviendas	81
7.4	Enlace personas y viviendas	82
7.5	Problemas detectados en la elaboración del fichero censal de viviendas	84
7.6	Identificador único	87
7.7	Variables de vivienda	87
8	Clasificación de las viviendas según uso a partir del consumo eléctrico	91
8.1	Viviendas principales y no principales	91
8.2	Clasificación de viviendas según consumo eléctrico	92
9	Sistema de difusión de resultados censales	94
9.1	Líneas generales del sistema de difusión censal (sdc-2021)	95
9.2	Preservación del secreto estadístico en el sdc-2021	97
9.3	Ficheros de microdatos	99
9.4	Publicación de resultados censales por parte de eurostat	99
10	Calendario de difusión	100
11	Más allá de 2021: el futuro de los censos de población	100
11.1	El marco europeo	101
11.2	Estrategia post-2021 en España	101

1 Nota sobre la versión provisional de la metodología

Esta versión de la metodología del Censo de Población y Viviendas 2021 se actualizará y completará en los próximos días con información sobre hogares y población residente en establecimientos colectivos.

2 Origen de la demanda y justificación de su necesidad

Los Censos de Población y Viviendas constituyen la operación estadística de mayor arraigo en la estadística pública; también son las operaciones más complejas y de mayor envergadura, razón por la cual dan nombre a muchos de los institutos de estadística de mundo.

En España, el primer censo moderno de población, entendiendo como tal el que utiliza a la persona como unidad de análisis, se realizó en 1768 por el Conde de Aranda bajo el reinado de Carlos III. También son de destacar por su interés el Censo efectuado en 1787 por Floridablanca y el realizado diez años más tarde por Godoy en tiempos de Carlos IV.

No obstante, la serie de censos de la organización estadística oficial se inicia en 1857 con el primero de la Comisión General de Estadísticas del Reino, al que siguió, en un lapso inusualmente corto, el de 1860. Después vinieron los de 1877, 1887 y 1897. A partir de 1900 ha habido Censo de Población cada diez años sin excepción alguna.

En definitiva, el Censo de Población de 2021 ha sido el decimoctavo de los Censos oficiales realizados en España.

Su realización se enmarca dentro del Programa Mundial 2020 que abarca el periodo 2015-2024 promovido por Naciones Unidas. El año 2021 ha sido el año elegido como referencia en el caso de los países europeos.

Como ya ocurriera con la edición de 2011, los censos en la Unión Europea (UE) se han realizado con una reglamentación comunitaria. El reglamento 763/2008 del Parlamento Europeo y del Consejo (junto con otros que lo desarrollan), además de implantar la obligatoriedad de realizar el Censo cada diez años (afecta a los censos de 2011 y 2021), asegura la comparabilidad de los resultados a nivel de la Unión Europea por lo que respecta a la metodología, definiciones y el programa de datos y de los metadatos estadísticos asociados.

Si bien el Reglamento 763/2008 del Parlamento Europeo y del Consejo ha servido de base, sin modificaciones, para los censos de 2011 y 2021, todos los reglamentos de ejecución que establecen los detalles de la operación se han revisado respecto de la edición de 2011. Los reglamentos de ejecución actualmente en vigor son:

- 2017/0881R Reglamento de Ejecución de la Comisión de 23/05/2017. Establece las modalidades y estructura de los informes sobre la calidad y formato técnico para la transmisión de los datos.
- 2017/0543R Reglamento de Ejecución de la Comisión de 22/03/2017. Establece especificaciones técnicas de los temas y sus desagregaciones.

- 2017/0712R Reglamento de la Comisión de 20/04/2017. Establece año de referencia y el programa de los datos y de los metadatos estadísticos.
- 2018/1799 Reglamento de la Comisión de 21/11/2018 relativo al establecimiento de una medida estadística directa temporal para la difusión de temas seleccionados del censo de población y vivienda de 2021 georreferenciados en una malla de 1 km².

3 Objetivos del proyecto

Aunque son muchos los objetivos que cubren los Censos demográficos, con el paso del tiempo estos han ido cambiando. Entre los objetivos que pretende cubrir el censo de 2021 cabe destacar:

- a) Recuento de la población. El Censo de Población determina el número de habitantes de España y de todas sus divisiones administrativas (las comunidades autónomas, provincias, municipios e incluso distritos y secciones censales). Si bien es cierto que ya existe un recuento de la población anual por medio de las cifras oficiales de población para los municipios, a partir del Padrón, el Censo realiza ese recuento con fines y metodología distintos del padrón.
- b) Conocimiento de la estructura de la población. El Censo de Población no solo cuenta el número de habitantes sino que investiga, para cada uno de ellos, una serie de características demográficas, económicas y sociales y proporciona una imagen de la estructura de la población.
- c) Proporciona información de base para la elaboración de estadísticas por muestreo. Los Censos de Población y Viviendas han suministrado tradicionalmente el marco para las encuestas por muestreo. La información obtenida en el censo de 2021 puede, de nuevo, cumplir esta función, al proporcionar información exhaustiva de población y viviendas.
- d) Recuento de las viviendas. El Censo de Viviendas, permite determinar el número de viviendas existentes y algunas características de las mismas así como de los edificios donde se encuentran.
- e) Atender las necesidades estadísticas de ámbito internacional. Los organismos internacionales solicitan información de los diferentes países para elaborar sus estadísticas demográficas y sociales, siendo los Censos una de las principales fuentes utilizadas.

Por último, ciertas funciones tradicionalmente encomendadas a los censos, como la actualización y normalización de instrumentos de infraestructura estadística (los callejeros, la cartografía), y en cierto modo el marco para el muestreo, que se renovaban o mejoraban con ocasión de la operación censal ya no precisan de los censos dado que se realizan mediante otras operaciones, con mucha mayor precisión y actualidad.

4 Un censo basado en registros administrativos

4.1 ANTECEDENTES: EL CENSO DE 2011 Y EL CONTEXTO INTERNACIONAL PREVIO AL CENSO 2021

El censo de población es seguramente la operación estadística que se realiza en un mayor número de países del mundo. Según datos de Naciones Unidas, en la ronda 2005-2014, un total de 214 países (que representaban un 93% de la población) realizaron un censo, la mayoría en los años 2010 y 2011, mientras que solo 21 países, muchos de ellos en situaciones políticas inestables o de conflicto, no lo pudieron llevar a cabo.

En España, en 2011 se realizó un censo de población que introducía una importante novedad respecto de todos los anteriores: por primera vez no se contemplaba en el proyecto el envío de cuestionarios a todas las viviendas del territorio nacional sino solo a una muestra de ellas, cercana al 10%.

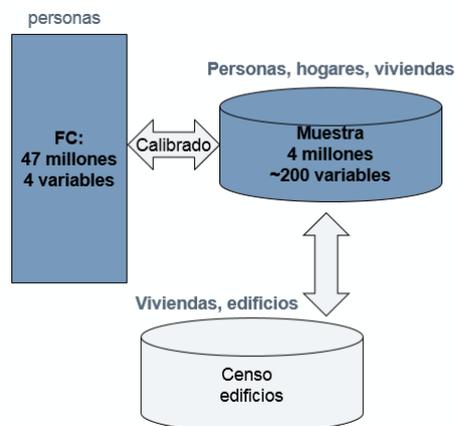
La operación de 2011 dejó de ser un “censo” desde la perspectiva del **proceso** seguido para obtener los datos, si bien sí puede considerarse que, al menos para una parte de las variables, el **producto** obtenido era un verdadero censo, en tanto que los datos procedían de un fichero que contenía todas las personas.

Siguiendo la clasificación sobre métodos censales que establecía el Reglamento 763/2008, el censo de 2011 debe considerarse un “censo combinado” en el que entran en acción registros administrativos y una encuesta, como se observa en la figura 1. Por un lado se construyó, mediante combinación de ficheros administrativos, el llamado “**fichero censal**” de población, que era un registro de población formado por casi 47 millones de individuos.

Por otro lado, se realizó una encuesta a una muestra de hogares, que recopiló información directa, mediante cuestionarios, de algo más de 4 millones de personas.

Además se realizó una operación de **censo de edificios**. Se recorrió todo el territorio nacional con el fin de identificar todos los edificios que contenían viviendas. En realidad, un 20% de las secciones censales no se visitaron porque se disponía de una información de partida suficientemente buena para ellas, lo que hacía innecesario el trabajo de campo.

Figura 1. Esquema global del Censo de Población y viviendas 2011



Este censo de edificios, que consumió la mitad del presupuesto global, era necesario para poder completar un marco de edificios y viviendas de calidad. Ya se disponía de un marco, a partir del censo de 2001 y de diversos métodos de actualización (sobre todo de la información proveniente de Catastro y del propio padrón continuo del INE). Pero era necesario completar esa información con un recorrido a pie de calle que permitiera resolver información incompleta y también proporcionar coordenadas geográficas a muchos edificios para los cuales no se disponía de ese dato.

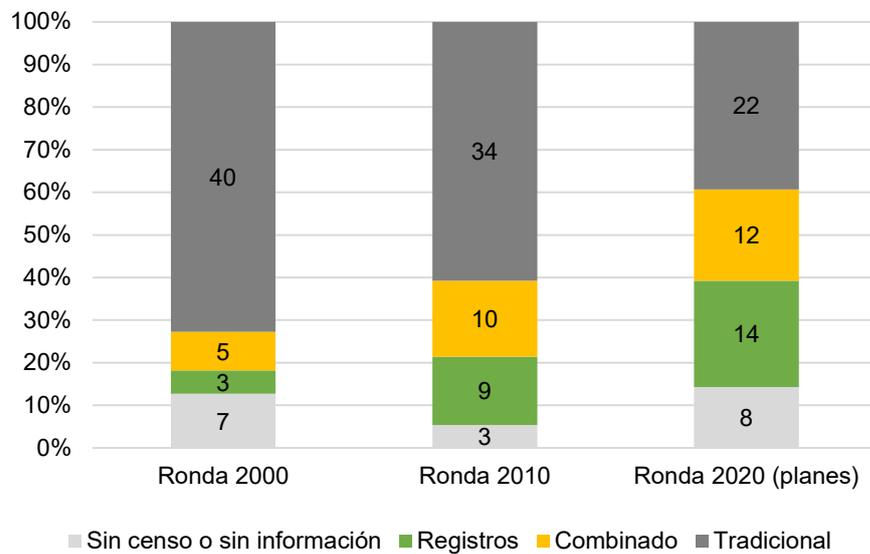
España no fue, ni mucho menos, el único país que renunció en 2011 a recoger cuestionarios de todos los hogares. Aunque el censo tradicional siguió siendo el método más frecuente (fue el seguido por un 85% de países), su importancia ha perdido peso en los últimos años.

Durante los años posteriores al censo de 2011 la tendencia a construir la información censal sin necesidad de enviar cuestionarios a toda la población, iniciada por algunos países pioneros en los años 80, ha ido creciendo paulatinamente. Se hace todavía más relevante en el ámbito UNECE, que está formado actualmente por 56 países¹, entre los que se encuentran muchos de los más desarrollados del mundo.

Así, según una encuesta realizada por UNECE en abril de 2015 sobre la metodología prevista para llevar a cabo la ronda de 2020 (comparada con las dos ediciones anteriores) se tenían los siguientes resultados:

Figura 2. Métodos censales en distintos países (Fuente: UNECE)

¹ UNECE está formado por todos los países europeos, Turquía, Rusia y todos los estados post-soviéticos, Canadá, Estados Unidos e Israel. En la ronda del 2000 aparecen solo 55 países debido a que Montenegro pasó a formar parte de UNECE en el año 2006.



A lo largo de las dos últimas décadas el número de países en el ámbito de UNECE que realizan un Censo tradicional prácticamente se ha reducido a la mitad mientras que el número de los que optan por basarlos en registros o mediante métodos combinados no ha parado de crecer.

En innumerables foros internacionales se constata el interés de todos los países por ir abandonando los métodos tradicionales de recogida de los censos e ir incorporando información de registros administrativos u otras fuentes, si bien son muy pocos los países que se muestran en condiciones de dar el salto a un censo que prescindiera de la recogida directa de información de los hogares y que base por tanto la información en tales registros.

Dentro del ámbito de la UE, como se observa en la figura 3, España se ha sumado en 2021 al reducido grupo de 8 países que han realizado el censo basándose en registros administrativos. Además hay doce países que han pasado del Censo tradicional en la ronda de 2011 al Censo combinado en la ronda de 2021. Apenas seis países de la UE mantienen métodos tradicionales.

Figura 3. Métodos de recogida de los Censos de Población en la UE-27 (2001, 2011, 2021)



Fuente: European Parliamentary Research Service (EPRS) a partir de datos de UNECE y de los INEs de Estados Miembros de la UE

Cabe comentar aquí que las decisiones finales sobre la metodología censal en España se tomaron mucho antes de que en marzo de 2020 surgiera el brote de COVID-19 que ha supuesto un enorme desafío a los institutos de estadística de todo el mundo. En efecto, los problemas derivados del confinamiento de la población, la población desplazada de su lugar habitual de residencia, la imposibilidad de realizar operaciones de campo o de entrevistar hogares durante muchos meses, entre otras muchas, han significado muchos retrasos en la ronda censal en todo el mundo.

En una encuesta realizada por la División de Estadística de Naciones Unidas a las oficinas de estadística¹ de todo el mundo, solo 15 países, entre ellos España, respondieron que el COVID-19 no afectaba a los planes de realización del censo. Cabe destacar que en el ámbito europeo muchos países han debido retrasar la operación y de hecho tres países, entre ellos Alemania, no han podido realizar el censo en 2021 y han tenido que hacerlo en 2022.

Esta circunstancia ha acrecentado sin duda el interés en todo el mundo por los métodos no tradicionales para confeccionar censos y probablemente se experimentará una

¹ Más información en <https://unstats.un.org/unsd/demographic-social/census/COVID-19/>

transición más rápida a censos basados en registros administrativos para la ronda 2030 de la que ya se aventuraba hace unos años.

4.2 LOS TRABAJOS PREVIOS PARA EL CENSO DE POBLACIÓN Y VIVIENDAS 2021

En el censo de 2011, si bien se disponía de dos fuentes con datos de personas –el fichero censal y la muestra–, el fichero censal solo se utilizó para obtener el recuento de población por municipios. Todas las explotaciones detalladas del censo se construyeron a partir de los cuestionarios de la muestra de hogares, que a efectos de los usuarios fue el único producto censal. Esta limitación se debió al hecho de que el fichero censal apenas tenía algunas pocas variables para cada persona: las que figuran en el padrón (sexo, edad, lugar de nacimiento, nacionalidad, clasificación básica de nivel de estudios).

Como planteamiento inicial para el censo de 2021, el objetivo que se marcó el INE al terminar los trabajos del censo 2011, fue estudiar hasta qué punto se podía aumentar el número de variables obtenidas a partir de distintos registros administrativos y disminuir, por tanto, la información contenida en los cuestionarios censales.

El análisis de fuentes no fue inmediato porque son muchas las variables recogidas en un censo como este, que no solo censa población sino también edificios, viviendas y hogares y que además recoge para cada una de estas unidades distintos tipos de información. Algunas de las variables son muy fáciles de obtener de registros, pero en otros casos el reto es mucho mayor.

Durante los años 2014 y 2015 se inició la investigación. Se contactó con multitud de departamentos de la Administración y se recibieron decenas de ficheros distintos con los que componer la información censal.

La conclusión a la que se pudo llegar ya desde finales de 2015 es que se disponía de mucha más información y de mucha mayor calidad de la que se preveía al inicio del análisis: con seguridad se podía construir más del 90% de la información que contenía el cuestionario usado en la muestra del censo de 2011 a partir de registros administrativos.

Esto implicaba ya una primera decisión estratégica en la mejora de calidad respecto del censo de 2011. Si el número de variables que puede obtenerse de registros administrativos es muy superior al de 2011, la muestra, tal cual estaba diseñada en 2011, con el objetivo de extraer de ella toda la información para toda la población, dejaba de tener sentido.

Los esfuerzos se centraron entonces en analizar la posibilidad de prescindir por completo de cuestionarios censales y construir el censo íntegramente a partir de fuentes administrativas.

En la primera mitad de 2017 se elaboró el Anteproyecto de Censos de Población y Viviendas 2021. En ese documento figuraban unas líneas estratégicas para el proyecto. Ya entonces se contemplaba la idea de basar el censo casi exclusivamente en datos administrativos, quizá combinado con alguna encuesta para completar lagunas de información.

Durante los años 2016 y 2017 se realizó un estudio variable por variable, denominado fichero precensal 2016 (FPC-2016). Este trabajo se completó a finales de 2018. Sus

conclusiones, que no estaban aún disponibles cuando se redactó el anteproyecto, permitieron avanzar algo más en la estrategia. Cabía destacar las siguientes:

- ✓ El censo de población **era perfectamente factible** basándose exclusivamente en registros administrativos, es decir, no era necesario enviar ningún cuestionario a los hogares ni solicitar colaboración de ninguna otra forma. Sería necesario completar la información con procedimientos estadísticos de imputación en ciertos casos, pero no en mayor medida de lo que se utiliza en censos clásicos.

El requisito mínimo, que es la obligación legal establecida por el Reglamento Censal¹ de la UE se cumplía sobradamente para todas las variables referidas a la población.

- ✓ Si bien subsistía antes de ejecutar el FPC-2016 la duda sobre la posibilidad de realizar operaciones de campo dirigidas a mejorar la información de registros administrativos en zonas específicas (zonas marginales, zonas con abundancia de población extranjera comunitaria) se descartó por considerar que serían operaciones muy poco o nada eficientes, que no mejorarían lo obtenido con registros más imputación.
- ✓ Se consideraba conveniente la realización de una encuesta complementaria, por muestreo, dirigida a toda la población de un tamaño cercano al 1% de la población. Esta operación tendría varios objetivos:
 - Mejorar las evidencias sobre ciertas variables (como el régimen de tenencia o las formas de convivencia en el hogar) como apoyo para la imputación necesaria en el censo.
 - Proporcionar ciertas variables que se deben publicar con nivel de detalle provincial o superior (como las del equipamiento de las viviendas) y para las que, por tanto, no es necesario recabar información exhaustiva.
 - Proporcionar algunas otras variables que un censo basado en registros no puede dar actualmente, pero que no forman parte de los requisitos imprescindibles del censo (uso de lenguas, movilidad cotidiana...)

Esto implicaba una decisión estratégica de gran calado. El INE se sumaba así a la decena de países (el mayor de los cuales, Holanda, tiene 17 millones de habitantes) que abandonan los censos basados en cumplimentación de cuestionarios por parte de los ciudadanos. El Censo de Población y Viviendas se convertía en una operación invisible para la sociedad en general, que se iba a realizar con un coste cualitativamente inferior, asumiendo muchos menos riesgos de los que supone una operación de recogida masiva de encuestación a los hogares ofreciendo unas garantías en cuanto a cobertura y calidad muy superiores.

Cabe considerar que el censo de 2021 en España ha sido el fruto de un proceso que se inicia en 1996 con el padrón continuo. Se debe entender, por tanto, como la culminación de un largo camino de 25 años. Además, se ha podido realizar porque en España confluyen otros dos elementos fundamentales que solo se dan en unos pocos países: abundancia de registros administrativos de calidad y base legal clara y suficiente para acceder a ellos.

¹ A lo largo del documento se menciona en muchas ocasiones, por simplificar, el “Reglamento censal”. Se trata, en realidad, del conjunto de reglamentos mencionados en el Apartado 2.

4.2.1 La elaboración progresiva de ficheros precensales

La elaboración del censo de 2021 se inicia con el fichero precensal 2016 (FPC-2016) que consistió en la construcción de una primera versión de lo que se podría considerar un censo de población completo basado íntegramente en registros.

Como ya ocurriera con el censo de 2011, el elemento principal sobre el que se construye todo el proyecto es el padrón continuo. El FPC-2016 parte, esencialmente, del padrón a 1 de enero de 2016, que constituye un primer esqueleto de fichero de población. A este fichero se le enlazan registros administrativos de diversa naturaleza con el fin de componer una información individual similar a la que se obtendría con un cuestionario dirigido a los hogares.

El FPC-2016, como primer “ensayo general” de un censo de población basado en registros era aún un producto incompleto. No se disponía de todos los registros administrativos: algunos de ellos se habían recibido demasiado tarde como para formar parte de ese primer análisis; en otros casos se comprobó que la petición de información por parte del INE podía redefinirse y mejorarse o podía recoger más variables; quedaban además ficheros ya detectados pero aún por incorporar o mejoras de cobertura y puntualidad de otros.

Además, el FPC-2016 era un ensayo del censo de población pero no aún del censo de viviendas. De hecho, el FPC-2016 inicia el estudio de un *Censo de Viviendas* pero no se trabajó con el objetivo de construir un fichero precensal de viviendas 2016. El “ensayo general” sobre viviendas se dejó para una fase posterior, con la construcción del fichero precensal 2018, que ya aborda un censo de viviendas completo.

A lo largo de los siguientes años se trabajó en la incorporación paulatina de mejoras generándose ficheros precensales anuales. Así, se construyó un FPC-2017 pero que solo se centró en algunas variables. El siguiente fichero completo fue el FPC-2018, que incorporaba ya muchas de las posibles mejoras identificadas en el FPC-2016 y, sobre todo, realiza por primera vez una aproximación completa al Censo de viviendas.

Se generaron también ficheros referidos a 2019 y 2020, luego todos los procesos de enlace y depuración de la información se han ido mejorando durante los últimos años. El Censo de 2021 no parte de cero sino que es una actualización de los procesos ya probados en los sucesivos ficheros precensales por lo que los riesgos de la operación censal de 2021 se minimizaron.

4.2.2 Elementos estratégicos del censo de 2021

Si hablamos de estrategia para construir un censo no se puede soslayar el elemento principal: la fuente principal de información sobre la población será el padrón continuo que se completará con otros muchos registros, pero usar el padrón como esqueleto supone tomar dos decisiones estratégicas:

- **Se considera exclusivamente población que, bien está empadronada, bien lo ha estado en algún momento**, y por tanto figura en la “base padronal”.
- **La población se contará en el lugar dónde está empadronada**. No se contempla modificar la residencia de las personas por la incorporación de datos de residencia

de otros ficheros salvo en casos muy particulares, como por ejemplo cuando se encuentra en un domicilio un menor supuestamente viviendo solo, sin adultos.

Más allá de esta cuestión, se resumen a continuación las principales características del censo 2021:

– **Se ha construido un verdadero “producto censal” referido a 1 de enero de 2021:**

El censo de 2021 está referido **a la fecha 1 de enero**. La operación censal da como resultado un conjunto de ficheros con datos individuales:

- un fichero completo de personas formado por tantos registros como habitantes
- un fichero de hogares que contiene a todos los residentes en viviendas familiares principales (es decir, donde reside población).
- un fichero de establecimientos colectivos que contiene a la población residente en ellos
- un fichero de viviendas, distinguiendo viviendas principales y no principales.

En el momento de publicar este documento de descripción metodológica del Censo (noviembre de 2022) solo se ha publicado una parte de la información relativa a personas. Los datos sobre hogares, establecimientos colectivos y viviendas se publicarán durante la primera mitad de 2023.

A partir de estos ficheros se calculan todos los resultados agregados que componen los productos de difusión del censo.

El producto que se ha construido se basa, por tanto, en microdatos que recogen a toda la población. Es decir, es un censo exhaustivo, construido con ficheros que contienen tantos registros como individuos (personas, hogares, viviendas) forman el colectivo estudiado. En ese sentido, es similar al que se dispondría en caso de procesar los cuestionarios de todos los hogares de España mediante un censo clásico.

Merece la pena destacar algunos elementos:

– **La información proviene casi exclusivamente de registros administrativos pero se completa a partir de encuestas ya existentes...**

Esta es la solución que ya se utiliza en los países que basan sus censos en registros administrativos para algunas variables que no se encuentran, o no se construyen del todo bien a partir de registros.

Por ejemplo, el tipo de hogar es una variable difícil de construir a partir de registros. En algunos casos es muy sencillo: hogares unipersonales o parejas casadas con o sin hijos. Suponen la mayoría de casos. En otros, el uso de fuentes administrativas como datos tributarios, permisos de residencia de extranjeros, datos de filiación del DNI, Registro Civil,...) permiten resolver la incógnita sobre parentescos.

Pero subsiste un grupo muy numeroso de hogares para los que no es posible determinar vínculos entre sus miembros. La solución utilizada por otros países en esos casos es la imputación probabilística para estos hogares a partir de datos de encuestas existentes. En ese sentido el INE dispone de muestra suficiente tomando datos de la

Encuesta de Población Activa (EPA) o la Encuesta Continua de Hogares (ECH), incluso acumulando muestra, de uno o varios años, de ambas.

– **y se ha realizado una encuesta específica para complementar la información sobre algunas variables.**

El proyecto censal se ha complementado con la realización de una encuesta específica, denominada Encuesta sobre Características Esenciales de la Población y las Viviendas (ECEPOV-2021). Se recogió durante 2021 y permite mejorar la imputación de algunas variables censales y proporcionar algunas otras que no se encuentran en registros administrativos. Se trata de una encuesta por muestreo dirigida aproximadamente a un 1% de la población, un tamaño muy inferior al de la encuesta de 2011 que fue de casi el 10%.

– **Se incorporan nuevos temas y nuevas formas de medir algunas variables.**

El uso de datos administrativos abre la puerta a publicar nuevas variables que no se habían recogido en censos anteriores por no sobrecargar el cuestionario o por las dificultades que suponía. Nos referimos, por un lado, a variables que se podrían recoger en el censo a nivel individual, como por ejemplo, la fecundidad. Se incorporará, durante 2023, información sobre los hijos tenidos desde 1996, primer año desde el cual se puede enlazar con información de nacimientos.

También habrá información nueva que, por diversas razones, no se va a poder incorporar de forma individual pero sí agregada, aunque con mucho detalle. Por ejemplo, no se puede proporcionar individualmente el consumo eléctrico de las viviendas o el nivel de ingresos del hogar o de la persona pero sí se proporciona el **nivel medio de ingresos de cada sección censal** o municipio o datos sobre **consumo eléctrico** para ámbitos geográficos muy reducidos (normalmente secciones censales). El consumo eléctrico aporta una nueva forma de estudiar las viviendas vacías que se detallará en la versión definitiva de este documento metodológico.

– **El censo 2021 es el inicio de un nuevo sistema de producción de información demográfica continua y actualizada**

El censo de 2021 es el último censo tal como se contempla ahora, en el sentido de ofrecer una foto detallada del territorio una vez cada diez años; junto con el censo de 2021 se ha diseñado una estrategia para ofrecer información del mismo nivel de detalle pero de forma más periódica, un “censo anual”. Esta nueva estrategia se describe con mayor detalle en el apartado 9.2

5 El censo de población: personas y hogares

5.1 EL PRIMER OBJETIVO: CONTAR LA POBLACIÓN

El objetivo original de los censos es contar el número de habitantes que reside en un territorio en una fecha de referencia, que en este caso es el **1 de enero de 2021**. Algunas variables en concreto pueden tener periodos de referencia específicos. Así, las variables relativas al mercado laboral estarán referidas a la semana anterior a la fecha de referencia (la última semana de 2020).

En la actualidad en España este objetivo principal pierde importancia, pues debido a la existencia del padrón continuo no se tiene la incertidumbre en las cifras de población que se daba en el pasado y que todavía ocurre en otros países durante los periodos intercensales, que normalmente son de diez años.

En efecto, en muchos lugares del mundo, durante el periodo que media entre dos censos, no hay más que unas proyecciones basadas, en algunos casos, en muy pocas evidencias. En España, la existencia de un registro de población permanentemente actualizado, como es el padrón continuo, hace que se puedan proporcionar cifras de población con mucha frecuencia y que, por tanto el objetivo original del censo, actualizar las cifras de habitantes, se convierta casi en secundario.

Estrictamente hablando, en un censo de población se deben incluir únicamente las personas, de cualquier nacionalidad, que tienen fijada su “**residencia habitual**” en el territorio nacional.

El concepto de “residencia habitual” figura definido en los manuales de recomendaciones de Naciones Unidas pero el Reglamento 763/2008 del Parlamento Europeo y del Consejo lo precisa algo más. En concreto, se establece lo siguiente:

«Residencia habitual»: el lugar en que una persona pasa normalmente el período diario de descanso, independientemente de ausencias temporales con fines de ocio, vacaciones, visitas a amigos o parientes, negocios, tratamiento médico o peregrinaje religioso.

Deberán considerarse residentes habituales de la zona geográfica en cuestión únicamente:

- i) aquellas personas que hayan vivido en su lugar de residencia habitual durante un período ininterrumpido de al menos doce meses antes de la fecha de referencia, o*
- ii) aquellas personas que hayan llegado a su lugar de residencia habitual durante los doce meses anteriores a la fecha de referencia con la intención de permanecer en él durante al menos un año.*

Cuando no se puedan determinar las circunstancias descritas en los incisos i) o ii), se entenderá por «residencia habitual» el lugar de residencia legal o registrado;

Por tanto, expresamente se contempla en el Reglamento (y en las recomendaciones de Naciones Unidas)¹ la opción de contar como población residente la población registrada (empadronada, en el caso español).

5.1.1 Los datos sobre población en España desde 1996

Durante los años 70 y 80 del siglo pasado en España se disponía de una cifra oficial de población cada diez años (el Censo de 1970 y el de 1981) y en los periodos intercensales se procedía a la actualización de esos datos mediante la renovación o rectificación padronal.

¹ Ver documento “Guidelines on the use of registers and administrative data for population and housing censuses”. Diciembre de 2018. <https://www.unece.org/index.php?id=50794&L=0>

Esencialmente el método consistía en solicitar a los ayuntamientos datos agregados de su población residente cada año, que éstos obtenían de los recuentos de sus padrones municipales, existentes en España desde hace siglos en muchos casos. Este sistema se conocía como “rectificaciones padronales”¹.

Así mismo, en el año central de cada periodo intercensal se realizaba lo que se conocía como “renovación padronal”. En 1986, entre los censos de 1981 y 1991, se produjo la última renovación padronal. La “renovación” de 1986, al contrario de la “rectificación” anual era un verdadero procedimiento censal. En 1986 se enviaron a los hogares los volantes de empadronamiento para actualizar sus datos de vecinos residentes en cada municipio.

En 1996 se cambió la normativa padronal y se puso en marcha un sistema muy novedoso de gestión continua e informatizada de los Padrones municipales, que pasaban a estar coordinados por el INE. Era el inicio de un registro de población continuo en España.

Las antiguas rectificaciones anuales dieron así paso a cifras anuales de población a partir del padrón, a uno de enero de cada año, desde 1998.

Pero a diferencia de cómo venía sucediendo hasta entonces, de acuerdo con la nueva normativa, el censo de población y viviendas no venía a corregir la serie padronal sino que las cifras del padrón se actualizaban por medio de actuaciones de los ayuntamientos, no mediante una operación de envío de cuestionarios censales.

Así pues, desde la entrada en vigor del padrón continuo, se sientan las bases para que en España haya dos series de datos de población independientes entre sí. Es a partir de 2001, con el censo, cuando se consuma esta separación en dos series de datos:

- Por un lado, cada año, se tienen las cifras poblacionales basadas en recuentos directos de **padrón** ([Estadística del Padrón Continuo](#)) que se publican, como cifras oficiales para todos los municipios de España a uno de enero.
- Por el otro, los datos del censo de 2001 requieren de una actualización con fines estadísticos. Se construye la que podemos denominar “**serie estadística**” ([Cifras de Población](#)). Durante la década 2001-2011 esta serie estadística se conoce como “*Estimaciones de la población actual*” (EPOBA) y desde 2011, con alguna mejora metodológica, bajo la nueva denominación de “*Cifras de Población*”. Esta serie se actualiza también cada seis meses, sumando a la cifra censal los flujos demográficos estimados.

En realidad esas dos fuentes no son del todo independientes: lo son en cuanto a que la segunda toma como punto de partida el censo, pero la serie estadística también se apoya en el padrón, dado que tanto los flujos migratorios como el propio censo 2011 se obtienen del padrón. No obstante, todavía a 1 de enero de 2021, los datos publicados del Censo son ligeramente diferentes de los datos del padrón. En el apartado 9, al final de este documento, se describen las líneas generales sobre las que puede basarse en un futuro próximo la desaparición de esta dualidad de cifras poblacionales que cabe considerar como una anomalía, aunque ya de pequeña magnitud, en el sistema estadístico español.

¹ Ley 70/1980, de 16 de diciembre, por la que se modifican las fechas de referencia para la formación de los Censos Generales de la Nación y de renovación del Padrón Municipal de Habitantes.

5.1.2 La cifra de población del censo 2011

El recuento de la población de España a 1 de noviembre de 2011 que se publicó como principal dato del censo (46.815.916 habitantes) se obtuvo mediante un sistema completamente nuevo: no era el producto de un simple recuento de cuestionarios recogidos, como ocurría tradicionalmente con los censos, ni tampoco era el número de habitantes que se obtenía del padrón. Se ideó un procedimiento específico para contar la población a partir del padrón pero usando otra información disponible para acercar esa cifra al concepto de población residente habitual, propia de unos censos de población, y no al de población empadronada, que cumple otros fines.

El procedimiento partía de considerar como “dudosos” a un porcentaje de la población que figuraba empadronada a partir de cruce de registros administrativos.

Los factores de recuento

A las personas, cuya residencia en España se consideró segura, se les asignó un “factor de recuento” igual a 1. Este “factor de recuento”, ideado para el censo de 2011, puede entenderse como la probabilidad de que una persona resida en España a partir de la evidencia encontrada en distintas fuentes. Un valor 1 es una probabilidad del 100% y por tanto, la persona se cuenta como residente.

Como consecuencia de este cruce de ficheros, hubo un porcentaje pequeño, pero no despreciable, de la población (inferior al 3%) para el cual no se disponía de evidencia alguna en los registros. La única presencia de una persona en el padrón, en algunos casos registrada con mucha antigüedad, no era suficiente para contarla como residente. Esta población se consideró “dudosa” y se introdujo un método estadístico para contarla.

Se contabilizó esa población asignándoles “factores de recuento” distintos de 1. Los factores de recuento se calcularon a partir de la encuesta dirigida a la muestra del 10% de la población que formó parte del proyecto censal. Los detalles del método se pueden encontrar en la metodología del censo de 2011. Pero cabe destacar que el factor de recuento medio resultó ser 0,424, lo que implica que los 1.046.000 dudosos se contabilizaron como una población de 444.000 personas aproximadamente.

Al final de este proceso todos los registros tienen asignado un factor de recuento, con valor 1 para el 97% de los registros, y con valor distinto de uno (casi siempre menor), para el resto. El fichero así formado se conoció como Fichero Censal Final Ponderado (FCFP), que contenía unos 47,4 millones de registros, pero que sumaban 46,8 millones de residentes estimados: la cifra censal

Este método introduce una complicación adicional cuyas consecuencias no se podían valorar suficientemente en aquel momento: el recuento de personas deja de ser un número entero y pasa a ser un número con decimales. Esto significa, entre otras cosas, que las tablas que se construyan muchas veces no suman, lo cual es fuente de innumerables pequeños problemas de cara a los usuarios durante la década siguiente. Un ejemplo sencillo, con datos reales, se puede observar en la figura 4.

Figura 4. Datos de población para un municipio según el censo de 2011

	Dato publicado (valores enteros)				Dato real (con decimales)			
	Ambos sexos	Hombres	Mujeres	Diferencia	Ambos sexos	Hombres	Mujeres	Diferencia
Total	180	87	94	-1	180,2266	86,671	93,5556	0
Nacidos en España	163	77	86	0	162,7705	77	85,7705	0
Nacidos en el extranjero	17	10	8	-1	17,456	9,671	7,785	0

5.1.3 Las cifras de población desde 2012

A partir del censo 2011, cuya fecha de referencia es 1 de noviembre, se construye la serie estadística de cifras de población, publicada cada semestre, con inicio en 1 de enero de 2012.

Esta serie se actualiza mediante un método “de flujos”, es decir, al stock poblacional referido a 1 de enero de cada año se le suman o restan (según corresponda) los siguientes flujos:

- Nacimientos y defunciones, obtenidos de las estadísticas de movimiento natural de la población.
- Migraciones interiores y exteriores: provenientes de la Estadística de migraciones, es decir, a partir de movimientos observados en el padrón.
- Adquisiciones de nacionalidad española de residentes (operación estadística del mismo nombre) para proporcionar cifras por nacionalidad.

Entretanto, cada año se siguen publicando las *Cifras oficiales de población de los municipios españoles* que se derivan del Padrón. Los resultados son ligeramente diferentes debido sustancialmente a que la serie estadística parte del censo, que ofrece cifras distintas al padrón pero evolucionan de forma similar.

5.2 LA CIFRA CENSAL Y EL MÉTODO DE SIGNOS DE VIDA

Quizá el elemento más destacado de todo censo es la determinación de lo que se conoce como la “cifra censal”, es decir, el número de habitantes, el universo de población a partir del cual se realizan todas las desagregaciones y análisis.

El método de estimación de la población residente usado para el censo de población de 2021 es completamente novedoso. Guarda ciertas similitudes con el usado en 2011 en el sentido de que se parte de los datos del padrón pero se ha corregido la cifra de forma diferente, no mediante una estimación de la población dudosa, sino mediante la aplicación de diversas técnicas que se detallan a continuación, la más importante de las cuales es la de los *signos de vida*.

Como se ha comentado, asignar factores de recuento a la población en 2011 trajo aparejado un recuento poblacional con decimales, lo cual resultaba poco natural para los usuarios menos expertos en estadísticas demográficas, acostumbrados a que la población solo puede medirse con números enteros.

Para 2021, el método planteado ha hecho innecesario el uso de factores de recuento y por tanto de decimales. Se trata de aplicar lo que se conoce en la literatura reciente de

censos de población como signos de vida (o más formalmente, como “signos de presencia”).

Si bien conceptualmente parece poco discutible que pueda utilizarse información adicional para mejorar el contenido de un registro de población, en realidad apenas se han encontrado ejemplos de aplicación práctica de este método en el mundo.

Se trata de partir de un registro de población, en nuestro caso el padrón, y de evaluar la verosimilitud de la residencia de cada persona a partir del cruce con otros registros. Así, no se tienen registros seguros y dudosos sino que, para cada individuo, a partir del cruce con otros registros, se tiene que decidir, si el registro se cuenta como residente o no.

Como resultado de todo este proceso se obtiene un fichero censal que contiene tantos registros como habitantes saldrían del recuento de población (la cifra censal). De esta manera cualquier recuento de otras variables, como el número de hogares en que viven, se haría a partir de ese fichero censal de forma natural y siempre usando números enteros.

El caso teórico y las limitaciones prácticas de un método de signos de vida

En teoría, en la sociedad actual es casi imposible vivir sin dejar rastro en algún tipo de registro luego si una oficina de estadística dispusiera de acceso ilimitado a tal información, que no es el caso, parecería que podría determinarse de forma bastante precisa para cada persona si reside o no en el país (o en un lugar concreto).

Si nos ceñimos estrictamente a registros administrativos, es decir, de propiedad pública (no considerando así accesos a redes sociales, uso de tarjetas de crédito o cuentas bancarias, telefonía móvil u otros sistemas de titularidad privada) la situación cambia radicalmente. Aun así, podría haber registros administrativos que ofrecieran información muy relevante sobre presencia: pensemos en títulos de transporte público, consultas médicas, adquisición de medicamentos, información de ayuntamientos, tráfico, servicios sociales...

En la práctica, a día de hoy, apenas existen ficheros de este tipo utilizables para un censo de población y se deberían vencer muchas barreras técnicas, de protección de datos o de opinión pública antes de que un censo de población pudiera hacer uso de estos medios para determinar la residencia de las personas, aunque sea con fines estadísticos como es el caso.

Pero existen ciertos ficheros administrativos a nuestra disposición que forman parte del proyecto censal y que pueden ayudar a determinar de forma algo más precisa la cifra de residentes que si nos atenemos a la cifra padronal. Pero no solo eso, es que si no los usamos, estaremos cometiendo errores e imprecisiones importantes no solo en la cifra de residentes sino también en algunas variables de las que recoge el censo, como por ejemplo las relativas a la actividad económica.

Ensayos previos al censo de 2021 de métodos de signos de vida

El trabajo se inició en el FPC-2016 y en el FPC-2018 y se usó un procedimiento, con varias alternativas, que se mostró poco realista porque obligaría de descontar demasiada población al no encontrarse –lógicamente– signos de vida para ella.

Se tomaba al Padrón como elemento básico de la estructura del fichero precensal y a partir de él se elaboró un procedimiento para efectuar un recuento de la población mediante cruce con otros ficheros.

Después se analizaban los signos de presencia que las personas de la base tienen en los registros administrativos disponibles y también los movimientos padronales en los meses posteriores a la fecha de referencia.

Dentro de la **información administrativa** que se utiliza para determinar estos signos de presencia de una persona en España, cabe destacar:

- Información procedente de la Agencia Tributaria, Haciendas Forales y Hacienda Territorial de Navarra.

Se tiene en cuenta tanto si la persona aparece como declarante del Impuesto sobre la Renta de las Personas Físicas del ejercicio anterior como la existencia de retenciones e ingresos a cuenta, distinguiendo entre los practicados por diversos conceptos (rendimientos del trabajo, rendimientos de actividades económicas, prestaciones recibidas, renta mínima de inserción, etc...).

- Información contenida en el fichero de asegurados y beneficiarios en alta a 1 de enero de cada año en la Base de datos de aseguramiento sanitario (BADAS) proporcionada por el INSS.

En este fichero los asegurados están clasificados en grupos de aseguramiento de acuerdo a su condición (trabajadores, pensionistas,.....). Existen grupos específicos que incluyen los asegurados de otros países y sus beneficiarios que están residiendo en España cuyos gastos sanitarios van a cargo del país de aseguramiento y para los cuales se ha recibido un formulario que garantiza que se hacen cargo de su cobertura sanitaria en España. Se considera que en estos casos la residencia en España es altamente probable.

Existe otro grupo en el que se incluye a las personas que perdieron su aseguramiento inicial por diversas causas, por ejemplo un trabajador que pierde su trabajo, y finalizado el periodo de prórroga no han encontrado otra vía de aseguramiento. Estas personas tienen una menor probabilidad de estar residiendo en España.

Pero en general el fichero BADAS ofrece poca utilidad para signos de vida en tanto que su naturaleza consiste en garantizar asistencia sanitaria y por tanto tiene menos tendencia a borrar población.

- Fuentes relativas al mercado laboral

Se analiza la información de los ficheros de demandantes de empleo a último día de los meses de noviembre y diciembre del año anterior a la fecha de referencia, proporcionados por el Servicio Público de Empleo Estatal, teniendo en cuenta que la persona cumpla las condiciones para ser considerada como parado a efectos del SEPE.

También se utiliza el fichero de afiliados a la Seguridad Social que estaban de alta durante la semana previa a 1 de enero de cada año, proporcionado por la Tesorería General de la Seguridad Social, así como el fichero de perceptores de prestaciones a 1 de enero de cada años del Instituto Nacional de la Seguridad Social.

Se consideran también los mutualistas, tanto asegurados como beneficiarios, a 1 de enero de las principales mutualidades (MUFACE, MUGEJU; ISFAS).

En general, la presencia de una persona, con ciertas salvedades, en estos ficheros puede tomarse como signo de presencia para el proceso de cálculo de la población.

- Otras fuentes complementarias

Información procedente del Registro Central de Extranjeros, como las fechas de caducidad y comprobación de residencia, fechas de solicitud, concesión y denegación de permisos de residencia, etc...

Inicialmente todos estos indicadores de presencia en los diferentes ficheros administrativos se cruzan a nivel individual pero esto da como resultado que en torno a un 10% de la población, bien no tiene signos de vida, bien los tiene pero resultan poco significativos. Lo segundo ocurre si contamos los registros que cruzan con el fichero BADAS, por su tendencia a no borrar población.

Una persona inactiva, que conviva con otra que tiene ingresos bajos, y por tanto no tenga que hacer declaración de la renta, no aparecería en ninguno de los registros anteriores pero esa no es razón suficiente para descontarla en un recuento de población (estaríamos borrando a varios millones de personas). En general la población no encontrada en signos de vida es mayoritariamente femenina, aumenta con la edad y afecta más a extranjeros que a españoles.

La aplicación de signos de vida a nivel individual tiene además una importante implicación a nivel de hogar. Por ejemplo, si una persona no aparece en signos de vida y se borra, esta persona desaparece de un hogar cuya composición se altera y puede incluso dar lugar a hogares sin sentido (hijos menores viviendo solos).

Por eso el análisis de signos de vida se enfrentó en una segunda revisión, a nivel de hogares. Se probaron diversas alternativas con la idea de que solo con que un miembro mostrara signos de vida el hogar entero se consideraba confirmado.

Aún con esta aproximación se tiene que en torno al 1,5% de la población (lo que significa unas 700.000 personas) no tienen signos de presencia. El análisis de esta población por distribución geográfica y por características sociodemográficas hacía pensar que más bien se estaría eliminando población por defectos en los cruces de ficheros que por verdaderos signos de vida.

El método de signos de vida utilizado en el censo 2021

El análisis de literatura en otros países no ayudó demasiado. Esencialmente no se conoce la aplicación práctica de ningún modelo que nos pueda servir de referencia.

Finalmente, el método usado para el censo de 2021 se aplica exclusivamente a población extranjera, dado que solo para ella se tienen los procesos de caducidad o comprobación de oficio que el método de signos de vida pretende corregir.

Además, se aplica a nivel de hogar, de modo que solo se dejan de contar hogares en los que de forma inequívoca todos los miembros del hogar están afectados o bien el hogar resultante no tiene una composición alterada, por ejemplo, los hogares en que no hay relaciones de parentesco entre sus miembros. El padrón municipal funciona en la actualidad aplicando a la población extranjera unos procedimientos de renovación que permiten mantener razonablemente actualizado el fichero. Es algo parecido a lo

que hacen otros países con registros de población o a cómo funcionan registros como el DNI, que obligan a la renovación cada cierto tiempo.

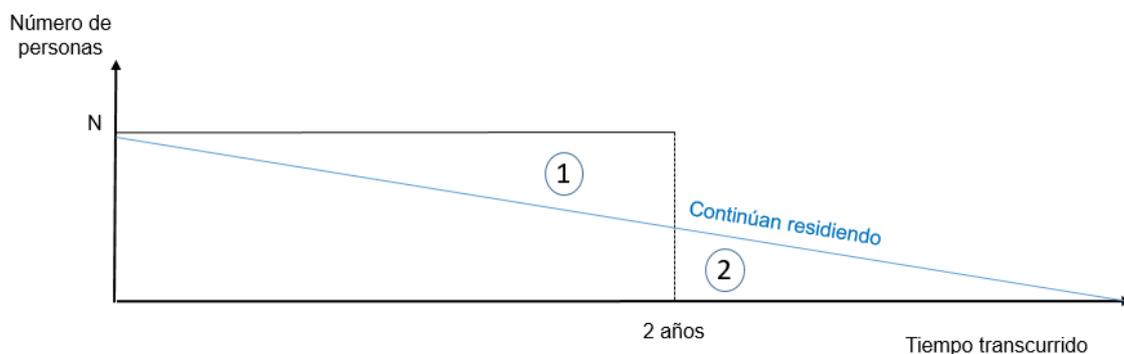
La población extranjera, a efectos de la gestión padronal, se divide en dos grupos, "ENCARP" y "no ENCARP". Los extranjeros no comunitarios sin autorización de residencia permanente (abreviados ENCARP), son las personas que residen en España sin tener la nacionalidad española ni la de ningún otro Estado de la Unión Europea y que no disponen de una autorización de residencia permanente expedida por la Dirección General de la Policía. Desde 2005 estas personas deben acudir a sus respectivos ayuntamientos para renovar su inscripción en el padrón municipal cada dos años. De no hacerlo se pone en marcha un proceso que termina dando lugar a su baja padronal por caducidad (llamada BBC en terminología del padrón).

El resto de población extranjera se conoce como "NO ENCARP". Estos ciudadanos extranjeros residentes en España no están obligados a renovar su inscripción padronal, pero desde 2008 se puso en marcha un procedimiento paulatino mediante el cual los ayuntamientos deben confirmar periódicamente su residencia. Al principio se hacía solo para un porcentaje de ellos pero en la actualidad se comprueba para todos ellos. Esta comprobación debe realizarse para todos los extranjeros de este grupo que no hayan producido movimientos en un período de cinco o dos años, dependiendo si están inscritos en el Registro Central de Extranjeros o no. En caso de no poder comprobar que la persona sigue residiendo en el municipio, el ayuntamiento correspondiente pondrá en marcha un procedimiento que culminará con una baja de oficio por inscripción indebida (BII).

En definitiva, este procedimiento da lugar a que para la población extranjera en su conjunto se pueda analizar si la inscripción del habitante ha alcanzado su fecha de "caducidad" (estrictamente solo caduca la inscripción de los ENCARP pero podemos hablar de caducidad para el resto cuando han transcurrido esos dos o cinco años). En principio podríamos dejar de contar como residentes a todos los extranjeros de la base padronal que tengan su inscripción caducada.

Si analizamos una cohorte de extranjeros que entra en España en un año dado, el proceso actual y suponemos que todos ellos caducan a los dos años tendríamos esquemáticamente la situación ilustrada en la figura 5.

Figura 5. Esquema del recuento de población extranjera según fecha de caducidad de su inscripción



Según este esquema, las N personas de la cohorte contarían mientras no hubiera caducado su inscripción y todas ellas automáticamente dejarían de contar transcurridos esos dos años.

Pero la población que inmigra abandona el país de forma más parecida a lo que muestra la línea azul descendiente. El triángulo 2 contiene la población que sigue viviendo en España pero cuya inscripción ha caducado. Si cruzamos esta población con los ficheros de signos de vida (laborales, tributarios) encontramos extranjeros que inicialmente estarían “caducados” pero que por presentar signos suficientes de presencia son “rescatados” y contabilizados en la cifra de población censal.

El triángulo 1 contiene la población que ya no reside en España pero cuya inscripción aún no ha expirado, por lo que siguen contando como habitantes. El método de signos de vida aplicado no guarda una analogía inmediata con el usado para el grupo 1, porque por las razones que apuntamos anteriormente la no presencia en estos registros no parece suficiente para dejar de contar a un empadronado. Pero sí se puede hacer de forma más clara para una parte de ellos. La decisión tomada para el censo 2021 es aplicar signos de vida solo a una parte de los extranjeros, aquellos que sí figuraban el año anterior en los ficheros de signos de vida. Así, si una persona extranjera aparecía en el fichero precensal a 1 de enero de 2020 como afiliado o como parado o pensionista, estudiante o en cualquier otra situación que denotara su presencia, pero a 1 de enero de 2021 ya no aparece en ninguno de estos ficheros, consideramos que esa persona ha abandonado el país, aunque su inscripción no haya caducado aún.

Una vez realizado esto en una primera fase se deben analizar los hogares resultantes de la aplicación de estos signos de vida, y con ello se aplica una segunda fase que corrige a la anterior.

En concreto, para los menores de 16 años de edad, dado que, en general, no están presentes en los registros administrativos utilizados en el procedimiento de signos de vida (salvo el grupo de estudiantes de 15 años que figuran matriculados en alguna enseñanza oficial) se considera como signo de presencia que haya al menos un adulto con signos de presencia en el domicilio en el que figuran empadronados. Es decir, se dejarían de contar los menores que tras aplicar el procedimiento de signos de presencia, queden empadronados en un hogar sin adultos.

Y por el contrario, se contabilizarían en la población censal los menores que, aunque inicialmente tengan su inscripción padronal caducada, residan en hogares donde al menos un adulto presenta signos de presencia en registros.

En el método de signos de presencia aplicado en el Censo 2021 se han utilizado los registros de temática laboral ya mencionados (INSS, TGSS, SEPE, Mutualidades), algunos grupos específicos del BADAS para los cuáles la residencia está garantizada y, excepcionalmente, los registros de temática educativa (Matriculados en enseñanzas universitarias y no universitarias). Aunque se disponía también de información tributaria, se ha optado por no utilizarla en el método. El motivo es que la información tributaria se recibe un año después de la fecha de referencia censal y, el objetivo a partir de 2021 es obtener cifras de población censales a 1 de enero de cada año que estén disponibles dentro del año en curso. Para que estas cifras de población censales sean coherentes y formen una serie con periodicidad anual desde 2021 necesitamos basar el método de signos de vida en las mismas fuentes todos los años. Por ello se ha decidido utilizar las fuentes laborales cuya disponibilidad es más rápida y descartar las tributarias.

En cuanto a las fuentes educativas, aunque por su calendario de disponibilidad tampoco van a poder ser utilizadas en los procesos anuales de signos de presencia que se apliquen en el futuro, ha sido conveniente tenerlas en cuenta en el censo 2021 para garantizar la máxima cobertura de las variables de estudios.

La población centenaria

Como ocurre a cualquier registro de población, el padrón acumula registros que pueden considerarse “residuos” de un procedimiento que no puede ser perfecto. Este es el caso de personas que fallecieron, incluso hace décadas y que nunca se borraron. Forzosamente se trata de españoles, dado que para los extranjeros hay procesos de caducidad. Se comprueba cada año que son muy pocos casos pero existen. Son diversas las razones que pueden llevar a esto pero fundamentalmente se trata de errores en la certificación de la defunción que impiden dar de baja a ningún empadronado.

En ese caso, el registro correspondiente a la persona fallecida se mantiene en el padrón y envejece indefinidamente. Estos registros afloran cuando van alcanzando edades elevadas, en las cuales el número de efectivos es ya muy bajo. Por eso se ponen en marcha procesos de control exhaustivos para la población de 100 años o más (llamados centenarios).

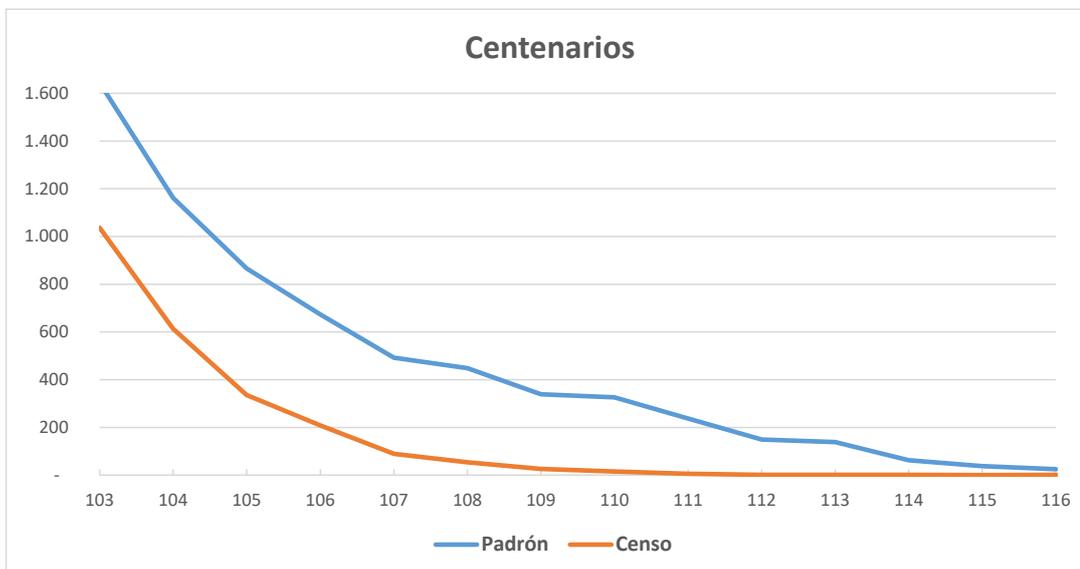
Evidentemente este problema de registros residuales que corresponden a población fallecida pero no eliminada del fichero no solo se da en los centenarios sino en todo el fichero del padrón pero resulta bastante indetectable y solo aflora en esas edades más avanzadas. Hay que reconocer la limitación que esto supone y que puede significar algunos miles de registros pero su impacto es muy pequeño dentro de grupos poblacionales mucho más grandes y por eso se concentra el esfuerzo en la población de 100 y más años.

Por ello se aplica un proceso de corrección de la población centenaria de modo que se elimina del fichero censal todo registro padronal de más de 99 años que no se pueda enlazar con ningún registro del INSS (pensionistas) o tributario, por considerarse altamente improbable que corresponda a una persona viva.

En 2021 esto supone la no contabilización de 5.228 registros de la base padronal.

El efecto de eliminar estos registros de centenarios es que se observa que los efectivos resultantes muestran un patrón mucho más verosímil en cuanto a tasas de mortalidad específicas por edad, como se observa en la figura 6, en la que, para mayor claridad, se muestran los efectivos mayores de 102 años antes y después de aplicar la corrección.

Figura 6. Personas mayores de 102 años antes y después de aplicar corrección por presencia en registros administrativos



El método de signos de vida aplicado en el censo de 2021 presenta los siguientes resultados:

Figura 7. Aplicación de signos de vida en el censo de 2021

Empadronados que no se contabilizan (recinto 1 de la figura 5)	No empadronados que se contabilizan (recinto 2 de la figura 5)	Población centenaria que no se contabiliza
152.599	158.044	5.228

¿Por qué es necesario aplicar signos de vida?

Como se puede observar en los datos anteriores, el resultado neto de aplicar signos de vida es casi cero. Es un esfuerzo grande cuyo resultado apenas significa aumentar la población en 5.445 personas. Pero sin embargo resulta muy necesaria la incorporación de esta fase, no ya por la calidad del recuento de la población sino por la calidad de las variables de la población.

En efecto, la mera aplicación del criterio de empadronamiento correcto para contabilizar la población introduce un problema adicional. Si no aplicáramos signos de vida tendríamos en el fichero censal 152.599 personas de las cuales su relación con la actividad sería "No consta" dado que no aparecen en ninguno de ellos mientras que tendríamos 158.044 personas sobre las cuales conocemos su relación con la actividad dado que figuran en los ficheros de afiliación a la seguridad social o de desempleados o pensionistas y no los podríamos incorporar al censo porque su inscripción padronal ha caducado (y por tanto esa información se despreciaría).

Los resultados sobre población extranjera afiliada o desempleada se resentirían gravemente respecto de datos fehacientes publicados para las mismas fechas por otras instituciones de referencia y en cambio tendríamos para decenas de miles de personas una actividad laboral no asignada, luego la calidad de esas variables quedaría muy afectada.

La aplicación de signos de vida viene por tanto a resolver más un problema de calidad en las variables que de cantidad de población.

Aplicación futura del método de signos de vida

La elección del método de signos de vida empleado en el censo de 2021 se ha tomado pensando en el mantenimiento futuro de una serie anual de población. No se puede usar un método de signos de vida para contabilizar la población a 1 de enero de 2021 que sea distinto del método que se use durante los años siguientes, lo que daría lugar a rupturas en las series de población. Así, tal como se describe en el apartado 9 sobre el futuro de los censos, se construirán cifras anuales de población plenamente comparables con la cifra censal.

Pero es evidente que el método de signos de vida empleado es solo una primera aproximación a lo que un método más ambicioso podría ofrecer. El método del censo no puede constituir una hipoteca para métodos futuros, si bien cuando estos se puedan emplear será necesario revisar series o bien constatar e intentar cuantificar la ruptura de series. Esto es especialmente relevante en el caso de la población de nacionalidad española, para la cual hay margen de mejora en la aplicación de signos de vida.

Otra población no incorporada en el Padrón: edades 0 y 1

Otro elemento que diferencia la cifra censal de la recogida en el padrón es la población de edades muy cortas, fundamentalmente menos de 2 años. Por el método administrativo usado en el padrón, una pequeña parte de los nacimientos se incorporan de forma algo tardía al padrón. Estos nacimientos están detectados y se tiene de ellos la información de inscripción en registro civil así como el boletín estadístico de parto. Obviamente esos registros se añaden en el recuento censal. En el caso concreto de 2021 suponen unos 22.000 niños de edades 0 o 1 año, casi todos ellos de nacionalidad española.

Finalmente, la cifra de población según el censo de población a 1 de enero de 2021 asciende a **47.400.798** personas. De ellas, **5.402.702** son extranjeros.

La comparación entre esta cifra y la del padrón a la misma fecha figura en la tabla 8 a continuación:

Figura 8. Comparación entre cifra de población censal y padronal a 1 de enero de 2021

	Censo 2021 (A)	Padrón 2021 (B)	– (B)
Total	47.400.798	47.385.107	15.691
Españoles	41.998.096	41.944.959	53.137
Extranjeros	5.402.702	5.440.148	-37.446

5.3 VARIABLES RELATIVAS A LA POBLACIÓN

A continuación se detalla la situación por variables o grupos de ellas. Para el análisis de cada caso se seguirá la misma estructura:

- Requisitos
 - Reglamento UE: Especifica el requisito mínimo para el proyecto

- Censos anteriores: el hecho de poder dar continuidad a variables contenidas en censos anteriores no es, estrictamente hablando, un requisito, pero sí un factor importante a considerar.
- Planteamiento propuesto para la variable: a la luz de los requisitos y las fuentes disponibles se muestra como se obtiene la variable en cuestión.
- Método y fuentes disponibles: se explica con qué fuentes se ha construido la información y cuál ha sido el método seguido.
- Análisis de calidad y mejoras previstas: para cada variable se analiza la calidad de la información obtenida, tanto desde la perspectiva de los resultados en la asignación de la variable (en cuántos casos se ha logrado asignar un valor a partir de las fuentes y qué fuentes han intervenido) como de los resultados, cuando puede cotejarse con otras fuentes.

Dado que el censo de 2021 marca el inicio de un proceso continuo de mejora de las estadísticas demográficas, cabe mencionar en ciertos casos las modificaciones que se podrán ir incorporando en los próximos pero que aún no se han podido poner en práctica por diversas razones.

- Residió siempre en el extranjero y llegó en 1980 o posteriormente
- Entre 2020 y 2021
 - 2021
 - 2020
- Entre 2015 y 2019
 - 2019
 - 2018
 - 2017
 - 2016
 - 2015
- Entre 2010 y 2014
 - 2014
 - 2013
 - 2012
 - 2011
 - 2010
- Entre 2005 y 2009
 - 2009
 - 2008
 - 2007
 - 2006
 - 2005
- Entre 2000 y 2004
- Entre 1995 y 1999
- Entre 1990 y 1994
- Entre 1985 y 1989
- Entre 1980 y 1984
- Residió en el extranjero y llegó en 1979 o anteriormente, o nunca residió en el extranjero
 - No consta

En el Reglamento, la desagregación año a año solo se exige en tablas a nivel nacional.

– **Lugar de residencia habitual un año antes del censo**

Constará de las siguientes categorías, que proporcionan una relación entre los lugares de residencia habitual en la fecha censal y un año antes del censo:

- No cambió de residencia habitual
- Cambió de residencia habitual
 - En el interior del país declarante
 - Misma región NUTS3 (provincia/isla)
 - Distinta región NUTS3 (provincia/isla)
 - Fuera del país
- No consta.

- No aplicable (niños de menos de 1 año)

Censos anteriores

Las variables relativas a migraciones se constrúan a partir de la siguiente información del cuestionario:

1 ¿Desde qué año reside en esta vivienda?

Desde el año Desde que nació
(Pase a la pregunta 3)

¿y en este municipio?

Desde el año Desde que nació
(Pase a la pregunta 3)

¿y en esta comunidad autónoma?

Desde el año Desde que nació

¿y en España?

Desde el año Desde que nació

¿Dónde residía antes de llegar por última vez a este municipio?

En otro municipio:

Provincia:

En otro país:

2 ¿Dónde residía hace 1 año?

En este municipio (o no había nacido aún)

En otro municipio:

Provincia:

En otro país:

¿Dónde residía hace 10 años?

En este municipio (o no había nacido aún)

En otro municipio:

Provincia:

En otro país:

b) Planteamiento propuesto para las variables y desglose

El análisis que puede realizarse sobre lugares de residencia anteriores y lugares de procedencia es mucho más rico que en el caso del censo de 2011 porque se puede reconstruir, con ciertas limitaciones como se detalla más adelante, la trayectoria de las personas desde la puesta en marcha del padrón continuo en 1996.

Para los movimientos previos a esa fecha solo se dispone de la información que aporta el censo de 2001 y, por tanto, la reconstrucción está limitada por las preguntas contenidas en aquel cuestionario censal, y solo para aquellos casos en que la información entre ambas fuentes sea coherente para los años comunes (1996 – 2001).

El resumen de las variables incorporadas y las fuentes que intervienen se encuentra en la figura 9.

Figura 9. Resumen de variables migratorias contempladas y fuentes utilizadas

Variable	No aplica (*)	Total a asignar	FUENTE CON LA QUE SE HA ASIGNADO			
			Base Padronal	Censo 2001	Imputación	No consta
Año llegada a España	83,8%	16,2%	14,6%	1,0%	0,6%	
Año llegada a la comunidad	77,2%	22,8%	10,9%	9,6%	2,3%	
Año llegada a la provincia	72,4%	27,6%	13,8%	9,2%	4,6%	
Año llegada al municipio	46,4%	53,6%	32,6%	19,7%	1,3%	
Año de llegada a la vivienda	15,2%	84,8%	62,9%			21,9%
País de procedencia	83,8%	16,2%	11,6%	1,0%	3,6%	
Comunidad de procedencia	77,2%	22,8%	11,1%	6,4%	0,07%	5,3%
Provincia de procedencia	72,4%	27,6%	14,0%	7,6%	0,03%	6,0%
Municipio de procedencia	46,4%	53,6%	33,0%	13,3%	0,01%	7,4%
Residencia hace 1 año	0,7%	99,3%	99,3%		0,01%	0,0003%
Residencia hace 10 años	8,8%	91,2%	91,2%		0,01%	0,002%

(*) Siempre residieron en ese lugar/no había nacido hace 1 año o 10 años

Leyenda:

Exigida por reglamento (y proporcionada en el Censo 2011)

Proporcionada en el Censo 2011

La variable “Año de llegada a España” se obtendrá únicamente para aquellas personas que hayan residido en alguna ocasión en el extranjero, a diferencia del Censo 2011, en que esta variable recogía además el año de nacimiento para aquellas personas que siempre habían residido en España.

Para la variable “Lugar de residencia habitual hace un año” (un año antes del censo) se amplía el número de categorías exigidas en el reglamento a las siguientes (sustituyendo la unidad geográfica NUTS3 por provincia, y distinguiendo, en caso de cambio de residencia, si es el mismo municipio, provincia, comunidad y, en caso de No consta, distinguiendo si siempre residió en España o no):

- 1- Misma vivienda
- 2- Distinta vivienda del mismo municipio
- 3- Distinto municipio de la misma provincia
- 4- Distinta provincia de la misma comunidad
- 5- Otra comunidad
- 6- Residía en el extranjero
- 7- No consta la residencia hace un año, pero siempre ha residido en España
- 8- No consta la residencia hace un año, y no siempre ha residido en España
- 9- No había nacido

Además de las variables *Año de llegada a España* y *Lugar de residencia habitual un año antes del censo*, obligatorias por reglamento, se proporcionarán una serie de variables adicionales, agrupadas en varios lotes y con ciertas categorías agregadas:

– Años de llegada y lugares de procedencia:

- *País de procedencia*, asociado al “año de llegada a España”, es decir, el país donde residía antes de trasladar a España su residencia habitual por última vez (en caso de que haya habido varias ocasiones).

- *Año de llegada a la comunidad autónoma y comunidad autónoma de procedencia.*
- *Año de llegada a la provincia y provincia de procedencia.*
- *Año de llegada al municipio y municipio de procedencia*, es decir, municipio de residencia anterior a la de la fecha censal. Además, *tamaño del municipio de procedencia.*
- *Año de llegada a la vivienda.*

Los años de llegada a cada región (junto con sus regiones asociadas de procedencia) pueden ser diferentes, puesto que no tienen por qué referirse todos simultáneamente al mismo municipio de residencia. Por ejemplo, una persona pudo moverse de la comunidad autónoma Madrid a la de Castilla y León en 2015 (en la provincia de Valladolid), pero moverse a la provincia de Zamora en 2018 dentro de la misma comunidad (suponiendo que sea la última vez que llegó a esa provincia). En este caso, llegó a Castilla y León en 2015 (con comunidad de procedencia Madrid) y a Zamora en 2018 (con provincia de procedencia Valladolid, que no pertenece a la comunidad de Madrid).

En los casos en que no ha habido variación de región, se considera como "No aplicable".

Los años de llegada a cada región se proporcionarán desglosados año a año desde 2016, agrupados para el quinquenio 2011-2015, en decenios desde el 1981 hasta el 2010, y una categoría abierta para años anteriores a 1981. El año de llegada a la vivienda, que se obtiene solo de los datos del padrón y no del Censo 2001, se proporciona con el mismo detalle desde 2001, y el resto agrupados en una sola categoría abierta hasta el año 2000.

– Último lugar de residencia anterior:

Se refiere al último municipio de residencia anterior al censal (si lo hubo; será "No aplicable" si siempre residió en el mismo), y sus regiones asociadas (provincia, comunidad autónoma y país). Si alguna de estas regiones geográficas asociadas al municipio anterior es igual a la del municipio la censal, se indica con la categoría "No aplicable", para indicar que no ha cambiado de región. Estas regiones no tienen por qué coincidir con las regiones de procedencia, donde se considera por separado la región anterior (diferente) a la censal, mientras que en esta variable consideramos las regiones asociadas al municipio anterior al censal en España.

Por ejemplo, si siempre ha vivido en la misma provincia, pero ha cambiado de municipio, existirá un municipio concreto anterior al censal, pero la provincia no habrá cambiado, y se indicará con la categoría "No aplicable".

Además de la variable que relaciona el lugar de residencia censal con la anterior, se proporcionan las regiones correspondientes a ese municipio anterior. En total:

- *Municipio, provincia y comunidad de residencia anterior* (las regiones correspondientes al municipio anterior), en su última estancia en España (desde que nació o llegó a España). La provincia y la comunidad son las relativas al municipio de residencia anterior, al contrario que las variables de procedencia, que pueden ser independientes. En ese caso, también se da el *tamaño del municipio*.

- *País de residencia anterior*: si en su última estancia en España (desde que nació o llegó a España por última vez) no vivió en más de un municipio, y procede del extranjero, se proporciona este país.
 - *Relación entre el lugar de residencia anterior y actual* (censal), como se ha explicado antes, ampliando las categorías que publicamos respecto a las pedidas en el reglamento.
- Lugar de residencia hace un año:
- *Municipio, provincia y comunidad de residencia hace un año* (desde la fecha de referencia censal). En ese caso, también se da el *tamaño del municipio*.
 - *País de residencia hace un año*.
 - *Relación entre el lugar de residencia en 2020 y actual* (censal), como se ha explicado antes.
- Lugar de residencia hace diez años:
- De modo equivalente al lugar de residencia hace un año, se generan las variables análogas para el lugar de residencia hace diez años.
- *Municipio, provincia y comunidad de residencia hace diez años* (desde la fecha de referencia censal). En ese caso, también se da el *tamaño del municipio*.
 - *País de residencia hace diez años*.
 - *Relación entre el lugar de residencia en 2011 y actual* (censal).

c) Método y fuentes

El padrón contiene la información necesaria para construir variables migratorias casi directamente, dado que se conserva para cada persona cuáles son sus lugares de residencia anteriores, pero la gran limitación es que inicia su historia en el año 1996.

En su instante inicial, el padrón continuo se puebla con las variables que se disponen para esa carga inicial, que son las variables identificativas mencionadas en el apartado anterior. Nada se sabe, por tanto, sobre los movimientos anteriores de los empadronados de 1996 más allá de su lugar y fecha de nacimiento. Estos casos se pueden tratar de enlazar con los recogidos en los cuestionarios del Censo del 2001, en los que se realizaban algunas preguntas sobre residencias anteriores.

Se ha realizado un enlace con el censo de 2001, que fue un censo exhaustivo, de forma que, siempre que se consigue asociar a un individuo con su cuestionario de 2001 se puede obtener información sobre su historia residencial anterior a 1996, siempre que la información contenida en los años comunes entre ambas fuentes (1996 - 2001) sea coherente. Esto ocurre en muchos casos, como se observa en la figura 9 y, por eso, aunque la principal fuente para obtener las variables migratorias sea el padrón, también tiene un fuerte peso censo el Censo de 2001, tanto mayor cuanto menos reciente sea la variable a la que nos referimos.

En caso de que no se consiga encontrar a la persona en el Censo de 2001, o la información que se obtenga sea incoherente o incompleta, se llevará a cabo un proceso de imputación. Si para una persona conocemos su lugar de nacimiento y es distinto de su lugar de residencia en 1996, se imputa un movimiento migratorio previo a ese año.

En cuanto a la variable "País de procedencia", han sido necesarias pequeñas depuraciones similares a las de las variables país de nacimiento y nacionalidad en cuanto a código de países. Si queda en blanco, se realiza un proceso de imputación en base al país de nacimiento.

Esta variable, junto con los años de llegada a cada región, se ha imputado en su totalidad, mientras que para los lugares de procedencia se ha permitido el valor en blanco. Mención especial cabe sobre el "Año de llegada a la vivienda". Dado que se ha obtenido solo a partir del padrón, y no se ha recuperado información del censo 2001, solo tenemos información desde 1996. Así, una persona con esta variable el blanco puede ser porque siempre haya vivido en España, pero desde antes de 1996, o porque realmente no conozcamos la información (anterior a 1996). Por este motivo, esta variable se proporciona con menor desagregación temporal.

Las variables de lugar de residencia habitual un año antes del censo y hace diez años se pueden obtener casi totalmente de la base padronal.

d) Análisis de la calidad de la variable y mejoras previstas

Entre las fechas que figuran en la base de datos de Padrón y las recuperadas enlazando con el Censo de 2001 se consigue asignar el año de llegada a la gran mayoría de los registros. En definitiva, el grado de imputación aplicado a este grupo de variables es muy bajo, como se observa en la figura 9. La peor situación se da en la variable "Año de llegada a la provincia", puesto que al no existir la pregunta correspondiente en el Censo de 2001, tiene mayor porcentaje de imputación.

No obstante, la información que se tiene en la base padronal sobre lugares de residencia previos desde 1996 es mucho más rica que lo que se ofrece en el Censo 2001, que solo sigue la lógica de un cuestionario y recoge para cada persona una serie de variables.

Posteriormente a la publicación del censo se prevé la producción de mejor información longitudinal sobre recorridos migratorios de las personas a partir de la información padronal.

5.3.3 Estudios en curso y nivel de estudios alcanzado

a) Requisitos

Reglamento UE:

Se debe proporcionar información sobre nivel de estudios alcanzado para todas las personas de 15 y más años en la fecha de referencia. Los menores de 15 años se clasificarán como “no aplicable”.

La variable nivel de estudios se desagrega de forma obligatoria con las categorías siguientes:

- Menos que primaria
- Educación primaria
- Educación secundaria baja
- Educación secundaria alta
- Enseñanza postsecundaria no terciaria
- Educación terciaria de ciclo corto
- Grado en educación terciaria o nivel equivalente
- Nivel de maestría, especialización o equivalente
- Nivel de doctorado o equivalente
- No consta
- No aplicable

El reglamento no impone ninguna solicitud sobre estudios en curso.

Censos anteriores:

Si bien por reglamento no se requería dar información de estudios en curso y naturaleza de estos estudios, el Censo 2011 incluyó estas preguntas por interés nacional y comparabilidad con censos anteriores.

En concreto, en el Censo 2011, las variables censales relacionadas con los estudios eran las siguientes:

- Para menores de 16 años información de si está acudiendo o no a un centro escolar (pregunta 9)

- Grados universitarios de 240 créditos ECTS, diplomados universitarios, títulos propios universitarios de experto o especialista, y similares.
- Grados universitarios de más de 240 créditos ECTS, licenciados.
- Másteres, especialidades en Ciencias de la Salud por el sistema de residencia y similares.
- Doctorado universitario.

La CNED-A contempla tanto programas vigentes como no vigentes, de educación formal. Hay que señalar que no tenemos una cobertura total de todas las categorías de la CNED-A, ya que, por ejemplo, no se dispone de información de títulos propios universitarios (CNED-A 52, 63, 75). Además contamos con una cobertura parcial de la Educación postsecundaria no superior (CNED-A 41), así como de niveles menores de los certificados de profesionalidad (CNED-A 23, 24).

La mayoría de estas categorías más problemáticas quedan inmersas en otras de la CNED-A adaptada (grado superior o grados universitarios en el caso de los títulos propios; Primera etapa de la educación secundaria en el caso de los certificados de profesionalidad 1 y 2).

Para estudios en curso, se proporcionan datos según doce categorías basadas en la Clasificación de programas en niveles educativos¹ (CNED-P):

- Educación primaria e inferior
- Primera etapa de educación secundaria y similar
- Segunda etapa de educación secundaria con OG: 4º ESO
- Segunda etapa de educación secundaria con OG: Bachillerato
- Enseñanzas de formación profesional, artes plásticas y diseño y deportivas de grado medio y similares ; Programas de cualificación profesional inicial y similares; F.P. básica
- Segunda etapa de Ed.secundaria con OP (y sin cont. la educación superior): Enseñanzas de Escuelas Oficiales de Idiomas
- Segunda etapa de Ed.secundaria con OP (y sin cont. la educación superior): Enseñanzas profesionales de música y danza
- Enseñanzas de formación profesional, artes plásticas y diseño y deportivas de grado superior y equivalentes
- Grados universitarios de hasta 240 créditos ECTS, diplomaturas universitarias y equivalentes
- Grados universitarios de más de 240 créditos ECTS, licenciaturas y equivalentes
- Másteres, especialidades en Ciencias de la Salud por el sistema de residencia y similares
- Doctorado

¹

https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177034&menu=ultiDatos&idp=1254735976614

Para las personas que estén cursando estudios universitarios, se proporciona información también de la naturaleza de los estudios según una clasificación basada en la Clasificación de programas y titulaciones en sectores de estudios¹ (CNED-F):

- Educación
- Artes y humanidades
- Ciencias sociales, periodismo y documentación
- Negocios, administración y derecho
- Ciencias
- Informática
- Ingeniería, industria y construcción
- Agricultura, ganadería, silvicultura, pesca, y veterinaria
- Salud y servicios sociales
- Servicios

c) Método y fuentes

La recopilación de información sobre nivel de estudios y estudios en curso ha sido uno de los elementos más laboriosos del censo 2021 y se ha logrado sobre todo gracias al liderazgo del Ministerio de Educación y Formación Profesional (MEyFP) durante los últimos años.

Ha sido necesario reforzar la colaboración con las consejerías de educación de las CCAA para mejorar la información administrativa con el fin de lograr asignar estos estudios en curso a la población joven como se detalla más adelante.

La estimación del **nivel de estudios** se lleva a cabo mediante la incorporación de la información administrativa relativa al nivel educativo alcanzado más alto que se encuentre en los distintos registros administrativos. Las fuentes con las que se cuenta son las siguientes:

Padrón. El título escolar o académico es una variable de inscripción padronal que deben recoger los ayuntamientos en su padrón municipal y, por tanto, figura en los formularios de empadronamiento. La incorporación en el Padrón está relacionada con la formación del Censo Electoral. Así pues, en la gestión continua del Padrón se recoge esta información de los ayuntamientos y se depura con la información que se recibe semestralmente por parte del MEyFP.

• **Títulos del MEyFP.** Se dispone de información de los ficheros de títulos no universitarios (LGE, LOGSE, LOE), universitarios y homologaciones de títulos extranjeros a títulos universitarios y no universitarios, que, de manera continua, el Ministerio de Educación envía al Instituto Nacional de Estadística.

1

https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177034&menu=ultiDatos&dp=1254735976614

- **Censo 2001.** Para los casi 34,5 millones de personas de 16 años o más que respondieron al cuestionario censal en 2001 con fecha de referencia 1 de noviembre de 2001, tenemos información sobre el nivel de estudios alcanzado hasta ese momento.
- **Censo 2011.** Para los casi 3,5 millones de personas de 16 años o más que respondieron al cuestionario censal en 2011 con fecha de referencia 1 de noviembre de 2011, se tiene información sobre el nivel de estudios alcanzado hasta ese momento.
- **Demandantes de empleo del Servicio Público de Empleo Estatal (SEPE).** Se dispone de información del nivel de estudios terminados para las personas inscritas en los servicios públicos de empleo.
- **Certificados de profesionalidad del SEPE.** Se dispone de un fichero de certificados de profesionalidad histórico que acreditan competencias reconocidas por la Administración educativa surtiendo los efectos de convalidación del módulo o módulos profesionales correspondientes.
- **Egresados universitarios.** Se dispone de información histórica de personas egresadas en la Universidad desde el curso 2010-2011 hasta el curso previo al censo.
- **Graduados no universitarios.** Se dispone de información histórica de personas graduadas desde el curso 2014-2015 hasta el curso 2020-2021, previo al censo.
- **Registro Central de Extranjeros.** Se dispone de información de los extranjeros con una autorización de estudios en vigor en España.
- **Contratos en vigor del SEPE.** Se dispone de información desde 2001 del último contrato de cada persona a 1 de enero del año de referencia donde figura el nivel educativo del trabajador.
- **Matriculados en el sistema educativo en enseñanzas no universitarias del MEyFP.** Disponibles desde el curso 2014-2015 hasta el curso 2020-2021.
- **Matriculados en enseñanzas universitarias del Consejo de Universidades.** Se dispone de todas las matrículas desde el curso 2010-2011 hasta el curso 2020-2021.

Todas las fuentes son consideradas para la estimación del nivel educativo; en el caso de la información de estudios en curso, solo se utilizan las dos últimas fuentes referidas a matrículas.

La cuestión más difícil es asignar un nivel educativo a toda la población dado que no existe a priori ningún registro administrativo que ofrezca esta información con la desagregación requerida.

El padrón es la fuente de partida; garantiza disponer para toda la población de un valor de nivel educativo alcanzado pero en principio, resulta poco útil. En efecto, toda persona que se empadrona debe rellenar un formulario de empadronamiento (hoja padronal) en el que se solicita que rellene el código de su nivel educativo así:

Esta clasificación permite proporcionar los datos según la CNED-A. No obstante, no toda

Códigos de Nivel de Estudios
00 Se consignará este código para menores de 16 años
11 No sabe leer ni escribir
21 Sin estudios
22 Enseñanza primaria incompleta. Cinco cursos de EGB o equivalente o Certificado de Escolaridad o equivalente
31 Bachiller Elemental. Graduado Escolar. EGB completa. Primaria completa. ESO
32 Formación Profesional de Primer Grado. Formación Profesional de Grado Medio. Oficialía industrial
41 Formación Profesional de Segundo Grado. Formación Profesional de Grado Superior. Maestría industrial
42 Bachiller superior. BUP. Bachiller LOGSE
43 Otras titulaciones medias (Auxiliar de clínica, Secretariado, Programador de informática, Auxiliar de vuelo, Diplomado en artes y oficios, etc)
44 Diplomados en Escuelas Universitarias (Empresariales, Profesorado de EGB, ATS y similares)
45 Arquitecto o Ingeniero Técnico
46 Licenciado Universitario. Arquitecto o Ingeniero Superior. Grado Universitario
47 Titulados de Estudios Superiores no Universitarios
48 Doctorado y Estudios de postgrado o especialización para Licenciados. Máster universitario

la población figura con ese nivel de detalle. Es importante resaltar que las normas de gestión del Padrón municipal, permiten dos tipos de clasificaciones de esta variable por parte de los ayuntamientos, por lo que no se dispone de información relativa al título escolar con el mismo nivel de detalle para todas las personas. Así, sobre todo en empadronamientos antiguos, figuran solo cuatro categorías.

Aproximadamente para un 74% de la población se dispone del nivel detallado y en el restante 26% solo figuran cuatro valores posibles de nivel de estudios agregados. Pero además, esta información queda obsoleta salvo que la persona cambie su empadronamiento y rellene una nueva hoja padronal.

Para actualizar el nivel de estudios, el Padrón se viene cruzando desde hace varios años con el registro de Títulos del MEyFP, corrigiendo la información de partida en muchos casos. Por tanto, el Padrón ofrece un punto de partida que precisa mayor refinamiento.

Para la asignación del nivel educativo, primero se analizan de manera conjunta todas las fuentes relacionadas con el nivel alcanzado, sin tener en cuenta la información de matrículas.

Puesto que una persona puede tener diferentes niveles educativos registrados en los diferentes registros, se consideran como válidos aquellos que son coherentes con su edad escolar. De este modo, en primera instancia se puede dar un nivel educativo ya sea de tipo intervalo (aquellos casos donde la persona puede tener un nivel entre varios valores) o una clasificación directa.

Además todas las personas se cruzan con la información de los ficheros de matrículas. Para cada persona se toma, si se tiene, su matrícula más reciente a la fecha del censo y la de mayor nivel, de tal modo que en algunos casos se puede imputar el nivel

educativo ya alcanzado porque es un requisito para la inscripción en determinados cursos en los que la persona se encuentra matriculada.

Finalmente, para aquellos casos residuales que carecen de nivel educativo asignado o aquellos casos que se mantienen en duda entre varios valores se efectúan imputaciones probabilísticas.

En la siguiente tabla ilustramos en la medida en que distintas fuentes tienen su aportación al proceso de la asignación del nivel educativo. Está reflejado el orden de preferencia de las fuentes, que es altamente influyente, puesto que de entre los registros con información en alguna fuente, el 86% consta en más de una fuente educativa.

Figura 10. Fuente ganadora para nivel educativo en el Censo 2021 (población de 15 años y más)

Fuente ganadora	Registros	Porcentaje
Total	40.687.137	100%
Egresados universitarios	2.096.293	5,2%
Graduados no universitarios	3.302.977	8,1%
Títulos MEyFP	7.873.911	19,4%
Certificados de profesionalidad	236.289	0,6%
Demandantes de empleo	2.212.506	5,4%
Contratos	7.145.138	17,6%
Padrón	3.415.935	8,4%
Registro central de extranjeros	2.584	0,0%
Censos pasados	9.079.601	22,3%
Fichero Precensal 2020	3.983.431	9,8%
Matriculas universitarias	12.448	0,0%
Matrículas no universitarias	441.097	1,1%
Cuerpo pertenencia mutualidades	101.309	0,2%
Imputación desde intervalo	750.799	1,8%
Imputación desde no consta	32.819	0,1%

Para la asignación de los **estudios en curso**, se contempla que cada persona puede estar cursando distintos tipos de estudios simultáneamente. Algo más de 100.000 personas cursan dos tipos de estudios durante el curso 2020-2021. No obstante, el número de personas para las cuales se detectan más de dos tipos de estudios en curso durante el curso 2020-2021 es muy reducido (no llega a 2.000 personas). Por esta razón se plantea ofrecer para cada persona hasta dos tipos de estudios en curso.

Se proporciona una “Enseñanza principal” que generalmente se corresponde con la enseñanza de Régimen General (E.S.O, Bachillerato, F.P, estudios universitarios etc.) de mayor nivel que está cursando la persona (si cursa más de una) y una enseñanza secundaria que, o bien es una enseñanza de Régimen no general (Escuela oficiales de idiomas, enseñanzas de música y danza...) o bien es una enseñanza de régimen general de menor nivel que la principal.

Figura 11. Fuente ganadora para estudios en curso en el Censo 2021 (población de 15 años y más):

Fuente ganadora	Registros	%
Total	4.186.237	100%
Matrículas universitarias	1.267.577	30,3%
Matrículas no universitarias	2.588.531	61,8%
Cursa estudios, pero no hay información sobre los mismos	20.564	0,5%
Otras matrículas*	309.565	7,4%

En “Otras matrículas” se han incluido los registros de los que se ha podido deducir que estaban matriculados en el curso 20-21, gracias a la explotación del fichero de Títulos MEyFP a partir de una descarga posterior a la fecha de referencia censal. También se ha hecho una explotación del Registro Central de Extranjeros para detectar nuevos estudiantes a partir de su autorización de residencia por estudios. En ambos casos la información presente en ellas nos ha permitido detectar al menos un intervalo posible de los estudios en curso.

La asignación de los estudios en curso para la población más joven se ha resuelto en un porcentaje pequeño de casos mediante imputación y no puede considerarse aún del todo resuelta de cara al futuro.

Por una parte, los ficheros recibidos desde las distintas CCAA, aunque han ido mejorando drásticamente durante los últimos años, no han logrado la exhaustividad total. Así, no se dispone de datos de algunos centros privados o de algunas enseñanzas en particular. Esto implica que no puede garantizarse disponer de datos del 100% del alumnado.

Por otro lado es preciso enlazar la información que proviene del sistema educativo con el fichero censal de población y el enlace tampoco es perfecto. Sobre todo para personas extranjeras, no siempre se dispone del identificador (DNI, NIE) que permite mejorar el enlace, sino que las personas figuran en los ficheros de matrículas con su pasaporte. Así, en los ficheros de matriculados en enseñanzas no universitarias se tiene un 1,6% de matrículas no enlazadas. Y, por el contrario, en los ficheros del sistema universitario no figuran los nombres y apellidos, que permitirían alcanzar un porcentaje de enlace ligeramente superior al actual, lo que supone que un 4,0% de las matrículas universitarias no se consiguen enlazar.

También hay que considerar la posibilidad de personas que estén matriculadas en universidades españolas pero residan en el extranjero y cursen formación a distancia. En total, queda un pequeño residuo del 2,5% (115.834 matrículas) de los ficheros de estudios en curso que no se logran enlazar con ninguna persona del fichero de población, lo cual en parte puede deberse a títulos obtenidos o estudios cursados por no residentes pero en otros casos es el problema típico de falta de enlace por no disponer de buenos identificadores (ausencia de DNI o de nombres y apellidos para muchos registros).

Pero también ocurre al revés: hay personas que con seguridad deben estar cursando estudios pero no se logran enlazar con ninguna matrícula. Hay que tener en cuenta que el Reglamento, aunque no pide datos de estudios en curso, sí requiere nivel educativo alcanzado por todas las personas de 15 y más años, razón por la cual todo el bloque de variables educativas del censo 2021 se ha construido para ese colectivo. En España, en

líneas generales, la educación es obligatoria hasta los 16 años. Esto implica que los alumnos que en la fecha censal tienen 15 años deben forzosamente cursar algún tipo de estudios. Pero por las limitaciones en los ficheros del sistema educativo y de su cruce con el fichero de población censal, en particular queda un 4% de jóvenes de 15 años, algo más de 20.000 personas, (ver figura 11) a los que no se les consigue encontrar qué estudios están cursando, aun cuando por Ley deben estar haciéndolo. A estas personas se les ha asignado una categoría específica (“Cursa estudios, pero no hay información sobre los mismos”).

Figura 12. Estudios principales en curso para población de 15 a 18 años

Enseñanza Principal	Registros	%
No estudia	203.663	10,5%
Cursa estudios, pero no hay información sobre los mismos	20.554	1,1%
Educación primaria e inferior	6.846	0,4%
Primera etapa de educación secundaria y similar	193.476	10,0%
Segunda etapa de educación secundaria con OG: 4º ESO	430.562	22,2%
Segunda etapa de educación secundaria con OG: Bachillerato	630.086	32,5%
Enseñanzas de formación profesional, artes plásticas y diseño y deportivas de grado medio y similares ; Programas de cualificación profesional inicial y similares ; F.P.básica	244.661	12,6%
Enseñanzas de escuelas Oficiales de Idiomas	2.242	0,1%
Segunda etapa de ed.secundaria con OP (y sin cont. la educación superior):Enseñanzas profesionales de música y danza	691	0,0%
Enseñanzas de formación profesional, artes plásticas y diseño y deportivas de grado superior y equivalentes	45.788	2,4%
Grados universitarios de hasta 240 créditos ECTS, diplomaturas universitarias y equivalentes	151.541	7,8%
Grados universitarios de más de 240 créditos ECTS, licenciaturas y equivalentes	11.538	0,6%
Cursando otras enseñanzas superiores	10	0,0%

d) Análisis de la calidad de las variables y mejoras previstas

La estimación del nivel educativo alcanzado supone el tratamiento de muchas fuentes, cada vez más pesadas en volumen, puesto que el fenómeno educativo exige recoger el pasado educativo de las personas y su evolución.

El censo 2001 es la fuente ganadora en personas que ya entonces superaban la edad de alcanzar nuevas titulaciones y para las cuales no hay otra información que actualice o mejore esta.

Es de destacar que menos del 2% de registros resultan imputados y esto se realiza asignándoles aleatoriamente un nivel educativo (dentro del intervalo correspondiente) mediante las distribuciones observadas (por provincia de residencia y edad) en el Censo 2011.

Comparando con el nivel de formación ofrecido por la EPA en el primer trimestre de 2021:

Figura 13. Comparación Nivel de formación T1 2021 EPA y Censo 2021

Nivel formación	% EPA	% Censo 2021	dif %
Total	100	100	0
Analfabetos	1,3	0,32	-0,97
Educación primaria incompleta	4,57	4,05	-0,52
Educación primaria	10,72	12,13	1,41
Primera etapa de educación secundaria	28,9	29,74	0,84
Segunda etapa de educación secundaria. Orientación general	14,04	12,86	-1,18
Segunda etapa de educación secundaria. Orientación profesional	8,31	9,01	0,71
Educación superior	32,16	31,87	-0,29

Se observa que las diferencias estructurales son mínimas: el censo arroja un menor porcentaje de población con niveles de formación bajos (Analfabetos y Educación primaria incompleta) que la EPA, situando a estas personas en el nivel de educación Primaria, quizá por una auto clasificación a la baja en la encuesta. También se observa cierto trasvase entre las categorías de orientación general y profesional en la segunda etapa de educación secundaria. La comparativa en la educación superior es muy cercana a la de la EPA, con lo que la captación en las categorías más altas también parece adecuada, si bien ligeramente inferior en el Censo que en la EPA.

En cuanto a los **estudios en curso**, ya se ha mencionado una limitación que afecta a los estudiantes de 15 años.

Siendo conscientes de que esta limitación debe afectar también a estudiantes de 16 o más años, no puede utilizarse esa misma categoría de estudios desconocidos, dado que a esas edades no hay ya obligación legal y esas personas pueden no estar cursando estudios de ningún tipo. Si se observa el porcentaje de jóvenes de 16 años que están cursando estudios (92% según la figura 14) y se compara con el de 15 años (100%) se produce un escalón que en parte está producido por el abandono educativo pero en parte también por la pérdida de enlace entre unos ficheros y otros y por la falta de exhaustividad en los ficheros de partida.

Figura14. Porcentaje de población cursando estudios por edad

Edad	%
15	100
16	92,27
17	88,32
18	76,98

Evidentemente, para edades más altas este componente tiene mejor peso frente al de la población que no estudia, luego el impacto principal se debe dar en la edad 16 años.

Las mejoras previstas están implícitas en la descripción del problema que se acaba de hacer. Aunque se han logrado grandes avances en la recogida de información de carácter educativo que no estaba disponible antes, como es el registro nacional de las personas matriculadas en enseñanzas no universitarias y universitarias gracias a la cooperación con el MEyFP, la cobertura de estos registros no es completa aún y se sigue trabajando de forma conjunta con el MEyFP para su mejora.

Conforme se vaya recibiendo más información acumulada en el tiempo sobre matriculas en el sistema educativo y graduados, la estimación del nivel educativo se prevé que mejore.

5.3.4 Lugar de estudios

a) Requisitos

Reglamento UE

El Reglamento no requiere indagar sobre el lugar de estudios sino solamente sobre el lugar de trabajo.

Censos anteriores

Para cumplir con los requisitos y las necesidades específicas nacionales, el censo de 2011 no solo indaga sobre el lugar de trabajo sino también sobre el lugar de estudios. En efecto, se incluyó la siguiente pregunta en el cuestionario:

19 ¿Dónde está su lugar de trabajo o de estudio?

Si trabaja y estudia, le preguntamos por el sitio donde trabajó la semana pasada

En el propio domicilio

En varios municipios (comercial, conductor...)

En este municipio

En otro municipio:

Provincia:

En otro país:

Código postal de su centro de trabajo o de estudios:

Ya ha terminado el Cuestionario Individual de la Persona 1

No se ofreció información sobre el lugar de estudios en particular sino sobre el lugar de “estudios o trabajo”, si bien se podía suponer que, en el caso de los estudiantes, este lugar era el de estudios. Se proporcionó detalle municipal, incluso por código postal, para las grandes ciudades.

b) Planteamiento propuesto para la variable y desglose

Dado que esta información se construye con datos administrativos, nada impide tratar por separado el lugar de estudios y el de trabajo. De hecho, este lugar de estudios se ofrece para todas las personas que cursan estudios, independientemente de si además

trabajan. Se ofrece el lugar de estudios asociado a la matrícula principal, en caso de que se disponga de dos tipos de estudios en curso.

Para cumplir con los objetivos propuestos para el Censo de 2021 y mantener la comparabilidad con los resultados de los censos anteriores, se propone llegar a una desagregación parecida a la del censo de 2011: se ofrece el municipio de destino, pero no es posible ofrecer el código postal, como ocurriera en 2011.

Loa análisis sobre movilidad cotidiana que puedan derivarse de esta variable se realizan normalmente junto con la información sobre lugar de trabajo (ver apartado 6).

c) Método y fuentes

Obtenemos la variable de lugar de estudios de las mismas fuentes que los estudios en curso, fundamentalmente de los ficheros de matrículas universitarias y matrículas no universitarias del MEyFP, También el Registro Central de Extranjeros contiene información para conocer la provincia de estudios de la mayoría de los estudiantes así detectados.

Para un 91,8% de los estudiantes ha sido posible extraer el municipio de estudios y para un 0,4% adicional el lugar de estudios se ofrece a nivel de provincia.

Figura 15. Porcentaje cobertura Lugar de Estudios

Lugar de estudios	%
Municipio de estudios	91,8
Provincia de estudios	0,4
No consta	7,8

d) Análisis de la calidad de la variable y mejoras previstas

Tanto el lugar de trabajo, aún no publicado, como el de estudios ofrecen ciertas carencias inherentes a los registros administrativos de las que se obtienen.

Por otro lado, no debemos olvidar que el uso principal que se da a estas variables es para determinar matrices de movilidad origen-destino y que también la variable “origen”, que proviene siempre del padrón puede no ser correcta y pueden darse flujos origen-destino aparentemente extraños. Pero también hay que considerar que el contexto de teletrabajo y formación a distancia, acrecentado durante 2020 por la pandemia, puede aumentar esas situaciones en principio extrañas. Intentar delimitar errores y proceder a correcciones podría resultar aventurado y se ha preferido no introducir correcciones.

En definitiva las matrices origen-destino construidas a partir de fuentes administrativas no dejan de ser una primera aproximación al fenómeno con estas fuentes. Ayuda el hecho de disponer de unos flujos observados a partir del estudio de movilidad a partir de la telefonía móvil EM1 que, con seguridad, cuantifica mejor los movimientos entre áreas, y lo hacen con mucho detalle geográfico, mucho mejor que el municipal en el caso de municipios grandes, aunque no ofrece detalles sobre las personas que se desplazan ni sobre los motivos (estudios, trabajo, otros).

Seguramente para poder afirmar algo sobre la calidad de la información sea necesario un análisis conjunto con ambas fuentes, censo y estudio de movilidad, que no se ha podido acometer en toda su extensión por parte del INE.

Será necesario mejorar en el futuro la información sobre lugar de estudios en fuentes administrativas para lo cual hay mucho margen y también será necesario probablemente repetir cada cierto tiempo un estudio de movilidad a partir de la telefonía móvil que ofrezca esa información complementaria.

5.3.5 Relación con la actividad económica

a) Requisitos

Reglamento UE

Se debe proporcionar información sobre la situación actual de toda la población de 15 o más años respecto a la actividad económica, basada en un periodo de referencia de una semana reciente, que normalmente es la semana anterior a la fecha de referencia del censo (por tanto, en este caso sería la última semana de 2020).

Las categorías obligatorias para esta variable son:

- Población económicamente activa:
 - Empleado
 - Desempleado
- Población económicamente inactiva:
 - Pensionista o rentista
 - Estudiante
 - Otras situaciones de inactividad
 - Sin alcanzar la edad mínima nacional para ejercer una actividad laboral

El Reglamento exige esta desagregación como máximo a nivel provincial (por sexo y edad simple). A nivel municipal solo exige dar ocupados, parados e inactivos (y “no consta”).

Censos anteriores

En el anterior censo de 2011 el Reglamento europeo (CE) No. 1201/2009 era algo más exigente que el del censo de 2021 y exigía desglosar a las personas desempleadas según si habían tenido un empleo previo o no; se incluía también como categoría de inactividad opcional si la persona llevaba a cabo tareas del hogar u otras situaciones de inactividad.

Para cumplir con los requisitos y las necesidades específicas nacionales, el censo de 2011 incluyó las siguientes categorías en su cuestionario:

14 ¿En qué situación laboral estaba la semana pasada?

Ocupado/a (es decir, trabajó al menos una hora) o temporalmente ausente del trabajo:

a tiempo completo a tiempo parcial

Parado/a que ha trabajado antes

Parado/a buscando su primer empleo

Persona con invalidez laboral permanente

Jubilado/a, prejubilado/a, pensionista o rentista

Otra situación

(Pase a la pregunta 18)

Las recomendaciones internacionales para la variable “Relación con la actividad” indican que se debe seguir la definición de la OIT, según la cual, para poder clasificar a una persona se deben investigar cuestiones como la búsqueda activa de empleo o la disponibilidad para trabajar, entre otras. Para construir con precisión esa información, las encuestas laborales (como la EPA) realizan una batería de preguntas. Tradicionalmente, en los censos de población clásicos se realiza una aproximación, basada en la auto clasificación del informante, con una pregunta como la n° 14 del cuestionario del censo 2011 presentada anteriormente.

b) Planteamiento propuesto para la variable y desglose

Para cumplir con los objetivos propuestos para el Censo de 2021 y mantener algún desglose de los censos anteriores, se han planteado estas siete categorías, es decir, las del Reglamento más una, que se daría para cualquier desagregación (municipal o inferior) siempre que sea posible:

- Población económicamente activa
 - Empleado
 - Desempleado
- Población económicamente inactiva:
 - Jubilado, prejubilado
 - Estudiante
 - Percibiendo pensión de invalidez permanente
 - Otras situaciones de inactividad
 - Sin alcanzar la edad mínima nacional para ejercer una actividad laboral

c) Método y fuentes

No existe una única fuente que proporcione información de toda la población en su conjunto. La estimación se aborda a través de la unión de distintos registros administrativos.

Se han tenido en cuenta las siguientes fuentes:

- Fuentes de tipo económico:
 - Trabajadores que cotizan a la Seguridad Social que estaban de alta (Tesorería General de la Seguridad Social) durante la semana previa a 1 de enero de cada año.
 - Demandantes de empleo del SEPE.
 - Funcionarios que no cotizan a la Seguridad Social y pertenecen a alguna de estas tres mutualidades: MUFACE, MUGEJU e ISFAS.
 - Personas del régimen de la Seguridad Social que cobraron una pensión y causa de la misma (jubilación, invalidez, orfandad, viudedad). (INSS)
 - Personas que cobran una pensión pero no son del régimen de la Seguridad Social (Clases Pasivas).
- Información educativa sobre estudiantes matriculados en el sistema educativo, tanto para enseñanzas universitarias como no universitarias.
- Información tributaria del modelo 190 de la Agencia Tributaria y modelos similares de las Haciendas Forales.

Para cada persona con 15 años y más, se busca información en las fuentes económicas y educativas, en primer lugar.

Al tener en cuenta tantas fuentes, es normal que coexistan diferentes situaciones que llevan a conflicto para cada persona. Por ejemplo, una persona que se inscribió como demandante de empleo y a su vez, se matriculó a estudiar un cierto curso.

Lo más común es que cada persona aparezca solo en una fuente o que, si aparece en varias, la información sea la misma. Pero en cerca de tres millones de casos hay que asignar una única situación para cada persona. Para ello, se establecen una serie de reglas de prioridad basadas en las recomendaciones de Naciones Unidas y el Reglamento Europeo para los Censos:

- Si la persona se encuentra como empleada en alguna fuente, se la considera empleada.
- Si la persona percibe varias prestaciones, la de jubilación prevalece frente al resto para los mayores de 65 años (de 55 en el caso de parado).

También se aplican estas reglas en función de la edad:

- Si la persona tiene más de 65 años y no se la encuentra en ninguna fuente administrativa, se la considera inactiva.
- Si la persona tiene 15 años exactos, no es encontrada en ninguna otra fuente administrativa, se la considera estudiante.

Es importante considerar la fecha de referencia para estos datos. El Reglamento establece lo siguiente:

La «situación de actividad actual» es la situación presente de una persona respecto a la actividad económica, basada en un período de referencia de una semana, que puede

ser una semana indicada, reciente, fija, de calendario, la semana pasada completa de calendario o los siete días previos al censo.

Dado que la fecha de referencia del censo es el 1 de enero de 2021 se trataría entonces de obtener la situación laboral de la población durante los días previos. Por ello se ha hecho una solicitud específica a la TGSS intentando reflejar la situación de afiliación durante la semana del 21 al 27 de diciembre de 2020 (última semana completa previa a la fecha de referencia), y se ha construido la información tomando como ocupados a todas las personas con algo de actividad durante esa semana.

d) Análisis de la calidad de la variable y mejoras previstas

Tomando la población de 15 y más años, el número de registros según la fuente ganadora a la hora de asignar la variable se distribuye así:

Figura 16. Fuentes ganadores en la asignación de la relación con la actividad económica

Fuente ganadora	Frecuencia	Porcentaje
Total de personas de 15 años o más	40.687.137	100
TGSS	17.582.887	43,21
INSS	8.456.582	20,78
SEPE	4.296.695	10,56
Sin información	3.769.575	9,26
Matrículas no universitarias	1.993.044	4,9
Imputación determinística	1.409.823	3,46
MUFACE	940.689	2,31
Tributarias	842.738	2,07
Matrículas universitarias	777.251	1,91
ISFAS	330.022	0,81
Otras matrículas	154.854	0,38
Clases Pasivas	78.613	0,19
MUGEJU	54.364	0,13

Se tienen, por tanto, 3.769.575 no asignados que se reparten así por sexo y grupo de edad (quinquenal):

Figura 17. Distribución por sexo y edad de las personas sin relación con la actividad asignada.

	Ambos sexos	Hombres	Mujeres
Total edades	3.769.575	1.565.069	2.204.506
De 15 a 19 años	219.568	120.896	98.672
De 20 a 24 años	457.229	240.633	216.596
De 25 a 29 años	444.792	229.401	215.391
De 30 a 34 años	392.224	192.502	199.722
De 35 a 39 años	371.298	169.284	202.014
De 40 a 44 años	380.705	163.439	217.266
De 45 a 49 años	364.541	139.847	224.694
De 50 a 54 años	362.247	118.256	243.991
De 55 a 59 años	373.521	100.147	273.374
De 60 a 64 años	403.450	90.664	312.786

El porcentaje de no asignados se sitúa en el 9,26% de la población de 15 y más para el censo de 2021. El hecho de que una persona figure como no asignada en una categoría como esta no debe resultar extraño. De hecho es lógico que muchos de los inactivos no aparezcan en esos registros. La curva por sexo y edad de los no asignados es, además, muy similar a la de los inactivos, lo que indicaría que una gran parte de estos no asignados son inactivos. Por esa razón se les imputa a todos ellos la categoría de *otras situaciones de inactividad*.

La distribución resultante, una vez realizada la imputación anterior, no se aleja mucho de la que ofrece la EPA centrada a 1-1-2021 (se toman datos medios de la EPA del 4º trimestre de 2020 y 1º de 2021). Se ofrece una comparativa de censos anteriores y del resultado del censo de 2021 con la EPA de ese momento.

Figura 18. Comparativa de estimaciones EPA y censos. Diferencias en tasas de variación. Años 2001, 2011 y 2021.

	EPA 4º trim 2001	Censos 2001	Dif	EPA 4º trim 2011	Censos 2011	Dif	EPAMM* (T4 2020 y T1 2021)	Censos 2021	Dif
Ocupados	16.348,2	16.329,7	-0,1%	17.807,5	17.514,6	-1,6%	19.275,6	18.636,6	-3,3%
Parados	2.283,9	2.692,8	17,9%	5.716,8	7.377,3	29,0%	3.686,9	3.811,6	3,4%
Inactivos	15.931,6	15.201,4	-4,6%	15.427,0	14.191,8	-8,0%	16.669,1	17.430,5	4,6%

El sesgo al alza por la auto clasificación como parado que aparece en los censos de 2001 y 2011 se corrige, y de hecho, como era de esperar, se observa el efecto contrario en los inactivos: los datos del censo de 2021 para inactivos son algo más elevados que los de la EPA. Las diferencias se concentran en el colectivo de personas jóvenes y en extranjeros. No obstante, para la correcta medición de la situación laboral, un censo basado en registros administrativos tienen una limitación inherente a las fuentes que solo puede subsanarse mediante imputación. La condición de parado se obtiene principalmente de encontrar a la persona registrada y considerada como tal en el SEPE y no se sigue exactamente la clasificación OIT que precisa conocer muchas más información sobre la persona, como si está buscando trabajo activamente

Con todo, un censo basado en datos administrativos ofrece una clasificación de la población según su relación con la actividad económica más ajustada a lo que nos da una encuesta como la EPA que la recoge un censo clásico basado en cuestionarios que clasifica esa variable usando una única pregunta (la denominada "auto clasificación"), como puede verse en la tabla anterior.

5.3.6 Variables relativas a los ocupados

a) Requisitos

Reglamento UE

Para todas las personas empleadas de 15 y más años durante la semana de referencia, se debe proporcionar información sobre su ocupación, sector de actividad económica, situación profesional y ubicación del lugar de trabajo.

Estas cuatro variables censales se deben asignar en base al mismo empleo de la persona, su empleo principal, que se puede determinar según el tiempo que pase en este empleo o, si no se conoce, su retribución económica.

El Reglamento exige dar las siguientes categorías a nivel nacional únicamente:

La variable ocupación se desagrega de forma obligatoria con las categorías de la CNO11¹ a un dígito:

- Total
 - Directores y gerentes
 - Profesionales científicos e intelectuales
 - Técnicos y profesionales de nivel medio
 - Personal de apoyo administrativo
 - Trabajadores de los servicios y vendedores de comercios y mercados
 - Agricultores y trabajadores calificados agropecuarios, forestales y pesqueros
 - Oficiales, operarios y artesanos de artes mecánicas y de otros oficios
 - Operadores de instalaciones y máquinas y ensambladores
 - Ocupaciones elementales
 - Ocupaciones militares

La rama de actividad económica se refiere al tipo de producción o actividad del establecimiento o unidad similar en que este situado el puesto de trabajo de la persona empleada. La desagregación obligatoria de esta variable es a nivel de secciones de la NACE² rev.2 (21 categorías).

1

https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177033&menu=ultiDatos&idp=1254735976614

2

https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177032&menu=ultiDatos&idp=1254735976614

La situación en el empleo tiene las siguientes categorías obligatorias:

- Total
 - Empleados
 - Empresarios
 - Trabajadores por cuenta propia
 - Otros trabajadores por cuenta ajena

La desagregación obligatoria de la **ubicación del lugar de trabajo** es a nivel NUTS-2, es decir, comunidades autónomas en el caso español:

- Total
 - En el territorio del estado miembro
 - Todas las regiones NUTS-2 del estado miembro
 - Se desconoce el lugar de trabajo en el estado miembro
- Fuera del territorio del estado miembro
- Sin lugar de trabajo fijo (dentro o fuera del estado miembro)
- Se desconoce el lugar de trabajo
- No aplicable (si no trabaja)

Censos anteriores

Para cumplir con los requisitos y las necesidades específicas nacionales, el censo de 2011 incluyó las siguientes preguntas relativas a los ocupados en su cuestionario:

15 ¿Cuál es (era) su ocupación?

Si está parado refiérase a su última ocupación
Búsquela en la TABLA DE OCUPACIONES del FOLLETO ADJUNTO y anote la letra y el número que le corresponde:

Letra Número *Si no ha encontrado su ocupación o tiene dudas, escribala a continuación:*

16 ¿Cuál es (era) la actividad principal del establecimiento o local donde trabaja (trabajaba)?

Si está parado refiérase a su última ocupación
Búsquela en la TABLA DE ACTIVIDADES del FOLLETO ADJUNTO y anote el número que le corresponde:

Número *Si no ha encontrado su actividad o tiene dudas, escribala a continuación:*

17 ¿Cuál es (era) su situación profesional?

Marque sólo una opción

Empresario, profesional o trabajador por cuenta propia:

que emplea personal que no emplea personal

Trabajador por cuenta ajena:

fijo o indefinido eventual o temporal

Otras situaciones:

ayuda familiar miembro de cooperativas

19 ¿Dónde está su lugar de trabajo o de estudio?

Si trabaja y estudia, le preguntamos por el sitio donde trabajó la semana pasada

En el propio domicilio

En varios municipios (comercial, conductor...)

En este municipio

En otro municipio: _____

Provincia: _____

En otro país: _____

Código postal de su centro de trabajo o de estudios: _____

Ya ha terminado el Cuestionario Individual de la Persona 1

Se proporcionó mucho más detalle del exigido en el Reglamento. En concreto, la ocupación se proporcionó a tres dígitos de la clasificación CNO2011, la actividad económica a tres dígitos de la CNAE09. Además, los datos se ofrecen con detalle municipal y el lugar de trabajo (o de estudios) incluso por código postal, para las grandes ciudades.

b) Planteamiento propuesto para las variables y desglose

Para cumplir con los objetivos propuestos para el Censo de 2021 y mantener la comparabilidad con los resultados de los censos anteriores, se publica la ocupación a tres dígitos de la CNO2011, la actividad económica a tres dígitos de la CNAE-09, ubicación del lugar de trabajo a nivel de municipio o en su defecto de provincia y, para la variable situación profesional, la información obligatoria incluyendo las subcategorías de empresarios según si emplean o no personal, y adicionalmente las subcategorías de asalariados según sean fijos/indefinidos o temporales/eventuales.

c) Método y fuentes

La estimación de estas variables se lleva a cabo mediante la incorporación de la información administrativa relativa al empleo de la que se dispone.

Por tanto, las fuentes con las que se cuenta para la estimación de estas cuatro variables son las ya consideradas para la estimación de personas empleadas de carácter económico, y adicionalmente otras fuentes que pueden contener información de la actividad económica del establecimiento:

- Trabajadores afiliados a la Seguridad Social que estaban de alta durante la semana previa a 1 de enero de 2021.
- Información de mutualistas de MUFACE, MUGEJU e ISFAS.
- Información relativa al último contrato vigente de cada persona a 1 de enero de 2021 en el SEPE.
- Información de censos previos 2011 y 2001.
- DIRCE (Directorio Central de Empresas).
- Retenciones por rendimientos del trabajo por cuenta ajena, donde se incluye la actividad económica de la empresa responsable (AEAT y Haciendas Forales)
- Registro Central de Extranjeros (D.G. Policía)
- Información de encuestas: EPA (Encuesta de Población Activa) y EES (Encuesta de Estructura Salarial)

Como punto de partida para estimar las variables relativas a los ocupados, se distinguen dos situaciones iniciales: si la persona es mutualista o si está afiliada a la Seguridad Social.

Si bien es cierto que un funcionario puede estar dado de alta ejerciendo otra actividad económica residual, se toma como empleo principal el recogido en el fichero de su mutualidad, prescindiendo en un principio de las posibles discrepancias de información respecto a la Tesorería o el fichero de contratos del SEPE.

En el caso de **los funcionarios**, por tanto, se parte inicialmente de la información de los ficheros de MUFACE, MUGEJU e ISFAS para estimar estas cuatro variables,

aunque adicionalmente, para dar el máximo detalle posible de algunas variables, se utilizan como fuentes auxiliares los censos pasados.

La estimación de la actividad económica según la CNAE se obtiene directamente de la información contenida en los ficheros de las mutualidades. Asimismo, en ellos se dispone de la ubicación del puesto de destino a nivel de provincia, tratando de extender al detalle de municipio con la información de las fuentes auxiliares. La ocupación se estima en función del cuerpo de pertenencia dentro de la Administración, excepto en el caso de los funcionarios civiles recogidos en ISFAS, de los que no se dispone de información sobre su ocupación. Y la situación profesional se asigna de manera directa como “empleados por cuenta ajena”.

Para las **personas empleadas según la Tesorería**, la estimación de estas cuatro variables se realiza principalmente con la información del fichero de Afiliados y Afiliaciones en alta recibido de la TGSS, del fichero de contratos del SEPE y de censos previos. Es un proceso más laborioso que el requerido para los funcionarios, donde además se distingue entre la situación profesional de las personas para captar su información de forma más coherente.

La situación profesional viene marcada por el régimen de cotización, el tipo de contrato, tipo de relación laboral y código de Cuenta de Cotización de la Tesorería.

La ubicación del lugar de trabajo se estima a partir de la provincia asociada a la cuenta de cotización secundaria en la que está dado de alta, si bien para los asalariados se cuenta adicionalmente con el domicilio de la actividad al detalle de municipio en el fichero de afiliaciones de la Tesorería. Por otra parte, para los que se encuentra información en los contratos del SEPE, consta también el municipio de trabajo.

Es importante destacar que el lugar de trabajo censal, al estar recogido en registros no va a reflejar situaciones como el teletrabajo o el trabajo en varios municipios distintos.

La actividad económica de los afiliados viene recogida en el fichero de altas de la Seguridad Social y consta también, aunque al nivel de detalle de 2 dígitos CNAE, en los contratos del SEPE. Para los autónomos, en caso de ser encontrados como Personas Físicas en el DIRCE, hay información de la actividad CNAE, del municipio de trabajo, y consta si emplean o no personal.

Para estimar la ocupación del puesto de trabajo se acude al fichero de contratos del SEPE, asignando la ocupación del último contrato que se le haya detectado. En caso de no encontrar información, o para los autónomos, también se acude a información en los censos previos o en su defecto es imputada a partir del resto de información laboral disponible.

Para aquellos **ocupados estimados a través de fuentes tributarias**, por construcción son clasificados como autónomos. Se busca su información en el DIRCE, donde es posible encontrar tanto su actividad económica como si emplean o no personal, así como en el resto de fuentes.

Situación laboral

La estimación de esta variable no supone mucha dificultad, puesto que las fuentes en las que aparecen los ocupados nos permiten una **clasificación prácticamente directa**.

Las tres mutualidades contienen información de asalariados, la Tesorería aporta información detallada de asalariados y autónomos, y los ocupados procedentes de fuentes tributarias son considerados autónomos por construcción.

En la Tesorería, tanto para la situación laboral como para la actividad CNAE, se toma la información de la afiliación más frecuente durante la semana de referencia. En caso de haber varias posibles afiliaciones para un mismo afiliado, en caso de empate, se prioriza la jornada completa.

Figura 19. Situación laboral según la fuente que detecta los ocupados.

Fuente de procedencia ocupados	Situación laboral		Frecuencia	Porcentaje por fuente
ISFAS	Asalariados	Fijos	135.237	67,2%
		Temporales	65.890	32,8%
MUFACE	Asalariados	Fijos	598.695	99,8%
		Temporales	1.212	0,2%
MUGEJU	Asalariados	Fijos	46.098	99,9%
		Temporales	67	0,1%
Treasorería	Asalariados	Fijos	10.674.286	73,5%
		Temporales	3.842.944	26,5%
	Autónomos	Empleadores	621.258	20,6%
		No empleadores	2.388.203	79,4%
	Otra situación	Otra situación	55.120	100,0%
Tributarias	Asalariados	Temporales	3.449	100,0%
	Autónomos	Emplean	39.026	15,8%
		No emplean	208.205	84,2%

Los asalariados procedentes de MUFACE y MUGEJU son por defecto fijos, si bien se han clasificado a algunos, cerca de 1.300, en temporales gracias a información complementaria de Tesorería.

Los autónomos, independientemente de si proceden de Tesorería o de fuentes tributarias, se buscan como Personas Físicas en el DIRCE. En caso de encontrar algún local con asalariados según esta fuente, se les clasifica como empleadores. Para los autónomos que proceden de fuentes tributarias, además, se ha estimado la subclasificación en base a su edad, actividad económica y ocupación a un dígito, imputando a más de 23.000 como empresarios.

Figura 20. Fuente ganadora de la Situación laboral.

Fuente ganadora de la Situación laboral	Frecuencia	Porcentaje
Treasorería	17.418.020	93,2%
MUFACE	598.695	3,2%
DIRCE	224.793	1,2%
ISFAS	201.127	1,1%
Tributarias	159.558	0,9%
MUGEJU	46.098	0,2%
Imputada	31.399	0,2%

Actividad económica del establecimiento según CNAE

La actividad del establecimiento de trabajo tiene muy buena cobertura en registros administrativos. Para los funcionarios la asignación es directa, y para los ocupados procedentes de Tesorería se tiene cobertura completa. Para los ocupados estimados por fuentes tributarias hay además multitud de fuentes en las que comprobar si existe información de su CNAE. La dificultad radica en la selección de la fuente con mayor calidad para cada ocupado. Nótese que la actividad económica, y por tanto su fuente ganadora, incide directamente sobre la variable de lugar de trabajo; y que la actividad y la ocupación de las personas también están relacionadas, de tal forma que la asignación de la actividad económica repercute sobre ambas.

Es frecuente que haya más de una fuente disponible, en general coincidente, para un mismo ocupado. Incluso hay fuentes, como la Tesorería, las fuentes tributarias o el DIRCE, en las que puede haber más de una actividad distinta. Se estima la actividad económica del establecimiento teniendo en cuenta las distintas fuentes involucradas y la consistencia con el resto de variables de la persona ocupada.

En los asalariados se concluye que, en ciertas ramas de actividad, la calidad como fuente de los contratos del SEPE supera a la de la Tesorería, si bien priorizar al SEPE como fuente ganadora supone la dificultad añadida de estimar el tercer dígito CNAE de la actividad, pues en los contratos del SEPE la actividad figura a dos dígitos. En autónomos la actividad procedente del DIRCE tiene la ventaja adicional respecto de la Tesorería de poder situar en su municipio de trabajo (y no solo en su provincia) a los trabajadores por cuenta propia.

Para las personas en las que se ha priorizado la actividad económica de su último contrato y, por tanto, nos falta el detalle del tercer dígito CNAE, se ha estimado su actividad económica completa de acuerdo al municipio de trabajo del SEPE. Así, si observamos que dentro de la división CNAE estamos infraestimando algún grupo CNAE, imputamos probabilísticamente el tercer dígito de acuerdo al reparto municipal de esa actividad económica observado en el DIRCE (o en su defecto, de acuerdo al reparto provincial).

Para las personas sin información, se hace una imputación probabilística basada en los repartos de la actividad económica por provincia de residencia del Censo 2011. Como se puede observar en las tablas siguientes, el porcentaje de registros en los que la fuente ganadora son los censos pasados o en los que es necesario imputar es del 2,2%.

Figura 21. Fuente ganadora de la actividad económica según situación laboral.

Situación profesional	Fuente ganadora de la actividad económica	Frecuencia	Porcentaje de la fuente de la actividad dentro de la situación profesional
Asalariados	Tesorería*	12.529.647	81,5%
	Contratos SEPE*	1.641.620	10,7%
	MUFACE	599.907	3,9%
	ISFAS	201.127	1,3%
	Censos pasados	199.816	1,3%
	Tributarias	142.593	0,9%

	MUGEJU	46.165	0,3%
	Otros / Imputado	7.003	0,0%
Autónomos	DIRCE**	1.483.150	45,5%
	Tesorería**	1.422.211	43,7%
	Tributarias	111.057	3,4%
	Censos pasados	105.014	3,2%
	Otros / Imputado	124.530	3,8%
	No consta	10.730	0,3%
Otros	Tesorería	29.279	53,1%
	Contratos SEPE	24.583	44,6%
	Censos pasados	544	1,0%
	Tributarias	498	0,9%
	Otros / Imputado	216	0,4%

* En asalariados, Tesorería prioriza sobre SEPE como variable ganadora; entre los que tienen fuente ganadora de Tesorería hay un 67% de coincidencias en la CNAE con el último contrato del SEPE.

** En autónomos, DIRCE prioriza sobre Tesorería como variable ganadora; entre los que tienen fuente ganadora DIRCE hay una coincidencia en la CNAE del 85% con la Tesorería.

Figura 22. Fuente ganadora de la actividad económica del establecimiento.

Fuente ganadora de la actividad	Frecuencia	Porcentaje
Tesorería	13.981.137	74,8%
Contratos SEPE	1.684.171	9,0%
DIRCE	1.488.100	8,0%
Censos pasados	305.374	1,6%
MUFACE	599.907	3,2%
Tributarias	254.148	1,4%
ISFAS	201.127	1,1%
Imputado	105.465	0,6%
MUGEJU	46.165	0,2%
No consta	10.730	0,1%
Otros	3.366	0,0%

Lugar de trabajo

El lugar de trabajo se deriva de la misma fuente de procedencia que la actividad económica. En caso de que la propia fuente ganadora de la actividad CNAE incluya el detalle del lugar de trabajo a nivel de municipio, éste queda asignado. En caso de que la fuente de la que se obtiene la actividad proporcione solo la provincia de trabajo, sin llegar a detalle municipal, dadas las circunstancias apropiadas se estima, extendiendo desde la provincia, el municipio de trabajo.

Para el 88% de las personas ocupadas se obtiene el municipio de trabajo, mientras que ese porcentaje aumenta hasta el 98% en el caso de la provincia de trabajo, lo que supone que solamente 381.000 registros queden sin dato.

Las condiciones que se imponen a una fuente distinta a la ganadora de la actividad económica para extender de provincia a municipio son, además de lógicamente la coincidencia de provincia de trabajo que figura en la fuente con la de la fuente ganadora de la actividad, la coincidencia de la actividad económica. La extensión se aplica especialmente a los funcionarios de las tres mutualidades, puesto que en esas fuentes no se tiene la información al detalle municipal. En el 85% de estas extensiones, se encuentra el municipio de trabajo en censos pasados.

Figura 23. Fuente ganadora del Lugar de trabajo

Fuente ganadora del lugar de trabajo	Frecuencia	Porcentaje	Porcentaje que es extendido
Tesorería	13.981.641	74,8%	1,5%
Contratos SEPE	1.681.048	9,0%	*
DIRCE	1.488.350	8,0%	*
MUFACE	599.551	3,2%	39,2%
No consta	381.014	2,0%	-
Censos pasados	300.799	1,6%	*
ISFAS	201.127	1,1%	21,0%
MUGEJU	46.160	0,2%	43,4%

* La fuente ya contiene el detalle del municipio

Figura 24. Cobertura del Lugar de trabajo a nivel municipio/provincia por Situación laboral.

Situación laboral	Porcentaje con provincia	Porcentaje con municipio
Total	98,0%	88,4%
Asalariados	99,0%	95,4%
Autónomos	92,9%	55,4%
Otros	99,4%	64,8%

Ocupación

La principal fuente de información que se ha utilizado para determinar la ocupación son los contratos de trabajo proporcionados por el SEPE.

Se ha construido un fichero histórico (desde 2001) que contiene, para toda la población inscrita en algún momento en ese fichero, información sobre el último contrato que tuvieron. Este fichero contiene datos de todas las personas que, desde 2001, tienen o

alguna vez han tenido un contrato de trabajo registrado. Obviamente muchas de esas personas ya han podido fallecer o emigrar pero ofrece una información casi exhaustiva para los ocupados presentes en España en la fecha censal. El fichero de contratos del SEPE contiene la ocupación a tres dígitos de la CNO-11.

Hay que señalar que la ocupación es la variable laboral de las personas ocupadas más difícil de estimar a través de registros administrativos y, de hecho, es la única variable en la que no nos restringimos a los registros de los que no consta información para hacer imputaciones.

Para los asalariados es conocido que existe el sesgo a la baja de la ocupación que figura en su principal fuente de información, los contratos del SEPE: se encuentran sobrerrepresentados los puestos de ocupación elemental en detrimento de puestos cualificados.

Uno de los motivos de este sesgo se atribuye a la antigüedad del contrato. Para aquellas personas que firmaran contratos hace años, la información que se tiene de ellas es la de la ocupación a la firma del contrato. Las variaciones que se hayan podido producir en cuanto a ocupación dentro de una misma empresa, que no den lugar a modificaciones contractuales, no tienen reflejo en este fichero.

Es por ello que el grado de imputación de la ocupación es mayor que en el resto de variables, no solo por tener una menor cobertura, sino también por aplicar imputaciones en las personas con dato administrativo disponible pero claramente sesgado a la baja. En total casi el 19% de la ocupación es imputada, aunque no de forma homogénea. La imputación de la ocupación es sensiblemente mayor en los autónomos.

Para profundizar en el estudio de este sesgo a la baja en la ocupación recogida en los ficheros de contratos del SEPE, desde hace años se vienen analizando las discrepancias a nivel de microdatos entre las ocupaciones estimadas a partir de fuentes administrativas en los ficheros precensales y las recogidas en las encuestas de mercado laboral (Encuesta de Población Activa y Encuesta de Estructura Salarial).

De este estudio de las diferencias entre ocupaciones depuradas recogidas a través de encuestas y las registradas en los contratos, se ha llegado a la conclusión de que no todas las diferencias eran explicadas por la antigüedad de los contratos. Las fuentes que mejor nos han permitido identificar esta problemática han sido los censos pasados, en particular el Censo 2001.

La comparación de microdatos entre censos pasados y contratos SEPE ha permitido detectar sesgos sistemáticos a la baja en la ocupación de los contratos, como trabajadores cualificados en actividades agrarias con contratos de peones agrarios, albañiles con contratos de peones de la construcción, o asistentes administrativos especializados, agentes comerciales y vendedores con contratos de administrativos sin atención al público.

La edad, el nivel educativo, la actividad económica y la situación laboral de las personas con contratos SEPE son tenidas en cuenta a la hora de evaluar la existencia de un posible sesgo en el dato registrado, y también a la hora de imputar de forma probabilística una nueva ocupación.

Así, el 31% de las imputaciones que se han efectuado en ocupación, se han hecho en personas ocupadas con información en registros, ascendiendo ese porcentaje al 52% en el caso de los asalariados.

Para las personas ocupadas sin información, o de cuya información se prescinde por no ser consistente con el resto de información, se intenta imputar la ocupación por otras vías. Se toma en cuenta la información básica de la persona ocupada y su nivel educativo, y especialmente en los autónomos, el detalle del cuarto dígito de la actividad según la CNAE disponible en el DIRCE y Tesorería permite hacer imputaciones sobre la ocupación.

En asalariados, además de lo anterior, se usa el grupo de cotización para perfilar mejor las imputaciones determinísticas. Adicionalmente, para los autónomos se hace una imputación probabilística basada en la actividad CNAE y el nivel de formación, que supone el 14% del total imputado en dicha situación laboral.

Finalmente, por las características especiales de las actividades sanitarias y educativas (ramas P y Q), se hacen imputaciones determinísticas adicionales, motivadas especialmente en la división Q87 de Asistencia en establecimientos residenciales por la detección de sesgos a la baja no resueltos anteriormente en comparación con datos de encuestas de mercado laboral.

Figura 25. Fuente ganadora de la ocupación.

Fuente ganadora de la ocupación	Frecuencia	Porcentaje
Contratos SEPE	11.432.500	61,2%
Imputado	3.518.575	18,8%
Censos pasados	2.263.424	12,1%
No consta	619.241	3,3%
MUFACE	599.185	3,2%
ISFAS	200.600	1,1%
MUGEJU	46.165	0,2%

d) Análisis de la calidad de los resultados obtenidos y próximos pasos

En esta apartado, las distintas variables censales de los ocupados de 16 años y más que residen en viviendas familiares (se excluyen los residentes en colectivos) se comparan con la media móvil (MM) de los 4 trimestres centrados en el 1 de enero de 2021 de la EPA (Encuesta de Población Activa).

Situación laboral

Figura 26. Comparativa con MM EPA de la Situación profesional (miles)

Situación laboral	Censo 2021		EPA		Dif	Dif %
	Total	%	Total	%		
Autónomo emplea personal	659,7	3,5	900,2	4,7	-240,5	-1,1
Autónomo no emplea personal.	2.587,0	13,9	2.093,8	10,8	493,2	3,1
Asalariado	15.334,9	82,3	16.240,3	83,9	-905,4	-1,6
Ayuda familiar	0,0	0,0	79,6	0,4	-79,6	-0,4

Cooperativas	55,0	0,3	23,1	0,1	19,1	0,1
Otros			12,8	0,1		

La mayor diferencia aparece en el colectivo de asalariados. En el Censo 2021 no se ha llevado a cabo ninguna corrección sobre posibles falsos autónomos, considerando que si una persona aparece registrada como autónomo, es autónomo a efectos censales.

Actividad económica

Figura 27. Comparativa con MM EPA de la rama de actividad CNAE (miles)

CNAE	Censo 2021		EPA Media Móvil		Dif	Diferencia %
	Total	%	Total	%		
No consta	10,7	0,06				
A	892,0	4,79	780,5	4,03	111,4	0,75
B	19,8	0,11	32,7	0,17	-12,9	-0,06
C	1.993,7	10,70	2.398,3	12,39	-404,6	-1,70
D	36,2	0,19	89,7	0,46	-53,5	-0,27
E	132,1	0,71	151,5	0,78	-19,4	-0,07
F	1.222,4	6,56	1.279,0	6,61	-56,6	-0,05
G	3.120,5	16,74	2.964,1	15,32	156,4	1,43
H	950,0	5,10	998,6	5,16	-48,5	-0,06
I	1.422,4	7,63	1.364,8	7,05	57,6	0,58
J	586,3	3,15	625,4	3,23	-39,1	-0,09
K	359,0	1,93	472,9	2,44	-114,0	-0,52
L	134,9	0,72	154,9	0,80	-20,0	-0,08
M	1.064,5	5,71	1.060,4	5,48	4,1	0,23
N	1.224,9	6,57	1.001,6	5,18	223,4	1,40
O	1.316,9	7,07	1.375,5	7,11	-58,6	-0,04
P	1.308,4	7,02	1.382,5	7,14	-74,1	-0,12
Q	1.573,9	8,45	1.821,8	9,42	-247,9	-0,97
R	369,4	1,98	365,3	1,89	4,1	0,09
S	508,5	2,73	467,8	2,42	40,6	0,31
T	386,3	2,07	558,3	2,89	-172,0	-0,81
U	3,8	0,02	4,2	0,02	-0,4	0,00

Se observa bastante consistencia en esta comparativa entre censos y EPA.

Las diferencias en los porcentajes de reparto entre las distintas ramas CNAE varían entre el -1,7% en la rama C (Industria manufacturera) y el +1,4% en la rama G (Comercio).

En diferencias absolutas, la rama C sigue teniendo las mayores diferencias negativas, y por el contrario la N (Actividades adm. y servicios auxiliares) las mayores diferencias positivas.

Ocupación

Figura 28. Comparativa con MM EPA de la ocupación a dos dígitos CNO (miles)

	Censo 2021	EPA		
CNO	Total	Total	Diferencia	Diferencia %
0	112,4	113,3	-0,9	0,02
00	112,4	113,3	-0,9	0,02
1	717,2	784,6	-67,4	-0,21
11	50,9	45,1	5,8	0,04
12	230,3	208,2	22,1	0,16
13	183,9	238,4	-54,5	-0,25
14	139,7	180,8	-41,1	-0,18
15	112,4	112,2	0,2	0,02
2	3.420,6	3.754,8	-334,2	-1,05
21	745,1	799,9	-54,7	-0,14
22	894,9	885,6	9,3	0,23
23	242,4	220,9	21,5	0,16
24	438,7	581,5	-142,8	-0,65
25	182,5	227,8	-45,3	-0,20
26	369,0	492,9	-123,9	-0,57
27	180,9	190,4	-9,5	-0,01
28	189,8	208,5	-18,7	-0,06
29	177,2	147,1	30,0	0,19
3	1.935,3	2.208,3	-273,0	-1,03
31	313,4	334,6	-21,1	-0,05
32	76,6	127,9	-51,4	-0,25
33	120,5	152,1	-31,6	-0,14
34	22,7	93,1	-70,4	-0,36
35	534,5	600,4	-65,9	-0,23
36	312,0	308,2	3,7	0,08
37	337,1	280,7	56,4	0,36
38	218,5	311,2	-92,7	-0,44
4	1.806,3	2.042,9	-236,7	-0,87
41	356,4	528,6	-172,2	-0,82
42	85,5	74,4	11,1	0,07
43	598,5	513,7	84,8	0,56
44	293,9	342,5	-48,6	-0,19
45	472,0	583,7	-111,8	-0,48
5	3.987,1	4.067,1	-80,0	0,37
50	281,6	230,9	50,7	0,32

51	730,1	663,2	66,8	0,49
52	954,0	854,4	99,6	0,70
53	351,6	331,6	20,0	0,17
54	170,0	115,1	54,9	0,32
55	120,8	163,3	-42,5	-0,20
56	412,8	467,8	-55,0	-0,20
57	189,0	348,4	-159,4	-0,79
58	414,0	438,9	-24,9	-0,05
59	363,1	453,5	-90,4	-0,40
6	370,8	433,1	-62,3	-0,25
61	229,5	279,9	-50,4	-0,22
62	81,6	95,7	-14,1	-0,06
63	23,7	24,8	-1,1	0,00
64	36,0	32,7	3,3	0,02
7	1.947,0	2.142,0	-195,0	-0,62
71	572,0	584,2	-12,2	0,05
72	222,9	282,6	-59,6	-0,26
73	219,2	259,5	-40,3	-0,16
74	275,6	308,8	-33,2	-0,12
75	269,0	331,1	-62,1	-0,27
76	67,4	71,7	-4,3	-0,01
77	189,4	210,0	-20,6	-0,07
78	131,4	94,1	37,3	0,22
8	1.138,5	1.483,2	-344,7	-1,56
81	267,5	410,2	-142,7	-0,68
82	63,7	141,1	-77,4	-0,39
83	122,8	210,2	-87,4	-0,43
84	684,5	721,8	-37,3	-0,06
9	2.585,6	2.320,7	264,9	1,88
91	347,8	384,0	-36,2	-0,12
92	732,1	650,1	82,0	0,57
93	96,6	143,4	-46,9	-0,22
94	187,2	160,7	26,4	0,17
95	473,7	359,9	113,7	0,68
96	87,7	128,2	-40,5	-0,19
97	375,9	197,0	178,8	1,00
98	284,8	297,3	-12,5	-0,01
No consta	615,9	.	.	.

En general, aunque la comparativa es bastante consistente, se observa una diferencia a la baja en la ocupación del Censo respecto a la recogida en la EPA.

Existe un exceso de ocupaciones elementales, especialmente en el código de 'Peones de la industria manufacturera', que contrasta con un déficit de ocupados en 'Operadores de instalaciones y maquinaria, y montadores'.

Otras ocupaciones infra estimadas en el Censo son los 'Técnicos y profesionales científicos e intelectuales' y los 'Técnicos; profesionales de apoyo'. Entre ellos, los códigos más afectados son los de 'Profesionales de las ciencias físicas, químicas, matemáticas y de las ingenierías', 'Especialistas en organización de la Administración Pública y de las empresas y en la comercialización', 'TIC' y 'Profesionales de apoyo en finanzas y matemáticas'. Se trata de ocupaciones técnicas, muy distribuidas entre las actividades económicas, difíciles por tanto de captar si no aparecen actualizadas en el registro de contratos del SEPE.

En otros códigos como en 'Directores y gerentes' y en otras ocupaciones profesionales del grupo 2 se observan menores diferencias y datos generalmente coherentes en esta comparativa.

Lugar de trabajo

Respecto al **lugar de trabajo**, hay que destacar que la diferente metodología de construcción usada en el censo actual y en el de 2011 dificulta las posibles comparaciones entre censos.

En el Censo 2021 se proporciona el municipio o, en su defecto, provincia, del establecimiento de trabajo que figura en fuentes administrativas, dejando de captarse situaciones como el teletrabajo, trabajo en el propio domicilio, el trabajo en varios municipios o trabajo en el extranjero por las que sí se preguntaba en 2011.

Por otra parte, la comparación entre censos también se dificulta debido a la forma de formular la pregunta en el censo de 2011, dado que en él, si un trabajador afirmaba trabajar en varios municipios no debía responder en qué provincia lo hacía, luego la provincia de trabajo solo se asignaba a los que lo hacían en un municipio fijo.

Con estas salvedades, podemos observar que en general aumenta significativamente respecto a 2011 el porcentaje de ocupados que trabajan en un municipio, provincia o CC.AA distinta de la de residencia.

Figura 29. Comparativa con el Censo 2011 del lugar de trabajo (miles)

Relación entre lugar de residencia y trabajo	Censo 2011		Censo 2021		Diferencia	Diferencia %
	Frecuencia	Porcentaje	Frecuencia	Porcentaje		
No consta	1.595	9,1	381	2,0	-1.214	-7,1
Mismo municipio	9.946	56,8	6.052	32,4	-3.894	-24,4
Distinto municipio, misma provincia	5.300	30,3	10.750	57,5	5.449	27,3
Distinta provincia, misma CCAA	290	1,7	530	2,8	240	1,2
Distinta CCAA	382	2,2	967	5,2	584	3,0

6 Composición de hogares / relaciones de parentesco

a) Requisitos

- Reglamento UE

Se establecen varias variables obligatorias relacionadas con la composición del hogar y las relaciones de parentesco entre sus miembros. Destacamos las dos más importantes:

Tipo de hogar (hogares privados)

Las categorías obligatorias para esta variable son:

- Hogar no familiar
- Hogar unifamiliar
 - a. Hogares formados por parejas
 - i. Pareja sin hijos que convivan en el hogar
 - ii. Pareja con hijos que conviven en el hogar alguno menor de 25 años
 - iii. Pareja con hijos que conviven en el hogar todos de 25 años o más
 - b. Hogares monoparentales
 - i. Padre que convive con algún hijo menor de 25 años
 - ii. Padre que convive con todos sus hijos de 25 años o más
 - iii. Madre que convive con algún hijo menor de 25 años
 - iv. Madre que convive con todos sus hijos de 25 años o más
 - c. Hogar con dos familias o más

El Reglamento exige publicar esta desagregación como máximo a nivel nacional.

Tipo de núcleo familiar

- Pareja casada o en unión registrada

- Pareja casada o en unión registrada sin niños residentes
 - Distinto sexo
 - Mismo sexo
- Pareja casada o en unión registrada con al menos un niño residente menor de 25 años
 - Distinto sexo
 - Mismo sexo
- Pareja casada o en unión registrada cuyo hijo o hija residente menor tiene 25 años o más
 - Distinto sexo
 - Mismo sexo
- Pareja en unión libre
 - Parejas en unión libre sin niños residentes
 - Parejas en unión libre con al menos un niño residente menor de 25 años
 - Parejas en unión libre cuyo hijo o hija residente menor tiene 25 años o más
- Familias monoparentales (solo un padre varón)
 - con al menos un niño residente menor de 25 años
 - cuyo hijo o hija residente menor tiene 25 años o más
- Familias monoparentales (solo madre)
 - con al menos un niño residente menor de 25 años
 - cuyo hijo o hija residente menor tiene 25 años o más

A niveles inferiores al nacional no es necesario desagregar en parejas del mismo o distinto sexo.

– Censos anteriores

En el anterior censo de 2011 las variables relativas a la composición de los hogares no se preguntaron directamente en el cuestionario, sino que se construyeron a partir de la siguiente pregunta:

6 ¿Viven en esta vivienda los siguientes familiares de la Persona 1?

*Cuando marque SI, escriba el número con el que este familiar aparece en la **Lista de personas** de la página 2:*

Padre	<input type="checkbox"/> NO	<input type="checkbox"/> SÍ	→	Persona nº	<input type="text"/>
Madre	<input type="checkbox"/> NO	<input type="checkbox"/> SÍ	→	Persona nº	<input type="text"/>
Cónyuge o pareja	<input type="checkbox"/> NO	<input type="checkbox"/> SÍ	→	Persona nº	<input type="text"/>
Otros familiares (hijos, hermanos, etc.)	<input type="checkbox"/> NO	<input type="checkbox"/> SÍ	→	Nº de persona de uno de ellos	<input type="text"/>

No obstante, en la publicación del Censo 2011 se proporcionan tablas sobre hogares utilizando varias clasificaciones, dando lugar a un número de categorías superior al exigido por el Reglamento. Cabe destacar:

Estructura del hogar

Censos de Población y Viviendas 2011. Hogares

Resultados nacionales

Hogares según la estructura del hogar (detallada)

Unidades: hogares

Hogar con una mujer sola menor de 65 años	1.054.513
Hogar con un hombre solo menor de 65 años	1.429.621
Hogar con una mujer sola de 65 años o más	1.279.486
Hogar con un hombre solo de 65 años o más	429.700
Hogar con padre que convive con algún hijo menor de 25 años	180.826
Hogar con madre que convive con algún hijo menor de 25 años	693.167
Hogar con padre que convive con todos sus hijos de 25 años o más	153.055
Hogar con madre que convive con todos sus hijos de 25 años o más	666.209
Hogar formado por pareja sin hijos	3.804.677
Hogar formado por pareja con 1 hijo menor de 25 años	2.023.841
Hogar formado por pareja con 2 hijos en donde algún hijo es menor de 25 años	2.552.052
Hogar formado por pareja con 3 ó más hijos en donde algún hijo es menor de 25 años	538.471
Hogar formado por pareja con 1 hijo de 25 años o más	919.634
Hogar formado por pareja con 2 hijos de 25 años o más	252.663
Hogar formado por pareja con 3 ó más hijos de 25 años o más	35.262
Hogar formado por pareja o padre/madre que convive con algún hijo menor de 25 años y otra(s) persona(a)	894.956
Otro tipo de hogar	1.175.560

Fuente: Instituto Nacional de Estadística

Tipo de hogar

- Hogares unipersonales
- Hogares multipersonales que no forman familia
- Hogares de una familia sin otras personas adicionales y ningún núcleo
- Hogares de una familia sin otras personas adicionales y sólo un núcleo
- Hogares de una familia sin otras personas adicionales y un núcleo y otras personas emparentadas
- Hogares de una familia sin otras personas adicionales y dos núcleos o más
- Hogares de una familia con otras personas adicionales
- Hogares de dos o más familias

En el caso de los núcleos se ofrece información como el tipo de núcleo (pareja con o sin hijos, padre con hijos, madre con hijos) y características de la pareja (de hecho/ de derecho, del mismo /distinto sexo según número de hijos):

b) Planteamiento propuesto para la variable y desglose

A la vista de la información administrativa disponible, las categorías que se van a establecer para las variables relativas a la composición del hogar y las relaciones de parentesco entre sus miembros en el Censo de 2021 coinciden con las exigidas en el Reglamento

c) Método y fuentes

Para poder determinar la composición de los hogares es necesario reconstruir las relaciones de parentesco entre sus miembros a partir de la información contenida en los registros administrativos disponibles.

Una primera fase consiste en agrupar las personas del FPC en hogares a efectos censales. La definición de hogar que se aplicará en el Censo 2021, al igual que se hizo en el 2011, es la de hogar-vivienda, es decir, todas las personas que viven en una misma vivienda son miembros del mismo hogar, por lo que existirá un hogar por cada vivienda ocupada (vivienda principal en terminología censal) o cada alojamiento encontrado.

En un principio el punto de partida es el Padrón, que indica, para cada persona, cuál es su hogar padronal. Este hogar padronal viene dado por las personas empadronadas en una misma dirección o, cuando ésta no está suficientemente determinada, ya que en un reducido número de casos puede faltar información como el código de vía o el número, por las personas inscritas en la misma hoja padronal.

Sin embargo existen casos concretos en los que se procede a modificar el hogar padronal inicial de la persona:

Menores empadronados en un domicilio sin ningún adulto (“menores solos”)

Se considera que por definición esta situación no es factible y es consecuencia de una incidencia temporal en la situación administrativa del menor. Tras descartar, mediante el cruce con los directorios de establecimientos colectivos disponibles, que se trate de un centro o vivienda tutelada de menores, se procede a buscar en Padrón a los progenitores del menor. Para ello, como se detallará más adelante, se utiliza la información administrativa disponible (boletines de parto, información tributaria, información del Registro Central de Extranjeros) o se realizan búsquedas por coincidencia de apellidos. En muchas ocasiones se comprueba que esta situación responde a errores en la normalización de las direcciones padronales. Se procede a asignar al menor el hogar en el que está/n empadronado/s sus progenitor/es.

Existe un residuo de casos para los que no se encuentra a ningún progenitor y en los que es necesario agrupar al menor con otros parientes (hermanos, abuelos).

Hogares que, tras contrastar con fuentes tributarias, se unen o se separan

Dado que se dispone del fichero precensal de viviendas y que esta ha de ser totalmente coherente con la información de las viviendas del fichero precensal de personas, la información de hogares padronales se contrasta y, en caso necesario, se corrige, para formar los hogares censales.

Como se detallará más adelante, cada vivienda principal del fichero precensal de viviendas va a tener asociado un identificador. Este identificador va a proceder, mayoritariamente, de las referencias catastrales de las viviendas declaradas por cada persona como vivienda habitual en el IRPF, siempre que estén situadas en la misma sección censal que la dirección de empadronamiento.

Así, cabe esperar que cada hogar padronal se corresponda con una única referencia catastral. Pero habrá casos en que un mismo hogar padronal apunte a dos referencias catastrales diferentes, porque las personas empadronadas en esa dirección declaren diferentes viviendas habituales, y casos en que dos o más hogares padronales se correspondan con una sola referencia catastral porque las personas empadronadas en ellos declaren tener la misma vivienda habitual.

En el primer caso el hogar padronal inicial se dividirá en dos (o varios) hogares censales. En el segundo caso, dos hogares padronales se unirán en un único hogar censal.

Una vez determinados los hogares censales, se procede a buscar para cada persona, dentro de su hogar censal, al padre, madre, cónyuge y otros posibles parientes, lo que nos va permitir generar, al igual que en censos anteriores, las familias, núcleos familiares y estructura de los hogares.

Búsqueda de progenitores:

Puesto que el padre o madre de una persona no cambia a lo largo del tiempo, se pueden utilizar la información sobre padres e hijos de los censos pasados.

Además se dispone de los Boletines de Partos del Movimiento Natural de Población, disponibles desde 2002, en los que figuran datos identificativos del padre y madre de cada nacido, lo que permite asignar a cada nacido el padre y madre que figuran en estos boletines.

También se explota información del Registro Central de Extranjeros, gestionado por la D.G. Policía, sobre Autorizaciones de residencia a extranjeros por motivos familiares (reagrupación familiar). Tenemos información de la persona que otorga el derecho de residir en España y también de los familiares que se benefician de este derecho y su parentesco con la persona de referencia.

Se dispone así de una **base de datos de madres-hijos y padres-hijos formada por 21.998.937 y 19.960.819 pares respectivamente** para realizar una primera búsqueda dentro de cada hogar.

Así, si alguien continúa compartiendo hogar en la fecha de referencia con alguno de los progenitores que figuraban en censos anteriores y cumple unas determinadas condiciones (sexo, diferencia de edad) se considera que ese es su padre o madre.

Para las personas a las que no hemos asignado un padre en el paso anterior buscamos un candidato a padre entre los registros del hogar para los que coincide el nombre del padre y comparten el primer apellido. La información sobre el nombre del padre es proporcionada por la D.G. Policía a partir del fichero del DNI (o NIE en el caso de los extranjeros) o bien recuperada de los Boletines de Partos.

Por último, si en los pasos previos para una persona no se ha localizado a su padre, se puede obtener dicha información a partir de un hermano, con el que comparta apellidos, que sí lo tenga asignado.

El algoritmo de búsqueda de la madre es similar al del padre, modificando algunas condiciones.

Búsqueda de cónyuge o pareja:

Un primer paso consiste en declarar cónyuges o parejas a aquellas personas que comparten vivienda y tienen un hijo en común.

El siguiente paso es recuperar las parejas registradas en los Censos de 2001 y 2011, así como los matrimonios recogidos en los Boletines de Matrimonio de MNP, los localizados

en los ficheros que la Agencia Tributaria y las Haciendas Forales y procedentes del Registro Central de Extranjeros.

Se dispone, por tanto, de una **base de datos, que se incrementa cada año, con 15.695.060 de potenciales parejas** formada por personas que se han encontrado como pareja en alguno de estos registros en un momento dado.

Estas parejas potenciales se dan por válidas en el caso de que continúen compartiendo vivienda en la base de datos padronal.

En los pocos casos en que una persona figure compartiendo domicilio con más de un posible cónyuge se establecen prioridades en función de la actualidad de la fuente y las diferencias de edad.

Tras estos pasos, comparando con los resultados ofrecidos por las encuestas a hogares, observamos que existen parejas no casadas que conviven y que no tienen hijos en común que no se van a encontrar en ningún registro.

Por ello, para la detección de nuevos cónyuges, es necesario establecer un **modelo de imputación** que permita localizar parejas en hogares en los que inicialmente no se tienen.

En los hogares de dos miembros el objetivo es detectar parejas sin hijos en los hogares clasificados inicialmente como “otro tipo”, por ejemplo formados por dos personas adultas compartiendo piso. Para hogares de tres miembros, el objetivo es detectar nuevas parejas en hogares clasificados como “padre o madre con hijo y otra persona” y hogares de “otro tipo”.

Para ello se ha trabajado tanto con la muestra acumulada de dos años de la Encuesta Continua de Hogares de 2020 como con la muestra de la ECEPOV, diferenciando parejas de diferente sexo y parejas del mismo sexo, femenino y masculino.

Con estos datos se ha estudiado qué características diferenciales (edad, diferencia de edad, nacionalidad, país de nacimiento, estado civil, régimen de tenencia de la vivienda, nivel educativo, años compartiendo vivienda,...) tienen las parejas frente, por ejemplo, a dos personas compartiendo piso. En un principio se ha tratado de ajustar modelos logit y probit para explicar la probabilidad de que dos personas sean pareja en función de estas variables, pero al no cumplirse las hipótesis estadísticas necesarias se ha optado por utilizar métodos no paramétricos, como los árboles de decisión. Las conclusiones son muy similares a las obtenidas por otros países, como Holanda, Bélgica, Austria, Noruega, Suecia o Finlandia, que tienen el mismo problema en la determinación de parejas a partir de registros.

Se ha concluido que para hogares de dos miembros un árbol de decisión sencillo, basado en la edad del menor miembro, la diferencia de edad y la combinación de estados civiles nos ofrece una buena detección de parejas, alcanzando una sensibilidad y especificidad superiores al 98%.

Búsqueda de otros parientes:

Los hermanos se localizan porque comparten padre o madre (o ambos) o porque presentan sus dos apellidos iguales. En el caso de nacionalidad extranjera, se exige un único apellido informado y coincidente.

Los hijos se detectan a partir de un algoritmo de reciprocidad: si la persona A es la madre de la persona B, entonces B es “otro pariente” (hijo) de A.

La misma técnica se utiliza para localizar abuelos: si A es padre de B y B es padre de C, entonces A es otro pariente (abuelo) de C y C es otro pariente (nieto) de A.

En general, una vez localizados padres y cónyuges, cualquier parentesco puede obtenerse aplicando estas reglas.

d) Análisis de la calidad de la variable

El número de registros según la fuente ganadora utilizada para decidir si la persona convive o no con su padre o madre se distribuye así:

Fuente ganadora	Personas conviven con su padre	%	Personas conviven con su madre	%
Total	11.264.258	100,0%	14.487.811	100,0%
Censos anteriores, MNP y RCE	9.128.104	81,0%	11.479.744	79,2%
Nombre de la madre (DGP)	1.776.141	15,8%	2.556.962	17,6%
Hermanos	121.795	1,1%	122.284	0,8%
Imputado	238.218	2,1%	328.821	2,3%

El número de registros según la fuente ganadora utilizada para decidir si la persona convive o no con su pareja se distribuye así:

Fuente ganadora	Personas conviven con su pareja	%
Total	20.119.362	100,0%
Hijo/s en común	12.606.196	62,7%
Censos pasados, MNP, AEAT y RCE	5.914.584	29,4%
Imputado	1.598.582	7,9%

Se observa que a partir de los censos anteriores y la información del nombre de padre y madre se consigue localizar casi totalidad de progenitores que conviven con sus hijos.

Para la detección de parejas, sin embargo, es necesario imputar un 7,9% de los registros a partir de un modelo, ya que, como ya se ha comentado, esta es la principal dificultad para determinar la composición de los hogares a partir de registros administrativos, tanto en España como en el resto de países.

La imputación se concentra en las parejas del mismo sexo y en las parejas de hecho, de las que se tiene menos información en los registros administrativos:

La información metodológica sobre los hogares se completará próximamente. Para más información sobre estas variables se remite al lector al proyecto técnico del censo de 2021¹ donde se describe exhaustivamente el alcance de la investigación.

7 Censo de viviendas

7.1 ASPECTOS GENERALES DEL CENSO DE VIVIENDAS

El censo de viviendas consiste en una cuantificación exhaustiva de todas las viviendas, teniendo en cuenta que, a los efectos de un censo de población, se entiende por **vivienda** todo recinto estructuralmente separado e independiente que, por la forma en que fue construido, reconstruido, transformado o adaptado, está concebido para ser habitado por personas o, aunque no fuese así, constituye la residencia habitual de alguien en el momento censal.

Tradicionalmente, en los censos anteriores basados en la recogida directa de información de las viviendas a partir del recorrido de los agentes censales, no se consideran viviendas los huecos que, a pesar de estar concebidos inicialmente para habitación humana e incluidos dentro de edificios principalmente residenciales, en el momento censal se constataba que estaban dedicados totalmente a otros fines (por ejemplo, los que estén siendo usados exclusivamente como locales comerciales, consultas médicas, despachos de abogados, ...).

Este un elemento de divergencia importante con el censo actual en el que, simplificando el proceso, podemos decir que se contabilizan como viviendas las registradas en Catastro como uso residencial y aquellas que, estando clasificadas con un uso catastral distinto del residencial (oficinas, hoteles, locales), se comprueba que constituyen la residencia de un hogar familiar en el momento censal.

Por tanto, a diferencia del censo anterior, un inmueble utilizado como consulta médica ubicado en un edificio residencial que en Catastro figura con clave de uso residencial se va a contabilizar como vivienda en el censo 2021.

Esta diferencia conceptual va a implicar una aparente ruptura en la series de datos de viviendas entre ambos censos.

No obstante, hay que señalar que en el censo de 2011 se contabilizaron (dentro de los portales) un total de 25,849.338 “inmuebles” pero de ellos solo 25.218.536 fueron consideradas como viviendas en tanto que se usaban como tales (consultar apartado de tablas predefinidas de censo 2011 relativas a Edificios, el número de inmuebles ubicados en edificios destinados a viviendas²).

¹ Más información en <https://www.ine.es/censos2021/>

²<https://www.ine.es/jaxi/Datos.htm?path=/t20/e244/edificios/p02/I0/&file=00004.px>

Los datos de viviendas 2021 no serán por tanto comparables a las viviendas de 2011 sino más bien a los inmuebles ya que con la información catastral no puede determinarse si algo que está declarado como vivienda se está usando como local.

Esta cifra de “inmuebles”, sin reflejar tampoco exactamente el mismo concepto que el usado en 2021, es la que más se asemeja y esa comparativa daría que el crecimiento en el número de viviendas en la década se situaría en una cifra cercana a los 774.370 (=26.623.708 – 25.849.338).

Otra cuestión novedosa, de creciente importancia en algunas ciudades, es que algunos locales comerciales de pie de calle se pueden reconvertir en viviendas. En tanto que esta conversión requiere de autorización municipal y que el cambio de uso se refleja en Catastro, esa información se incorpora al censo. Aparecerán (o desaparecerán) nuevas viviendas, por tanto, en edificios terminados.

En resumen, en el censo 2021 se considera vivienda todo aquello que figure en Catastro como tal, independientemente de su uso real, que no podrá determinarse a partir de registros administrativos.

A continuación, podemos observar un pequeño esquema de los distintos tipos de viviendas usados en el Censo de 2021.

Tipos de vivienda

- Viviendas familiares
 - Viviendas convencionales.
 - Viviendas principales.
 - Viviendas no principales.
 - Alojamientos.
- Establecimientos colectivos.

Vivienda familiar: es aquella destinada a ser habitada por una o varias personas, no necesariamente unidas por parentesco, y que no constituyen un establecimiento colectivo.

Vivienda familiar convencional: Es una vivienda familiar que cumple todos los requisitos para ser habitada y en la fecha censal no se utiliza totalmente para otros fines. La vivienda convencional puede ser **principal** cuando es la residencia habitual de sus componentes. Si está destinada a ser ocupada sólo ocasionalmente o cuándo permanece sin ser ocupada se denomina **no principal**.

Alojamiento: Vivienda familiar que presenta la particularidad de ser móvil, semipermanente o improvisada, o bien que no ha sido concebida en un principio con fines residenciales, pero, sin embargo, constituye la residencia de una o varias personas en el momento del censo (por tanto, los alojamientos vacíos no se censan).

En el censo 2021 fundamentalmente se han clasificado como alojamientos asentamientos de chabolas o “invernaderos”, casas prefabricadas, zonas de acampada permanente de caravanas,...en las que se ha encontrado población empadronada.

Las casa cuevas, que en censos anteriores se clasificaron como alojamientos, y en general cualquier inmueble registrado en Catastro con uso residencial (o con otro uso pero con un hogar familiar residiendo en él) en se ha considerado vivienda.

Establecimiento colectivo (o vivienda colectiva): Es un lugar donde reside un colectivo de población, es decir, un grupo de personas sometidas a una autoridad o un régimen común no basados en lazos familiares ni de convivencia. La vivienda colectiva puede ocupar sólo parcialmente un edificio o, más frecuentemente, la totalidad del mismo. Por ejemplo, son viviendas colectivas las prisiones, las residencias de personas mayores o los conventos; también lo son los pisos tutelados.

Este punto de la metodología está centrado en las viviendas familiares convencionales.

7.2 PROCESO DE OBTENCIÓN DEL DIRECTORIO DE VIVIENDAS

La fuente administrativa a partir de la cual se elabora el Censo de Viviendas es fundamentalmente el Catastro. Para ser precisos hay que hablar de “los catastros” dado que los territorios con régimen fiscal foral tienen sus propios catastros, con distintas denominaciones por lo que el FCV se construye a partir del Catastro del territorio fiscal común y los cuatro catastros forales (las tres provincias vascas y Navarra).

Las principales ventajas de usar los catastros consisten en que, por un lado, todas las viviendas tienen un identificador único, la referencia catastral (RC), aunque esta afirmación tiene matices, que se describen más adelante.

Además, para todas las viviendas se dispone, en principio, de ubicación geográfica precisa (coordenadas), está bastante actualizado y debería estar libre de errores en gran medida, al tener repercusiones tributarias.

La información recibida por el Instituto Nacional de Estadística por parte de la Dirección General de Catastro es el “Fichero Informático de Remisión de Alteraciones de Catastro” (Datos referidos al territorio fiscal común, excluyendo País Vasco y Navarra).

Los directorios de viviendas de País Vasco y Navarra se han conformado con la colaboración de sus institutos de estadística (EUSTAT y NASTAT). Ambos han remitido al INE información tanto de las viviendas como variables asociadas a ellas. Esta información se ha cruzado con la información de hogares y personas del censo. En determinados casos, se obtiene o modifica la sección censal asociada a la vivienda y se depuran las coordenadas geográficas. Además, al igual que en territorio fiscal común, se crea un identificador único que se asocia a los hogares del territorio, por lo que viviendas y hogares quedan perfectamente enlazados.

Este fichero Incluye varios tipos de registros. Los tipos utilizados para la obtención del directorio de viviendas son los siguientes:

- Tipo 11: Registro de finca. Cada registro hace referencia a una finca o parcela catastral, que es porción de terreno delimitada en la que se desarrollan los bienes

inmuebles y construcciones asociados a la misma. Incluye la referencia catastral de la parcela a catorce posiciones. Este tipo de registro contiene otro tipo de información importante que ya se ha mencionado: Las coordenadas geográficas. Los registros se identifican a partir de la referencia catastral a catorce posiciones (RC14).

- Tipo 13: Registro de unidad constructiva. Representa un edificio o un conjunto de construcciones particularizadas dentro de un edificio. Contiene el año de construcción. Además de incluir la referencia catastral a catorce posiciones, incluye el código de la unidad constructiva que sirve para identificar las distintas unidades constructivas
- Tipo 14: Registro de construcción. Identifica cada uno de los locales existentes en un bien inmueble, con su descripción física: superficie, antigüedad, tipología, destino. Incluye la referencia catastral (RC o Rc20) y un número de orden del elemento constructivo (num-o) para distinguir las distintas construcciones asociadas a un mismo bien inmueble.
- Tipo 15: Registro de inmueble. Identifica mediante la referencia catastral a veinte posiciones cada uno de los bienes inmuebles dentro de una parcela catastral. Cada bien inmueble (concepto jurídico) incluye una o varias construcciones (concepto físico). Podríamos pensar, a priori, que cada bien inmueble con tipología vivienda es una vivienda en sí, pero no siempre ocurre en la realidad. Por ejemplo, hay una serie de bloques de viviendas que aparecen como un único bien inmueble en catastro, pero en los que hay más de una vivienda, es lo que se conoce como falta de división horizontal.
- Tipo 46: Registro de situaciones Finales de Titularidad. Contiene datos identificativos de bienes inmuebles, referencia catastral a veinte posiciones, junto con datos identificativos de sus correspondientes titulares catastrales. Además, incluye datos del derecho del titular sobre el bien.

La principal dificultad para elaborar un censo de viviendas a partir de este conjunto de ficheros con información catastral es pasar de “bienes inmuebles de uso residencial” (concepto jurídico de Catastro) a “viviendas” (concepto físico Censo).

El Catastro contiene los denominados “bienes inmuebles”, rústico y urbanos, con sus características físicas, económicas y jurídicas, como la localización y la referencia catastral, la superficie, el uso o destino, el año de construcción, la calidad de las construcciones, o información sobre los propietarios (con su número de identificación fiscal).

Es decir, el Catastro no contabiliza “viviendas” sino “bienes inmuebles de uso residencial”. Parece un matiz menor, pero no son exactamente lo mismo dado que, como se explica más adelante, el bien inmueble es más bien una unidad jurídica mientras que la vivienda a efectos censales tiene una definición de unidad física. Para llegar al concepto de “vivienda” usado en el censo a partir de las unidades catastrales, se han realizado diversos procesos, para lo que ha sido necesaria una estrecha colaboración entre el INE y la DG del Catastro con el fin de interpretar mejor las complejas estructuras de los ficheros catastrales

Variables de Catastro utilizadas para construir el directorio de viviendas:

- Clave de uso. La clave de uso clasifica el uso del bien inmueble: vivienda, comercial, almacén, etc.
- Tipologías constructivas (tip_con). La tipología constructiva informa sobre el tipo de construcción de acuerdo con sus características arquitectónicas y de funcionalidad de uso.
- Código de destino de la construcción. El código de destino indica la clasificación del destino de la construcción.
- Identificadores de todos los tipos de registro.

La clave de uso puede tomar los siguientes valores:

Figura 30. Clave de uso de bien inmueble

Código	Uso
A	Almacén, estacionamiento
C	Comercio
E	Cultural
G	Turismo
I	Industrial
K	Deportivo
M	Solares sin edificar
O	Oficina
P	Publico
R	Religioso
T	Espectáculos
V	Residencial
Y	Sanidad

En cuanto a los códigos de destino de la construcción, la clasificación es:

Figura 31. Códigos de destino de las construcciones.

Código de destino	
AXX	Almacenamiento
BXX	Almacenamiento
CXX	Comercio
DXX	Comercio
EXX	Enseñanza
FXX	Enseñanza
GXX	Hotelero
HXX	Hotelero

IXX	Industrial
JXX	Industrial
KXX	Deportivo
LXX	Deportivo
M	Suelo sin edificar
N	Suelo sin edificar
OXX	Oficina
PXX	Público
QXX	Público
RXX	Religioso
SXX	Religioso
TXX	Espectáculos
UXX	Espectáculos
VXX	Vivienda
WXX	Vivienda
XXX	Otros usos
ZXX	Otros usos

XX indica que puede tomar distintos valores, dentro de un subconjunto perteneciente a la categoría general indicada por la primera letra. Por ejemplo, dentro de A (Almacenamiento), AAL indica almacén, AAP indica aparcamiento, AES indica estación.

Dentro de la cantidad total de tipologías constructivas incluidas en los ficheros de catastro, las utilizadas para construir el directorio de viviendas son:

Figura 32. Subconjunto de tipologías constructivas utilizadas para formar el directorio de viviendas.

1.1 VIVIENDAS COLECTIVAS de CARÁCTER URBANO	1.1.1 EDIFICACIÓN ABIERTA	
	1.1.2 EN MANZANA CERRADA	
	1.1.3	GARAJES
		TRASTEROS
LOCALES EN ESTRUCTURA		
1.2 VIV. UNIFAMILIARES de CARÁCTER URBANO	1.2.1 EDIFICACIÓN AISLADA O PAREADA	
	1.2.2 EN LÍNEA O MANZANA CERRADA	
	1.2.3	GARAJES
PORCHES EN PLANTA BAJA		
1.3 EDIFICACIÓN RURAL	1.3.1 USO EXCLUSIVO DE VIVIENDA	
	1.3.2 ANEXOS	
7.1 CON RESIDENCIA	7.1.1 HOTELES, HOSTALES, MOTELES	
	7.1.2 APARTHOTELES, BUNGALOWS	

Fuente: https://www.catastro.minhap.es/documentos/preguntas_frecuentes_formato_cat.pdf

La información aparece en los ficheros de catastro en forma de vector, de modo que nos quedamos con las cuatro primeras posiciones del vector (tip_con_4), que, por ejemplo, en el caso de viviendas colectivas de carácter urbano edificación abierta, será '0111', en el caso de unifamiliares, edificación aislada o pareada es '0121', etc.

Para formar el fichero inicial de viviendas, partimos de los ficheros de Catastro de 2021 correspondientes a los registros de bien inmueble y a los registros de construcción, es decir, tipos 14 y 15 de las versiones rústicas y urbanas.

Se unen los registros rústicos y urbanos de bienes inmuebles para añadir los registros de rústico que no sean iguales a otros que ya estén en urbano (iguales hasta el secuencial de bien inmueble dentro de la parcela, cuatro últimas posiciones de la referencia catastral, sin dígitos de control, en adelante secu-bi). Se restringe esta unión a los registros con clave de uso compatible con vivienda.

Para cada parcela de este fichero, se calcula si presenta división horizontal (**dhv**, toma el valor 1 si hay división horizontal y 0 en caso contrario), considerando que la hay cuando aparece más de un valor diferente de secuencial de bien inmueble dentro de la parcela.

Del mismo modo, se unen los registros rústicos y urbanos de construcciones. Se restringe esta unión a los registros con códigos de destino que significan vivienda (V, W y equivalentes, se normaliza el código de destino).

Se integran estos ficheros para asignar a cada bien inmueble el conjunto de construcciones del que está formado. Se elegirán determinados registros del mismo (o una agregación de registros representados por uno solo de ellos) para formar parte del directorio de viviendas.

Además, hay una parte de construcciones con tipología vivienda que cruzan con un bien inmueble cuya clave de uso es distinta de vivienda, estos se incorporan también en el caso en que se compruebe que realmente es vivienda.

Primero. Se seleccionan para formar parte del directorio todos los registros de tipo 15 que no han encontrado su pareja a través del secu-bi con los registros tipo 14.

Segundo. Para los registros de la mezcla se comprueba si son únicos (hasta secu-bi) o si son duplicados (coincidentes hasta secu-bi y que se diferencian en el num-o).

Los registros son únicos pasan directamente a formar parte del directorio de viviendas.

Se trabaja con los duplicados calculando para cada secu-bi los valores de **num_edi** (se cuentan los registros con tip_con_4 con valor '0111' o '0112' correspondientes a edificios de varias viviendas) y **num_uni** (registros con tip_con_4 con valor '0121','0122','0131','0132' correspondientes a unifamiliares). También se calcula **num_edi_3** (registros con tip_con_4 con valor '0113') y **num_uni_3** (registros con tip_con_3 con valor '012','013'). Siendo tip_con_4 y tip_con_3 las primeras cuatro o primeras tres posiciones respectivamente de la tipología constructiva.

Según las combinaciones de **num_edi** y **num_uni** se decide si un registro es una vivienda y debe formar parte del directorio o si lo que es una vivienda es la agregación de varios registros por lo que solo se incluye uno de ellos en el directorio (sumando las superficies de los que agrega). Se consideran estas situaciones:

Situación 1: Si num_edi>0 y num_uni=0

Identificamos este caso con dos situaciones: viviendas en edificios sin división horizontal o áticos y dúplex en edificios con división horizontal.

- Si **dhv=0**, se eligen todos los registros con **tip_con_4** igual a '0111' o '0112' (el número de registros elegidos coincide con **num_edi**). Esto es lo que identificamos con viviendas en edificios sin división horizontal. Aunque de esta forma no se capta toda la falta de división horizontal, estos serían en los casos en los que solamente aparece una construcción para todo el edificio o una sola construcción para cada planta.
- Si **dhv=1** y **num_edi**≤3, (estos son los que identificamos con áticos y dúplex), se elige un registro entre los registros que tengan **tip_con_4** igual a '0111' o '0112'.
- Si **dhv=1** y **num_edi** >3 se eligen todos los registros con **tip_con_4** igual a '0111' o '0112'.

Situación 2: Si num_edi =0 y 0< num_uni <=3, se elige solo uno. Identificamos este caso con unifamiliares de varias plantas. El registro elegido debe tener tipología '0121', '0122', '0131' o '0132'.

Situación 3: Si num_edi =0 y num_uni >3, se eligen los registros que se consideren de planta baja. Para ello se normaliza el campo planta. Si falta algún secu_bi sin planta baja se elige un registro de los duplicados y se selecciona como vivienda. Esto puede ocurrir en el caso de viviendas que no empiecen en la planta baja sino en la primera por tener debajo un garaje u otro local. Identificamos este caso con urbanizaciones de chalés (en los que cada planta aparece separada) o con unifamiliares con más de tres plantas o anexos.

Situación 4: Si num_edi=0 y num_uni=0 vemos para cada secu-bi cuantos registros tienen **tip_con_3** es igual a '012' o '013' (se resume en un campo **num_uni_3**) y cuantos tienen **tip_con_4** igual a '0113' (se resume en un campo **num_edi_3**):

- Si **num_uni_3**> 0, se elige uno de los registros como vivienda. Después de analizados se identifica este caso con unifamiliares de varias plantas.
- Si **num_uni_3**= 0 y **num_edi_3**>0, se consideran todos los registros. Después de analizados se identifica este caso con pisos en edificios.
- Si **num_uni_3**= 0 y **num_edi_3**=0, se elige uno de los registros para formar parte del directorio de viviendas.

Situación 5: Si num_edi >0 y num_uni >0

Son pocos registros, se hace una selección inicial, y se va depurando. La selección inicial consiste en:

Si **dhv=0**, se eligen todos los registros con **tip_con_4**= '0111', '0112', '0121', '0122', '0131', '0132'.

Si **dhv=1**, se elige un registro entre los registros que tengan **tip_con_4**= '0111', '0112', '0121', '0122', '0131', '0132'.

Tercero. Para los registros de construcciones para los que no se encuentra su pareja entre los registros de bienes inmuebles, se observa que casi el 80% tienen el secu-bi en blanco (se excluyen todos estos registros con secu-bi en blanco porque en general representan elementos comunes a todos los inmuebles de una finca o parcela pero no viviendas (accesos, locales comunitarios, espacios de tránsito...) y un 8% con valor '0001'. De los cuales:

- La gran mayoría son registros que tienen un código de parcela que aparece en los ficheros de bienes inmuebles, pero figuran con otro secu-bi diferente, muchas veces en los registros tipo 14 aparece el secu-bi en blanco. Parece un problema de actualización de los registros de construcciones. Por lo tanto, estos registros no se incluyen en el directorio de viviendas, ya están incluidos.
- Para el resto se van incluyendo las viviendas en función de las características que presenten, de forma similar a los anteriores, y entrando a un proceso de depuración manual en muchos casos. En determinadas zonas turísticas, hay apartamentos que realmente se están usando como viviendas, por lo que se incluye las tip_con_4='0712' en estos casos (Siempre que se pruebe que realmente se están usando como vivienda).

7.3 ASIGNACIÓN DE SECCIÓN CENSAL A CADA UNA DE LAS VIVIENDAS

El proceso de asignación censal a cada una de las viviendas se ha llevado a cabo en distintas fases:

- **Intersección geográfica.** Intersección de las capas de parcela urbana y rústica de Catastro, con la capa de seccionado censal a fecha 01/01/2021. Solo en este paso se consiguen el 99,46% de la asignación de sección censal a la vivienda. Los siguientes métodos que se describen a continuación se han realizado para conseguir aumentar ese porcentaje, quedando, después de aplicarlos, solamente 2471 viviendas (0,009%) sin sección censal, que se han depurado/asignado a alguna de las secciones censales correspondientes al municipio en el que se encuentran.
- **Asignación por coordenadas.** Intersección de la capa de puntos creada a partir de las coordenadas de las viviendas, obtenidas a partir de los registros tipo 11 con el seccionado censal a fecha 01/01/2021. Cada punto representa una parcela. El resultado de la intersección nos da, para cada parcela, la sección o secciones en las que se encuentra. En esta fase, la siguiente más numerosa en cuanto a asignación de seccionado se refiere, se consigue un 0,28% de viviendas sobre el total.
- **Municipios de una sola sección censal.** Las viviendas pertenecientes a municipios en los que sólo hay una sección censal se han asignado directamente a esa sección.
- **Intersección viales de DGC con seccionado.** Intersección de las capas de viales de Catastro, con seccionado censal a fecha 01/01/2021. El resultado de la intersección nos da para cada vial de Catastro, identificado por el código de vía, la sección o secciones en las que se encuentra. Se calculan las frecuencias para seleccionar aquellas vías que se encuentran en una sola sección. Se asigna a partir del código de vía de la vivienda.
- **Intersección viales de Cartociudad con seccionado.** Intersección de las capas de viales de Cartociudad, con seccionado censal a fecha 01/01/2021. El resultado de la

intersección nos da para cada vial de Cartociudad, identificado por el código de vía, la sección o secciones en las que se encuentra. Se calculan las frecuencias para seleccionar aquellas vías que se encuentran en una sola sección. Se realizan dos cruces. Uno por código de vía, otro por literal de vía, siempre exigiendo coincidencia de municipio.

- **Intersección apps de MDG.** Intersección de la capa de APPS (aproximaciones postales) del MDG (Marco de Direcciones georreferenciadas del INE), con la capa de seccionado censal a fecha 01/01/2021. El resultado de la intersección nos da para cada APP, y por tanto para su código y número de vía, la sección o secciones en las que se encuentra. Se calculan las frecuencias para descartar las APPS que se encuentran en más de una sección, también se descartan los casos en los que un código de Catastro se corresponda con más de un código de vía INE.
- **Asignación a partir del callejero del Censo Electoral.** Este proceso es similar al punto anterior. Se busca coincidencia a partir de la equivalencia de códigos de vía de INE y Catastro y número, asignando la correspondiente sección censal según el fichero de tramos de vía del callejero de Censo Electoral.

7.4 ENLACE PERSONAS Y VIVIENDAS

Aunque a continuación trataremos el tema del enlace de personas y viviendas, y posteriormente los problemas encontrados y como se han solucionado a la hora de confeccionar el directorio de viviendas, cabe destacar que no ha habido una linealidad en los trabajos, se han trabajado en paralelo, de forma iterativa, hasta converger en el producto censal que disponemos ahora mismo.

El objetivo es enlazar viviendas del FCV (Fichero Censal de Viviendas) con las correspondientes viviendas principales del fichero de personas, el FCP (Fichero Censal de Personas) Dicho de otra forma, se trata de asignar, a cada hogar (vivienda principal) del FCP cuál es la referencia catastral de esa vivienda.

Al hacer esto, estamos ubicando perfectamente en el territorio a toda la población porque sabremos las coordenadas geográficas de la parcela catastral en la que se ubica cada vivienda, y además, podremos relacionar las variables de las personas del FCP con las variables de las viviendas del FCV.

Las fuentes principales son el modelo 100 de declaración de la renta de la Administración Estatal de la Agencia Tributaria, y los registros tipo 46 de Catastro.

Para este proceso se ha ideado un algoritmo en varias fases:

- Fase 1. Se toma la información de la referencia catastral (RC) del modelo 100 de la AEAT. Todos los hogares que hagan declaración de la renta (aproximadamente 2/3 de los hogares de España) deben consignar en el apartado de vivienda habitual la RC de la misma. La dirección que una persona declara a efectos fiscales puede ser distinta de la dirección donde las personas figuran empadronadas. Por tanto, solo en aquellos casos en que ambas direcciones se ubiquen en la misma sección censal se asignará al hogar del fichero FCP (fichero censal de personas) la RC de la vivienda según el FCV (Fichero censal de viviendas). Esto permite que a nivel nacional (salvo País Vasco y Navarra), un 50,3% de hogares tengan asignada una RC. Además, con

el FCV distinguiremos qué referencias catastrales se refieren a priori a viviendas y cuáles no.

Podemos, por lo tanto, dividir los hogares del FCP en tres grupos: los que se han conseguido enlazar con una RC considerada vivienda, los que tienen alguna RC pero ninguna de ellas es considerada vivienda (grupo 2) en los ficheros de catastro y los hogares para las que no se dispone de ninguna RC en el fichero de la AEAT (grupo 3).

- Fase 2. Para las viviendas que no se enlazaron (grupo 3 de la fase 1) se realiza el cruce del FCP con los registros tipo 46 de titulares de catastro. Si una persona se encuentra entre los registros tipo 46 como propietario de una vivienda en la misma dirección que tiene en el FCP se asigna la RC20 a la vivienda. Para las que no se consiga asignar pasan a la siguiente etapa. Esto permite enlazar otro 12,4% a nivel nacional.
- Fase 3. Cruce por correspondencia a nivel de edificio entre FCV y FCP. Si para un edificio, identificado por calle y número, ya hay viviendas a las que se ha conseguido asignar RC por los pasos anteriores, indirectamente se ha logrado enlazar el propio edificio y se puede asignar la RC14 a las restantes viviendas principales que, según el FCP, se sitúan en el edificio (comparten la misma dirección postal hasta el nivel de escalera). También se consigue asignar RC20 para algunas viviendas en este paso, cuando se consigue cruzar por la dirección postal completa. Esto proporciona otro 17,3% de enlace a nivel nacional. Para las que no se consiga asignar se pasa a la siguiente etapa.
- Fase 4. Cruce a nivel de edificio del FCV con el MDG, por calle y número. Se busca la calle en fichero de vías del MDG y se localiza la calle y número. Si para un edificio ya hay viviendas que tienen RC14 en el MDG se puede asignar esa RC14 a las restantes viviendas del edificio. Esto proporciona otro 15,4%. Para las que no se consiga asignar se pasa a la siguiente etapa.
- Fase 5. Para todos los hogares que llegados a este punto siguen sin tener su referencia catastral, ya sea a 14 posiciones o 20, se establecen cruces por literales de vías a partir de la dirección postal donde están empadronadas las personas y la dirección postal de catastro. Los cruces establecidos son los siguientes:
 - Coincidencia exacta de literal de vía y número.
 - Utilizando las equivalencias detectadas entre los códigos de vía de Catastro e INE.
 - Usar el nombre del paraje de Catastro como nombre de vía alternativo cuando el literal de vía es “diseminado” o “proc. Simplificado”.
 - Cuando el número es “0” comprobamos si el bloque puede incluir lo que realmente es el número, en caso de coincidencia, asignamos.
 - En determinados casos, el número de uno de los dos ficheros no existe, por ejemplo, un 3 que por error han puesto como 13. Estos casos se depuran y se asignan en huecos disponibles de la misma vía.

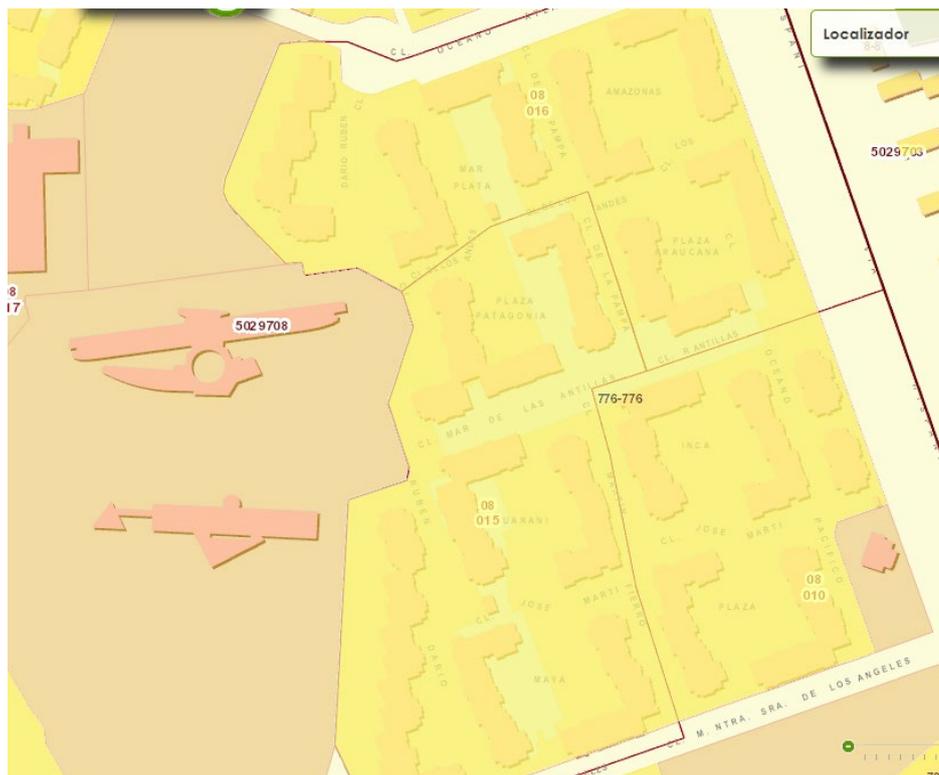
Con todo esto, hay un 1,8% de hogares de todo el territorio nacional a los que no se ha conseguido asignar por los métodos anteriores una Referencia Catastral. En estos hogares se ha estimado la vivienda a la que pertenecen dentro de la sección censal.

7.5 PROBLEMAS DETECTADOS EN LA ELABORACIÓN DEL FICHERO CENSAL DE VIVIENDAS

Aunque a continuación trataremos el tema del enlace de personas y viviendas, y posteriormente los problemas encontrados y como se han solucionado a la hora de confeccionar el directorio de viviendas, cabe destacar que no ha habido una linealidad en los trabajos, se han trabajado en paralelo, de forma iterativa, hasta converger en el producto censal que disponemos ahora mismo.

- Parcela que se extiende a lo largo de varias secciones. En ocasiones una parcela es tan grande que se extiende a lo largo de varias secciones. Estas parcelas se dividen en tantos “trozos” como secciones abarca teniendo en cuenta la información padronal de las personas asociadas a dicha parcela.

Figura 33. Detalle de parcela que se extiende por varias secciones (IDE-INE).



- Incongruencias entre la cartografía de las secciones censales y ficheros alfanuméricos de seccionado. (Direcciones de padrón cuya dirección postal tiene otra sección en la cartografía). Se asocian los hogares a la sección en la que se encuentra la cartografía.
- Falta de división horizontal dentro de una parcela catastral. El análisis realizado para el Censo 2021 nos ha permitido observar que, en algunos casos, un edificio completo tiene varias viviendas principales pero todas ellas comparten la misma

RC. Para poder asignar un identificador único se ha ideado un nuevo identificador, usado para este proyecto, que se ha denominado RC24 (se explicará de forma detallada en el siguiente apartado). Este identificador consiste en la referencia catastral, es decir, los primeros 20 dígitos coinciden con ella, y se añaden otros 4 para poder distinguir viviendas que comparten la RC. Así pues, al igual que cada persona se identifica internamente por su NIF o pasaporte, cada vivienda internamente, a efectos del Censo 2021, se identifica por su RC24.

En la siguiente figura se observa una manzana de Madrid. Los portales de la izquierda contienen del orden de 50 viviendas cada uno; los de la derecha, similares, solo tienen una única RC para cada una de las viviendas contenidas en ellos: no hay “división horizontal”.

Figura 34. Detalle de una sección censal con parcelas sin división horizontal (IDE-INE)

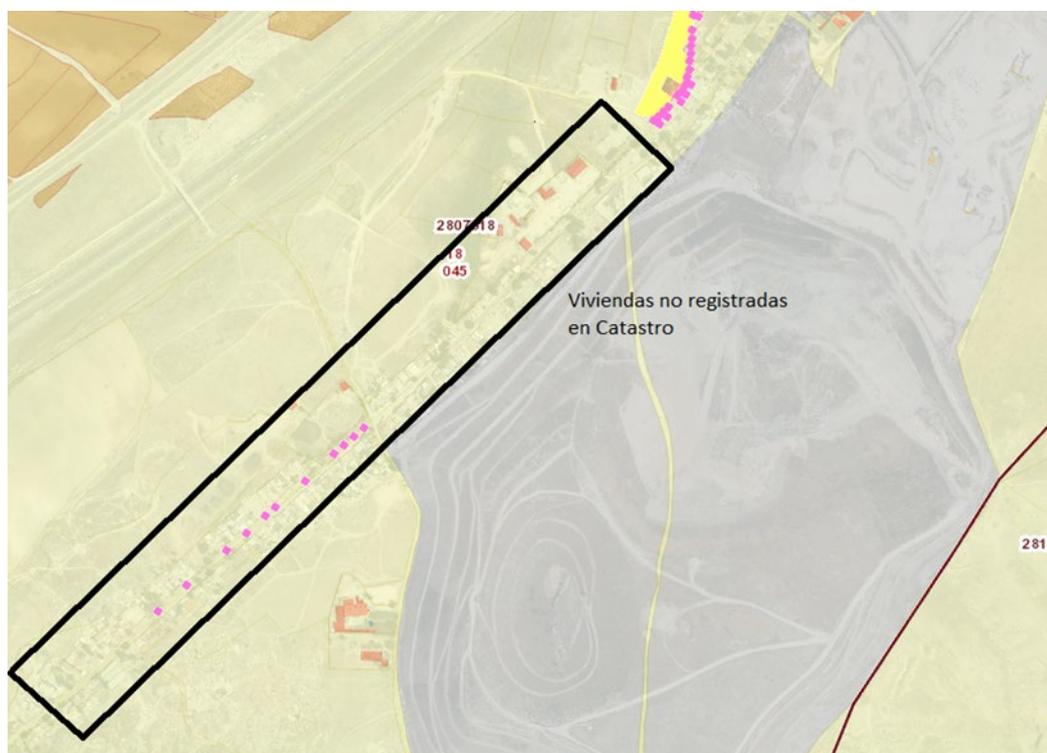


A veces, este problema se resuelve directamente utilizando únicamente la información catastral, cuando hay distintas construcciones asociadas a un bien inmueble, tal y como se ha visto en el apartado anterior. Para los casos que no se consiguen solucionar de esta forma, se utiliza la información del enlace entre

personas y viviendas, de modo que hay sospechas de falta de división horizontal cuando una vivienda, con su referencia catastral, se asocia a más de un hogar. Se han encontrado casos de falta de división horizontal de distinta naturaleza: dentro de la totalidad del edificio, dentro de la planta, y dentro de la propia vivienda.

- Zonas no catastradas/mal catastradas. En determinadas zonas hay viviendas sin catastrar, por ejemplo, la Cañada Real de Madrid. Estas zonas se detectan utilizando la información padronal de las personas y la ortofotografía. Se crean parcelas “ficticias” de forma manual, a las que se asocia unas coordenadas geográficas, así como la información de hogares.

Figura 35. Sección censal con presencia de viviendas no registradas (IDE-INE)



- Bienes inmuebles o construcciones no clasificadas como viviendas en los ficheros catastrales, pero se usan como tal. Hay bienes inmuebles que no están clasificados como viviendas, por ejemplo, oficinas, pero en la que habitan personas, por lo que deben incluirse en el censo. Estos inmuebles se obtienen de fundamentalmente porque se declaran como vivienda habitual en la declaración de la renta. Son los registros del grupo 2 de la fase 1 del procedimiento de enlace. En algún caso se obtienen por direcciones postales. También hay casos clasificados como suelo sin edificar que realmente están edificados, parece una falta de actualización de los ficheros de catastro. Se incluyen.

Figura 36. Sección censal con población empadronada en parcelas no destinadas a viviendas (IDE-INE).



- Depuración de almacenes y plazas de garaje clasificados con destino vivienda. Se descartan algunos casos mal clasificados, para evitar la inclusión de garajes y aparcamientos o trasteros.

7.6 IDENTIFICADOR ÚNICO

Una vez completado el proceso de asignación de referencias catastrales y solucionados todos los problemas, de las RC asignadas a hogares, unas estarán a 20 posiciones y otras a 14, pero el objetivo era tener un identificador único para cada vivienda y que se asocie a cada uno de los hogares: con este objetivo se crea artificialmente la “referencia catastral a 24 posiciones”, la RC24.

La RC24 se ha construido en función de si el bien inmueble presenta división horizontal o no, si la misma RC24 de la parcela aparece en distintas provincias, o si la parcela ha sido dividida.

Esta RC24 se asigna a cada uno de los hogares, según la RC20 o RC14 que tenían inicialmente. En determinados casos, cuando era posible, cruzando por planta y puerta. De modo que a cada hogar se ha asociado un código único de vivienda.

7.7 VARIABLES DE VIVIENDA

A continuación, se detalla cómo se ha construido cada una de las variables relativas a viviendas del censo.

- **Coordenadas geográficas.** Se obtienen a partir de los ficheros de Catastro, y de los ficheros recibidos de País Vasco y Navarra. En aproximadamente el 0,36% de las viviendas se han tenido que estimar, por diversos motivos.

Entre las que se han tenido que estimar, el 50,7% (0,18% del total) se han estimado mediante el centroide de la parcela. Son parcelas que no tenían coordenadas en los ficheros originales o estaban desviadas. El 36,4% (0,13% del total) mediante polígonos del SIOSE (Es el Sistema de Información sobre Ocupación del Suelo de España, integrado dentro del Plan Nacional de Observación del Territorio (PNOT) cuyo objetivo es generar una base de datos de Ocupación del Suelo para toda España). En SIOSE el territorio nacional está dividido en polígonos con una serie de tipologías, se seleccionan polígonos con tipologías compatibles de incluir viviendas y se asignan las coordenadas en función de esos polígonos y la sección censal. El resto de las coordenadas, 0,05% del total, se han obtenido de forma manual, mediante ortofotografía.

- **Tipo de vivienda según uso.** Las viviendas se clasifican en principal, si es vivienda principal para un hogar en cuestión, y, en caso contrario, será no principal. Esta variable se ha obtenido una vez finalizado el enlace de personas y viviendas, de modo que todas las viviendas que incluyen un hogar se marcan como principales.
- **Tipo de edificio en el que se encuentra la vivienda.** La clasificación de esta variable es:
 - Vivienda en edificio residencial con una vivienda.
 - Vivienda en edificio residencial con dos viviendas.
 - Viviendas en edificio residencial con tres o más viviendas.
 - Viviendas en edificio no residencial.

Para obtener la variable, se ha tenido en cuenta las claves de uso, destino y tipologías constructivas.

Si clave de uso o destino de la construcción está categorizado como vivienda en catastro y según las tipologías constructivas son edificaciones aisladas o pareadas, o bien, viviendas en línea o manzana cerrada, en este caso, se cuenta el número de viviendas que hay en la parcela. Si es una se categoriza como residencial con una vivienda, si son 2 se categoriza como residencial con 2 viviendas, y si son 3 o más, no se categorizan y pasan al grupo siguiente.

Para las viviendas que no se han clasificado anteriormente, se realizado una intersección de la capa edificios de catastro con la capa parcelas de catastro. Se cuenta el número de viviendas y edificios por parcela, prescindiendo de aquellos edificios que tienen un área muy pequeña, suelen ser anexos, cobertizos, etc. Se calcula el número de viviendas por edificio, y se asignan según el resultado: residencial con una vivienda, residencial dos viviendas, o residencial de tres o más viviendas.

Si la tipología constructiva es 011 o 013 (pisos en bloques de edificios o viviendas individuales con un solo edificio) entonces se cuenta el número de registros de viviendas por parcela, según el caso, se clasifica residencial con una, dos o tres o más viviendas

Si la tipología constructiva es distinta de los tres casos anteriores (011,012,013), es no residencial. En estos casos las claves de uso y destino son distintos de vivienda, es decir, por ejemplo, oficinas que se usan como viviendas.

Si la tipología constructiva es nula tenemos en cuenta la variable clave de uso, si es 'V', se cuenta el número de registros de viviendas por parcela y se actúa de forma similar a los casos anteriores.

Si la tipología constructiva es nula y clave de uso distinto de vivienda, se trata de viviendas en edificio no residencial.

– **Superficie útil de la vivienda.**

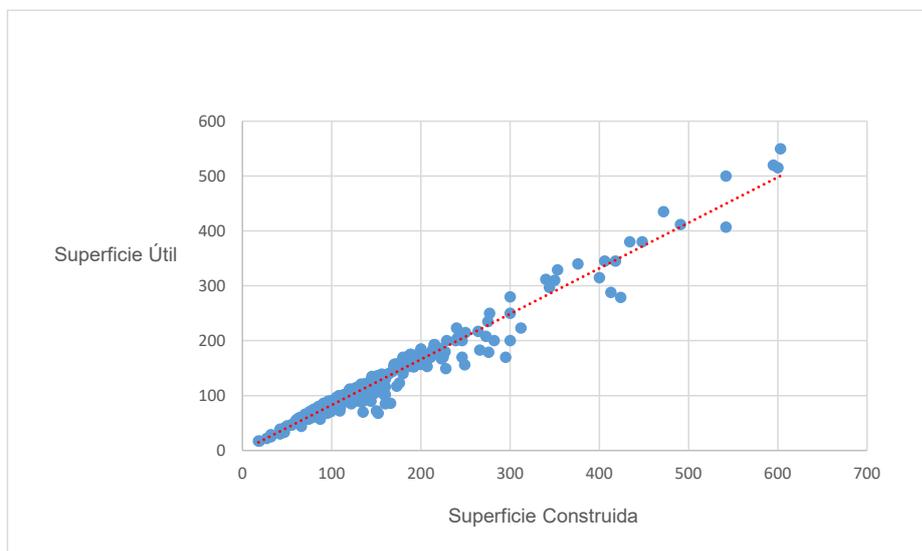
Los datos de Catastro no contienen información sobre la superficie útil, contienen información sobre la superficie construida.

Para calcular la superficie construida por vivienda, en el mismo proceso por el cual se obtiene el directorio de viviendas a partir de Catastro, se obtiene la superficie construida de las distintas construcciones que pasan a formar parte de la vivienda. Una vez agregadas, el resultado será la superficie construida de cada una de las viviendas. En los casos en los que solo uno de los registros pasa a formar parte del directorio de viviendas. Por ejemplo, una unifamiliar de tres plantas, tendrá tres registros como mínimo, uno por cada planta, y aunque solo se seleccione uno para formar parte del directorio de viviendas, se obtiene la superficie sumando las superficies de las distintas construcciones, siempre que las tipologías constructivas sean compatibles.

Ante el requerimiento por parte de Eurostat de obtener la superficie útil de la vivienda, ha sido necesario realizar una estimación a partir de un modelo que relacione ambas superficies. Para ello, se ha seleccionado una muestra actualizada con representación de todas las provincias, de distintos portales inmobiliarios de manera que las viviendas de la muestra contenían información tanto de superficie útil como construida.

Una vez depurada la muestra, y analizada la relación entre ambas variables se ha encontrado que el mejor modelo para estimar la superficie útil a partir de la construida, resulta ser la recta de mínimos cuadrados que pasa por el origen. El gráfico de dispersión es el siguiente:

Figura 37. Gráfico de dispersión. Superficie útil y superficie construida.



– **Régimen de tenencia de la vivienda**

- Vivienda en propiedad (ya sea por compra, pagada o con pagos pendientes, o por herencia o donación)
- Vivienda en alquiler
- Vivienda con otro régimen de tenencia (cedida gratis o a bajo precio por otro hogar o la empresa, u otra forma)

En el mismo procedimiento de asignación de referencias catastrales a partir de la información tributaria, se va incorporando el régimen de tenencia de las viviendas a partir de la información de las variables propiedad y usufructo (indicadoras de si el titular del bien inmueble posee la propiedad y el usufructo del bien) y el vector uso.

Estas variables se reciben por parte de la Administración Estatal de la Agencia Tributaria (modelo 100).

También se utiliza información sobre la titularidad de los inmuebles procedente del Registro de situaciones Finales de Titularidad. Contiene datos identificativos de bienes inmuebles, referencia catastral a veinte posiciones, junto con datos identificativos de sus correspondientes titulares catastrales.

El vector uso es un vector de cuatro posiciones que puede tomar el valor 0, o 1 según presente la característica o no. Así, el primer valor indica si es vivienda habitual o no, el segundo indica si es una vivienda en la que, en los casos de separación o divorcio, residen los hijos y/o cónyuge, el tercero si está a disposición de los titulares, y el cuarto si es arrendamiento.

Con esta información, se ha conseguido régimen de tenencia para algo más del 75% de los hogares. Para el resto, se ha realizado un proceso de imputación automática,

utilizando las frecuencias externas de la ECEPOV (Encuesta de Características Esenciales de Población y Viviendas).

- **Año de construcción.** El año de construcción se obtiene a partir de los registros tipo 13 de unidad constructiva, vía referencia catastral de la parcela o finca y el código de la unidad constructiva que tenga asociado la vivienda.
- **Plantas sobre y bajo rasante.** Para obtener las plantas sobre y bajo rasante del edificio en el que se encuentra la vivienda primero se normaliza el campo planta.

Se considera que una planta está sobre rasante cuando todo su pavimento inferior se encuentra por encima, o a la misma altura, de la rasante que determina la acera o el terreno.

Todas las plantas que estén por debajo de dicha altura se consideran bajo rasante.

En el caso de plantas bajo rasante, solo se incluyen las plantas si hay locales comerciales, oficinas o viviendas. Se excluyen plantas dedicadas a trasteros, almacenes, garajes.

8 Clasificación de las viviendas según uso a partir del consumo eléctrico

El método clásico, usado hasta el censo de 2011, para la determinación de la clase de vivienda convencional según su grado de utilización (principal, secundaria, vacía) se basa en la cumplimentación de esta información por parte de un agente censal. Este método es altamente subjetivo. Por ejemplo, muchas veces, aunque el agente se encuentre enfrente de la vivienda no hay evidencias para discernir si se trata de una vivienda secundaria o de una vacía.

La clasificación utilizada en el censo de 2021 es mucho más precisa, ya que está basada en información y datos más objetivos, como los del consumo eléctrico que la vivienda ha realizado a lo largo de todo el año precedente al momento censal (1 de enero de 2021).

Los resultados presentados tienen aún la consideración de provisionales dado que es preciso profundizar aún en la mejora de las fuentes y en su análisis.

8.1 VIVIENDAS PRINCIPALES Y NO PRINCIPALES

La distinción fundamental, y aquella a la que obliga el actual Reglamento comunitario para las viviendas convencionales es la de vivienda principal y no principal.

Determinar las viviendas principales de forma consistente con el proyecto censal en su conjunto pasa, como ya se ha mencionado, por tener un identificador que permita asociar el fichero de personas (FCP) con el fichero de viviendas (FCV). El número de

viviendas principales en el FCV coincide por definición con el número de hogares en el fichero censal de personas.

En cuanto a las viviendas no principales, en censos anteriores se acudió a una desagregación en secundarias, vacías y otras. Esta última categoría, utilizada durante la recogida del censo de 2011, se integró en los resultados junto con las viviendas secundarias dada la cierta arbitrariedad que tuvo su definición y por tanto la heterogeneidad en su uso por parte de los agentes censales.

8.2 CLASIFICACIÓN DE VIVIENDAS SEGÚN CONSUMO ELÉCTRICO

La clasificación utilizada en el censo de 2021 es mucho más precisa, ya que está basada en el dato objetivo del consumo eléctrico que la vivienda ha realizado a lo largo de todo el año 2020. Pero tiene el inconveniente de que no permite identificar, para cada vivienda concreta del fichero censal de viviendas, el consumo realizado y por tanto su clasificación según uso. Por tanto, el censo de viviendas y la clasificación de las viviendas según el uso se deben analizar por separado.

Se van a proporcionar datos agregados para todos los municipios mayores de 1.000 habitantes (que representan el 97% de la población), agrupando a nivel provincial los resultados de los menores de 1.000 habitantes.

Fuente de datos

La fuente de información sobre consumo eléctrico utilizada es la contenida es la contenida en el modelo 159 “Declaración anual de consumo de energía eléctrica”¹, una declaración informativa anual que deben presentar las compañías comercializadoras de los suministros de energía eléctrica a las autoridades tributarias (AEAT, HH.FF), correspondiente al ejercicio 2020.

Para cada CUPS (código universal punto suministro, equivalente a un contador) y cada contrato (ya que puede haber más de uno durante el año) los ficheros tributarios contienen información sobre el titular del contrato, fechas de alta y baja, la dirección del suministro y el consumo realizado.

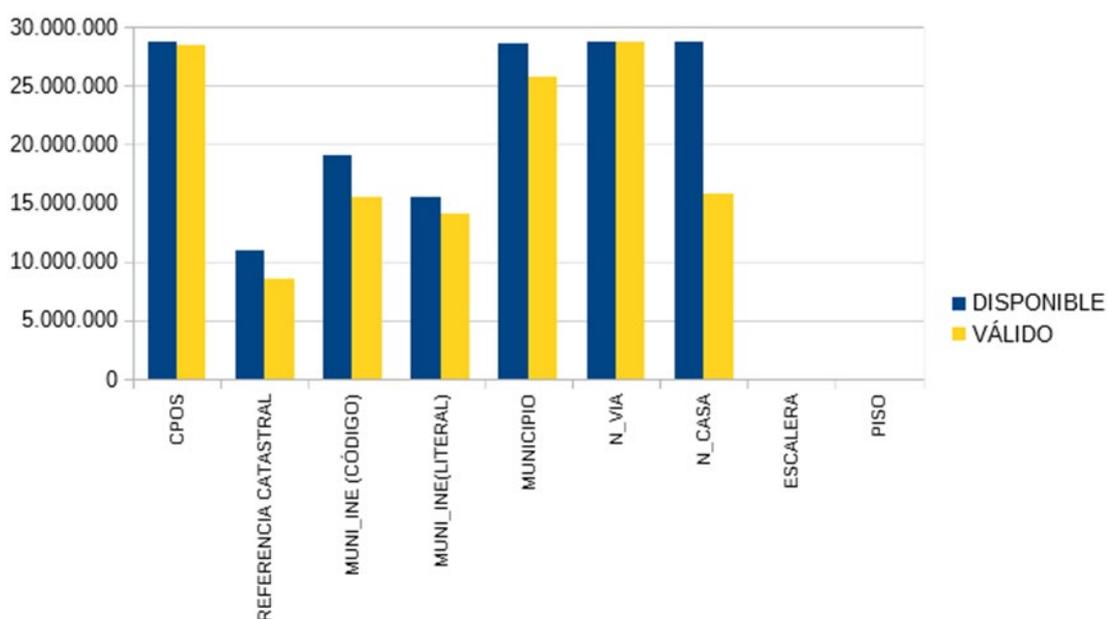
Datos del contrato	NIF del titular
	Tipo de NIF (Persona física, S.A, S.L, Comunidad Propietarios, Comunidad de Bienes, Corporación Local,)
	Razón social o nombre y apellidos del titular Fechas alta y baja contrato
Datos de la dirección del suministro	CUPS (Código universal punto de suministro)
	Provincia, municipio, código postal, tipo y nombre de vía, número, bloque, portal, escalera, piso, puerta
	Referencia catastral (de finca o individualizada)
Datos del consumo	Consumos mensuales, tipo de lectura (real, estimada), unidad de medida (Kwh, Megawh, Gigawh, Terawh)
	Importe total, potencia contratada

<https://www.boe.es/eli/es/o/2009/07/16/eha2041>

La información explotada se construye con las declaraciones de más de 200 comercializadoras de suministro eléctrico y contiene datos individuales de 32 millones de contratos correspondientes a 28,8 millones de contadores, de los cuales se determina que a 1 de enero de 2021 un total de 25.094.891 corresponden a viviendas.

Tras el análisis de la calidad de los datos de cada uno de los 28,8 millones de contadores en cuanto a la referencia catastral y elementos de la dirección, se concluye que, a día de hoy, no es factible enlazar individualmente la totalidad de los contadores con su correspondiente vivienda del fichero censal.

Se determina que se va a ofrecer una explotación agregada del número de viviendas clasificadas según su consumo para los municipios mayores de 1.000 habitantes y un indicador de intensidad de uso, la mediana del consumo anual, para los distritos censales de aquellos municipios en los que se haya conseguido asignar distrito a más del 75% de contadores.



Clasificación de las viviendas según consumo eléctrico: Viviendas vacías y de uso esporádico

A partir de la distribución de las viviendas según consumo eléctrico se ha construido una primera clasificación de las viviendas según su uso considerando cuatro grupos: *viviendas vacías*, *viviendas con muy bajo consumo*, *viviendas de uso esporádico* y *resto de viviendas*.

Se ha optado por considerar como **vivienda vacía** a aquella que:

- no dispone de contrato de suministro eléctrico, o

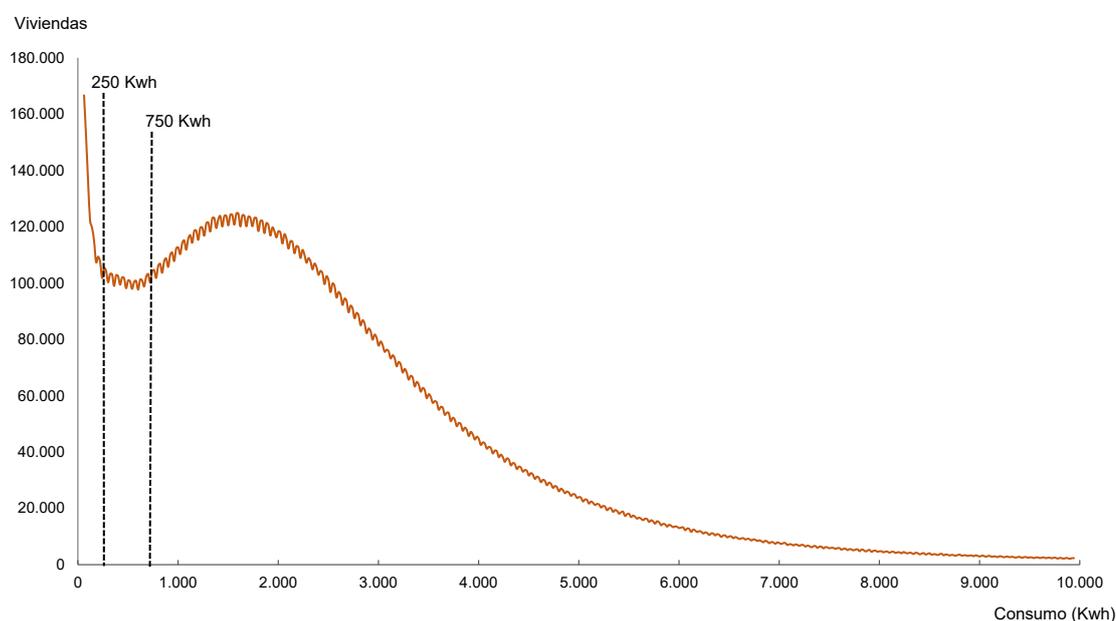
- su consumo total registrado en el año precedente ha sido menor al consumo que una vivienda media de ese mismo municipio dividido por 24 (lo que significa contar el consumo medio de **15 días en todo el año**).

La determinación de umbrales específicos para cada municipio se debe a que los consumos medios varían mucho por zonas climáticas, siendo en general superiores en zonas con veranos cálidos. Para determinar estos umbrales se ha considerado la mediana del consumo de aquellas viviendas con consumo superior a 250 kWh, tomando un umbral mínimo de 75kWh y máximo de 180 kWh.

No obstante, para facilitar el análisis, se ha establecido otro umbral –fijo para todos los municipios- de 250 kWh que, de media y de forma aproximada, correspondería al consumo de una vivienda que estuviera ocupada durante un mes en el año. A las viviendas con consumo por debajo de este umbral se las ha catalogado como de **viviendas de muy bajo consumo**.

Por su parte, las viviendas con un consumo entre 251 kWh y 750 kWh (correspondería, siempre de media, a un uso de la vivienda durante uno a tres meses durante todo el año) se han denominado viviendas de **uso esporádico**.

El resto de las viviendas se clasifican por tramos de consumo hasta más de 10.000 kWh



9 Sistema de difusión de resultados censales

Los principales objetivos en la difusión de los resultados del Censo de 2021 son:

- Maximizar el contenido de la información ofrecida a todos los usuarios teniendo en cuenta la preservación del secreto estadístico

- b. Aprovechar las nuevas posibilidades de análisis geográfico de la información que se derivan de la georreferenciación de la población
- c. Acortar los plazos de disponibilidad de la información de la operación censal ofreciendo los distintos resultados, a medida que vayan estando disponibles

La difusión de los censos de población y viviendas en España se ha situado en la vanguardia entre los países de nuestro entorno desde que durante la década de los 90 se editaran los primeros CD-ROM (la colección de publicaciones CERCA) con resultados de los censos de 1991.

El censo de 2001 estrenó un sistema de difusión, por entonces muy novedoso, que no ofrecía a los usuarios una lista de tablas predefinidas sino que les proporcionaba una herramienta para que, con ciertas limitaciones, el usuario construyera su propia consulta.

Esta misma fue la estrategia del sistema de difusión de los censos de 2011, si bien condicionada por el hecho de que los datos provenían de una muestra al 10% de la población, lo que limitaba la potencia a la hora de cruzar información muy detallada.

Se ha aprovechado la experiencia previa de los Censos de 1991, 2001 y 2011– tanto del INE como productor como de los usuarios a la hora de explotar la información censal–. El sistema de difusión censal de 2021 (en adelante SDC-2021) es un sistema de tipo data warehouse, en el que el usuario elige la combinación de variables y el área geográfica con mucha mayor libertad que en un sistema cerrado de tablas predefinidas.

Las posibilidades de difusión que permite el Censo 2021 son muy superiores al del Censo 2011 debido a que se recupera el carácter exhaustivo (el censo de 2011 se basó en una muestra cercana al 10% de la población). Por tanto, los límites que se imponen a las consultas de los usuarios se parecerán más a los del SDC de 2001 (impuestos por la preservación del secreto estadístico), al no tener cabida las restricciones por falta de representatividad muestral que se aplicaron en el SDC 2011.

Pero el SDC-2021 presenta una diferencia cualitativa con el SDC-2001 y es que toda la población estará georreferenciada, con unas coordenadas puntuales, lo que permite seleccionar información no solo basada en límites administrativos (como son los municipios o las secciones censales) sino referida al territorio de forma libre, independiente de los límites administrativos.

En la publicación del 30 de junio de 2023 se ofrecen resultados definitivos relativos a la población, hogares y viviendas hasta el detalle de sección censal. En las próximas semanas se habilitará la funcionalidad de navegación libre.

9.1 LÍNEAS GENERALES DEL SISTEMA DE DIFUSIÓN CENSAL (SDC-2021)

La funcionalidad completa del SDC se desplegará tras la publicación completa de los datos censales.

El SDC-2021 es un sistema interactivo en el que, en primer lugar, el usuario debe elegir entre consulta de tablas (es decir, navegación alfanumérica a partir de variables) o navegación a partir de mapas.

Consulta de tablas

Si se elige la primera opción, la navegación por tablas, el SDC-2021, como ocurriera en censos anteriores, ofrece, en primera instancia, una lista bastante completa de tablas predefinidas. Esta lista de tablas es el equivalente, en cualquier otra estadística, al plan de tablas que habitualmente se genera como producto de difusión. En los censos de 1991 y anteriores el censo daba lugar a una importante colección de tomos donde se detallaban los resultados de forma temática y por territorios.

Asimismo se ofrece al usuario la posibilidad de crear tablas. El usuario debe elegir en primer lugar el colectivo sobre el cual efectúa la consulta, define el ámbito geográfico (nacional, CCAA, provincias, municipios, secciones censales) y elige las variables que intervienen mediante un procedimiento interactivo típico en el cual compone las filas y columnas de la tabla a construir.

Consulta de mapas

El SDC-2021 ofrece al usuario que la primera interacción con el sistema sea a través de mapas. En este caso, de forma análoga a como ocurre con la consulta de tablas, el sistema permite en primer lugar acceder a unos mapas predefinidos en los se ofrecen los principales indicadores en forma de mapas así como la posibilidad de elegir su ámbito geográfico. Se ofrecen los siguientes tipos de acceso:

- **Mapas temáticos:** Se ofrece una lista de mapas predefinidos de forma análoga a las tablas predefinidas, en los que se incluyen datos a cualquier nivel geográfico, hasta sección censal. Tras la publicación de resultados definitivos del 30 junio se habilitarán los correspondientes mapas temáticos.

- **Análisis por áreas definidas por el usuario:**

La principal utilidad de una consulta de mapas radica en la posibilidad de que el usuario pueda elegir el ámbito geográfico de su consulta.

- En ese sentido, la consulta de mapas permite que el usuario navegue hasta encontrar el ámbito objeto de su interés y lo seleccione. El usuario puede estar interesado, no en una unidad administrativa dada sino en un grupo de ellas, lo que también puede seleccionar mediante herramientas típicas de selección geográfica de cualquier sistema de información de este tipo.

A partir de ese momento el sistema envía al usuario a la consulta de tablas en la cual el ámbito geográfico es el seleccionado en el mapa.

- Pero la potencia mayor de un sistema como este se alcanza cuando se permite al usuario seleccionar en el territorio de forma independiente de los límites administrativos. Aquí es donde surge una limitación muy importante impuesta por el secreto estadístico (ver sección siguiente sobre este tema).

En el censo de 2011 se ensayó una nueva forma de difusión de forma más libre que la de selección de ámbitos administrativos, pero que no suponía una selección libre. Se partió de la rejilla de 1 km² y se construyeron rejillas de tamaño variable en función de la cantidad de observaciones (hay que tener en cuenta que los datos provenían de una muestra). Este sistema basado en rejilla de tamaño variable venía obligado por la necesidad de preservar el secreto estadístico.

El SCD-2021 permite navegación libre sin límites de ningún tipo, aunque introduce ciertas limitaciones a la información que se ofrece precisamente para preservar el secreto estadístico.

El SDC-2021 permite, por tanto, responder a preguntas como cuáles son las características de la población en un radio de 500 metros en torno a un determinado punto o a una distancia dada de un tramo de carretera o de una playa.

Esta forma de difusión –no basada en divisiones administrativas– puede realizarse dado que cada vivienda tiene asignadas unas coordenadas GPS aproximadas (las del edificio donde está situada).

Se han diseñado también servicios web que permiten realizar explotaciones automatizadas de la información para facilitar la reutilización de la información censal en otras aplicaciones que estarán operáticos con la tercera fase de difusión de la información censal.

9.2 PRESERVACIÓN DEL SECRETO ESTADÍSTICO EN EL SDC-2021

En toda operación estadística, un elemento esencial a considerar en el diseño de los resultados que se ofrecen, consiste en evitar desvelar datos confidenciales, es decir, impedir que directa o indirectamente se pueda identificar a una unidad de la que se proporciona información, sea esta una persona o una vivienda, en el caso del censo.

Las técnicas de protección del secreto estadístico se basan generalmente en actuar sobre un conjunto dado de tablas de resultados; cuando se trata de ficheros de microdatos completos de una encuesta, se evita dar ciertos detalles de variables o de ubicación geográfica –por ejemplo, solo se proporciona la provincia, pero no el municipio–, siempre con el fin de evitar la posible identificación directa o indirecta de las observaciones.

Nada de esto es posible en el censo de población. La protección de la confidencialidad se dificulta dado que el SDC-2021 permite que cada usuario efectúe su consulta con las variables que considere y eligiendo el ámbito geográfico, luego no existe una cantidad finita de tablas que analizar a priori para preservar el secreto estadístico.

Una restricción sencilla sería evitar la consulta de áreas con muy pocas observaciones; por ejemplo, el sistema podría impedir consultas sobre cualquier recinto que encerrara menos de 50 observaciones. Pero el problema se produce con las consultas sucesivas por diferencias entre dos áreas. Así, un usuario podría estar tentado de consultar un área que encerrase 100 observaciones y a continuación, otra similar en la que solo selecciona una observación más, de forma que ambas consultas son válidas (superarían el umbral de 50) y no desvelarían a priori secreto alguno, pero la diferencia entre los resultados que ofrece la segunda y la primera serían debidos a la observación adicional, que quedaría completamente desvelada.

Se deben imponer restricciones de secreto estadístico, que juegan en contra de la potencia que los usuarios necesitan y esperan de un sistema basado en la libre consulta de variables y selección de ámbitos geográficos. Para poder conjugar el equilibrio entre maximizar la capacidad informativa del SDC-2021 y preservar la confidencialidad, el SDC-2021 impone tres tipos de limitaciones:

a) Variables que intervienen simultáneamente en una consulta

Con objeto de preservar la confidencialidad de la información individual en las consultas que se realicen al sistema, el menú de variables que se ofrecerá dependerá del número de unidades objeto de consulta (personas, viviendas, hogares) por lo que el sistema debe, en primer lugar, calcular el número de ellas contenido en el ámbito geográfico de consulta.

Se homogeneiza el tratamiento del secreto estadístico de tal forma que tanto la navegación temática como la geográfica por límites administrativos den la misma información.

Las variables se clasifican según el grado de información que ofrecen en dos categorías de sensibilidad:

- Se consideran variables con sensibilidad 0 aquellas cuyo detalle es tal que puede proporcionarse en cualquier nivel de consulta. Por ejemplo: sexo, edad en grandes grupos, nacionalidad (español/extranjero).
- Se consideran variables con sensibilidad 1, 2 y 3 según el número de categorías que intervienen en cada una de ellas y cómo discriminan a la población.

En función del número de variable que intervienen en una consulta y de su grado de sensibilidad se calcula un indicador sobre nivel de precisión de dicha consulta y se compara con el umbral de precisión permitido que se ha prefijado en función del número de observaciones implicadas.

El sistema decide así, para cada consulta, si se sobrepasa el umbral de precisión e impedirá efectuar consultas que sobrepasen el umbral, advirtiendo al usuario con este texto:

“Por razones de confidencialidad de la información Por razones de confidencialidad, no es posible consultar un mayor detalle de la información”

Si no ha sido satisfactoria la ecuación anterior (y solo en ese caso), entonces se aplicará en segundo lugar el siguiente criterio:

Número de unidades de una celda sea igual o superior a 5.

Es decir, la consulta también es presentada si, a pesar de no cumplir el primer criterio, alguna de las celdas de la misma tiene 5 o más observaciones.

b) Redondeo a tres

Un método clásico de preservación del secreto estadístico es el del redondeo aleatorio. Se implantó ya en el censo de 2011, donde se impuso un redondeo a 5, que para este censo se ha decidido modificar por un redondeo aleatorio a múltiplos de 3. Esto significa que, todo dato, antes de ofrecerse por el SDC-2021 se somete a un redondeo aleatorio al múltiplo de tres más cercano por arriba o por abajo. Esto hace que las tablas pierdan aditividad pero impide desvelar información de una única unidad.

El redondeo se hará de forma aleatoria en cada consulta. Por tanto, una misma celda podrá tomar diferentes valores en consultas diferentes.

Para evitar confusión a los usuarios, en cada consulta el sistema mostrará la siguiente nota:

“Para garantizar el secreto estadístico todos los datos están redondeados aleatoriamente a múltiplos de tres”

Las únicas tablas que no se someten a este proceso son las tablas predefinidas. Así, la población extranjera de España en el Censo 2021 según las tablas predefinidas, que es la correcta, es de 5.402.702 personas pero cualquier consulta a una tabla en el SDC-2021 podrá arrojar como resultado el valor 5.402.700 o 5.402.703, que son los múltiplos de tres más cercanos.

9.3 FICHEROS DE MICRODATOS

Para los usuarios más avanzados, como ya se viene haciendo en censos anteriores, se generarán ficheros de microdatos públicos. Estos ficheros contendrán aproximadamente el 10% de la población y se evitará proporcionar el detalle geográfico máximo con el fin de preservar el secreto estadístico.

Estos ficheros se generarán una vez completada la publicación definitiva del censo de 2021.

9.4 PUBLICACIÓN DE RESULTADOS CENSALES POR PARTE DE EUROSTAT

Los datos de los censos de población de los Estados Miembros (EE.MM.) de la UE no solo son objeto de publicación por parte de los institutos nacionales de estadística sino que también Eurostat pone en marcha un sistema de difusión a escala europea. Los resultados son proporcionados por los EE.MM. a partir de los censos de cada país, luego son productos censales, pero el sistema europeo contiene sus propias tablas y algunos elementos que merece la pena destacar.

En particular, una novedad respecto del Censo de 2011 es la existencia de reglamentación específica para la información a difundir mediante esta rejilla. Eurostat ha solicitado de los INEs la entrega, ya desde diciembre de 2022, de unos primeros datos de población usando una **rejilla censal (Census grid)**. Esta división del territorio, promovida por Eurostat¹ desde el censo 2011, permite la difusión de un conjunto de datos europeos armonizados basados en una red uniforme y atiende a las recomendaciones de la Directiva Comunitaria INSPIRE.

La difusión de datos de los Censos sobre una malla ofrece un gran valor analítico dado que la malla es regular, el tamaño de las celdas es constante y son continuas para todo el territorio. Ofrece una subdivisión del territorio que no depende de límites administrativos, por tanto invariable, lo cual permite un fácil estudio de una variable a lo largo del tiempo. Para ello es necesario disponer de los datos de los Censos georreferenciados, y se asocian por sus coordenadas a la celda correspondiente de la malla elegida.

La publicación de resultados censales por parte de Eurostat, tanto en forma de rejilla como alfanumérica en forma de tablas (el Census Hub) será paulatina conforme se vayan recibiendo datos de los EE.MM. y no estará completa hasta finales de 2024. En

¹ REGLAMENTO DE EJECUCIÓN (UE) 2018/1799 DE LA COMISIÓN de 21 de noviembre de 2018 relativo al establecimiento de una medida estadística directa temporal para la difusión de temas seleccionados del censo de población y vivienda de 2021 geocodificados en una malla de 1 km²

el caso de los datos españoles, estos estarán disponibles en los sistemas de Eurostat prácticamente al mismo tiempo que en los propios sistemas del INE, si bien lo que publica Eurostat es un subconjunto pequeño de la publicación del INE.

10 Calendario de difusión

Aunque ya se ha completado la difusión del grueso de la mayoría de la información censal según el calendario inicialmente previsto:

- Noviembre 2022: variables relativas a la población
- Febrero 2023: resto de variables relativas a la población como la rama de actividad económica, la ocupación o el lugar de trabajo.
- Junio 2023: Publicación de resultados completos el censo relativos a los hogares y las viviendas y remisión de todos los resultados a Eurostat.

Es necesario generar aún algunos productos adicionales de difusión, como los ficheros de microdatos para los usuarios, que se publicarán en julio 2023.

Está previsto incorporar información adicional sobre fecundidad 1996-2021 a finales de 2023

11 Más allá de 2021: el futuro de los censos de población

El método mediante el cual se ha construido el Censo de Población y Viviendas 2021 lleva implícita la estrategia de futuro más allá de esa fecha. El censo se elabora mediante una generación sucesiva de ficheros precensales anuales. Para poder disponer de cada uno de esos ficheros se precisa la recepción y tratamiento de decenas de fuentes administrativas de todo tipo.

No parece razonable desmontar toda esta maquinaria –que proporciona información tan útil– una vez terminado el censo y volver a ponerla en marcha cuatro o cinco años después para preparar el censo de 2031. Pero además no resulta lógico seguir pensando en elaborar esta información cada diez años cuando se dispone de ella de manera continua.

De cara al futuro parece lógico descartar la frecuencia decenal para la información censal; pero el problema es fijar una nueva estrategia, lo que no es fácil. Por un lado está el problema de la puntualidad en las fuentes. Los datos del censo de 2021 estarán disponibles dos años después de la fecha de referencia, en parte debido a la disponibilidad de ficheros, que en muchos casos se tienen más de un año después de la fecha de referencia, en parte debido a los pesados procesos de enlace y depuración de la información. Diseñar una nueva estrategia que supusiera disponer las cifras de población anuales mediante signos de presencia, como en el Censo de 2021, daría lugar a publicar dos años después las cifras de población y sería un paso atrás respecto de lo que ya tenemos hoy, unas cifras de población publicadas con menos de un año de retraso.

Cabe pensar también si tiene sentido publicar cada año muchos de los detalles que ofrece un censo, o si esa información competiría con estadísticas sectoriales mucho

mejor diseñadas desde el punto de vista temático. Pensemos en las estadísticas del sistema educativo o las laborales.

Otro elemento digno de consideración es la estrategia europea a partir de 2021, que está siendo objeto de debate en la actualidad.

11.1 EL MARCO EUROPEO

Si analizamos la situación de los países en el ámbito europeo nos encontramos con gran heterogeneidad. Hay países que mantienen censos clásicos, para los que es un reto mayor abandonar ese sistema, países que con distinto grado de avance ya han empezado la transición y países que han completado la transición a fuentes administrativas o lo esperan hacer en 2021.

Teniendo en cuenta esta diversidad, el sistema estadístico europeo está debatiendo en la actualidad cuál debe ser la estrategia post-2021. Hay que tener en cuenta también que el actual Reglamento de Censos expira en 2023 y se necesita un nuevo marco legal.

Desde 2018 se trabaja en un nuevo reglamento europeo que unifique la demanda de estadísticas de stocks y flujos demográficos, que ahora mismo están dispersas en tres reglamentos distintos (uno cubre las estadísticas migratorias anuales, otro los fenómenos demográficos y stocks anuales y un tercero es el de los censos decenales).

En el momento de elaborar este documento metodológico no se dispone aún de una propuesta concluyente pero ya hay un borrador suficientemente maduro, que situaría la mayoría de las actuales demandas del censo con una frecuencia anual; en particular este sería el caso de las cifras de población a nivel municipal, que ahora se recopilan cada diez años. Además, se solicitaría esa información con un plazo de menos de un año. La entrada en vigor del nuevo reglamento se espera para 2025.

Proporcionar anualmente datos de población municipal, incluidas bastantes variables sociodemográficas (nivel educativo, estado civil, tipo de hogar...) no supone un desafío inabordable para la estadística demográfica en España pero sí obliga a plantear una estrategia de publicación de datos anuales inmediatamente después de la publicación del censo de 2021.

11.2 ESTRATEGIA POST-2021 EN ESPAÑA

El marco europeo fija unos mínimos a tener en cuenta, pero dada la situación en España, donde en 2021 se habrá logrado la transición completa a registros administrativos, cabe ir más allá en la producción de información sociodemográfica a partir de las nuevas fuentes.

La disponibilidad, en noviembre de 2022, de una nueva cifra de población estadística consolidada a 1 de enero de 2021 obliga a rehacer las series de población hacia el pasado (lo que tradicionalmente se ha conocido como *estimaciones intercensales*). Dado que se dispone de una serie de población anual, llamada "Cifras de Población", lo que se hace es una revisión hacia atrás de esta serie.

A partir de 2023 se inicia el proceso de publicación de resultados censales anuales. No todas las variables del censo 2021 se publicarán anualmente; en particular no se dan todavía las condiciones para publicar algunas variables del censo de viviendas. Si bien se ha barajado anteriormente la posibilidad de denominar a esta operación anual de alguna otra forma, finalmente, por lo acuñado del término, se ha optado por mantener el nombre de *censo* para las publicaciones anuales. Así, el censo dejará de ser una operación decenal y pasará a ser anual.

El censo de población anual será objeto de publicación de un proyecto específico, por lo que aquí solo se citan las líneas generales. Este censo anual podrá ofrecer datos de población con doce meses respecto de la fecha de referencia que será 1 de enero de cada año, pero ciertas variables, sobre todo las relativas a composición de los hogares, no podrán estar en ese plazo, y se publicarán con un plazo algo mayor (previsiblemente 18 meses).

Esto implicaría que a finales de 2023 se tendría que hacer un esfuerzo extraordinario para publicar el censo anual de 2022 y el de 2023 por lo que quizá la frecuencia definitiva no se alcance del todo hasta 2024.

Este censo anual obliga a rediseñar otras estadísticas anuales de población. En particular, los productos que explotan de manera directa el padrón como la Estadística del Padrón Continuo o la Estadística de Variaciones Residenciales se tienen que rediseñar a la luz de esta nueva situación.

Además, disponer de cifras censales anuales da lugar a una **explotación estadística longitudinal** no contemplada hasta hoy. Esta nueva explotación, que será parte del censo anual, se iniciará en 2023 con referencia a 2022, afrontando temas nuevos como la formación, modificación y disolución de hogares (emancipación, duración de las parejas, formas de convivencia...).