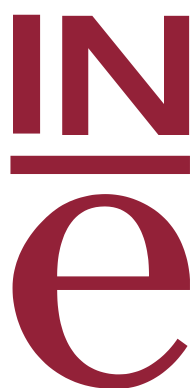


INSTITUTO NACIONAL DE ESTADISTICA



Censos de Población y Viviendas 2021

Proyecto técnico

Subdirección General de Estadísticas
Sociodemográficas
Septiembre 2019

Índice

1	Identificación de la operación	5
2	Origen de la demanda y justificación de su necesidad	7
3	Objetivos del proyecto	8
4	Clase de Operación	9
5	Contenido	9
5.1	POBLACIÓN OBJETO DE ESTUDIO	9
5.2	ÁMBITO GEOGRÁFICO O TERRITORIAL	10
5.3	PERIODO/S DE REFERENCIA DE LOS DATOS	10
5.4	VARIABLES DE ESTUDIO Y CLASIFICACIÓN	10
6	Características del proyecto	11
6.1	ANTECEDENTES: EL CENSO DE 2011 Y EL CONTEXTO INTERNACIONAL ACTUAL	11
6.2	EI CENSO DE POBLACIÓN Y VIVIENDAS 2021: PERSPECTIVA GENERAL	14
6.2.1	La elaboración progresiva de ficheros precensales	16
6.2.2	Elementos estratégicos del Censo 2021	17
7	El censo de población: personas y hogares	21
7.1	EL PRIMER OBJETIVO: CONTAR LA POBLACIÓN	21
7.1.1	Los datos sobre población en España desde 1996	22
7.1.2	La cifra de población del censo 2011	23
7.1.3	Las cifras de población desde 2012	26
7.2	LA CONSTRUCCIÓN DE UN FICHERO CENSAL DE POBLACIÓN PARA EL CENSO DE 2021: EL MÉTODO DE SIGNOS DE PRESENCIA	27
7.3	EL FICHERO CENSAL DE POBLACIÓN (FCP): QUÉ VARIABLES INTERVIENEN	32
7.3.1	Variables demográficas (sexo, edad, nacionalidad, lugar de nacimiento)	34
7.3.2	Variables relacionadas con las migraciones	35

7.3.3 Estado civil	40
7.3.4 Composición de hogares / relaciones de parentesco	44
7.3.5 Estudios en curso y nivel de estudios alcanzado	52
7.3.6 Relación con la actividad económica	58
7.3.7 Variables relativas a los ocupados	63
7.3.8 Estudio de las variables relacionadas con la movilidad	68
8 Censo de viviendas	71
8.1 ASPECTOS GENERALES DEL CENSO DE VIVIENDAS	71
8.2 EL FICHERO CENSAL DE VIVIENDAS (FCV)	73
8.2.1 Construcción del FCV a partir del Catastro	75
8.2.2 Enlace entre FCV y viviendas principales	80
8.3 DETERMINACIÓN DEL TIPO DE VIVIENDA SEGÚN USO	84
8.3.1 Viviendas principales y no principales	84
8.3.2. Clasificación de viviendas según consumo eléctrico	84
8.3.3. Locales	87
8.4 VARIABLES PARA TODAS LAS VIVIENDAS	88
8.5 VARIABLES DE LAS VIVIENDAS PRINCIPALES	90
8.5.1. Régimen de tenencia	90
8.5.2. Otras variables sobre las viviendas principales	93
8.6 ALOJAMIENTOS Y PERSONAS SIN HOGAR	94
9 Establecimientos colectivos	97
9.1 LA POBLACIÓN RESIDENTE EN ESTABLECIMIENTOS COLECTIVOS EN EL CENSO DE 2011	97
9.2 PLANTEAMIENTO PARA EL CENSO 2021	98
10 Incorporación de nuevas variables no investigadas en censos anteriores	100
10.1 NUEVAS VARIABLES DE CONTEXTO	101
10.2 LA ENCUESTA COMPLEMENTARIA (ECEPOV-2021)	102
11 Evaluación de la calidad	103

11.1 LA CALIDAD EN UN CENSO BASADO EN REGISTROS	103
11.2 INDICADORES DE CALIDAD DEL FICHERO CENSAL	105
11.3 INFORMES DE CALIDAD	107
12 Plan de difusión de resultados	108
13 Calendario de tareas	110
14 Más allá de 2021: el futuro de los censos de población	111
14.1 EL MARCO EUROPEO	112
14.2 ESTRATEGIA POST-2021 EN ESPAÑA	114
15 Estimación de costes	116
Anexo: Lista de variables de los Censos de Población y Viviendas 2021	117

Censos de Población y viviendas 2021.

Proyecto técnico

Versión septiembre 2019

1 Identificación de la operación

- Código y denominación PEN de la operación estadística que se presenta a dictamen.

7861 Censos de Población y Viviendas

- Unidad responsable de la operación.

Subdirección General de Estadísticas Sociodemográficas (INE)

- Organismos/Unidades colaboradores.

Por la magnitud de la operación, son decenas las instituciones que, con distinto grado de participación, colaboran o proporcionan información para elaborar los Censos. A lo largo del documento se menciona de forma más detallada. Se destaca en esta lista solo aquellas unidades que entregan registros administrativos al INE, si bien en algunos casos, para ello deben recopilarlos a su vez de otras fuentes.

- Agencia Estatal de Administración Tributaria (AEAT)
- Haciendas Forales de Bizkaia y Gipuzkoa¹
- Hacienda Tributaria de Navarra
- Ministerio de Trabajo, Migraciones y Seguridad Social (INSS, TGSS, SEPE)
- Ministerio de Justicia. (MUGEJU, DGRN)
- Ministerio de Hacienda (DG Catastro, clases Pasivas)
- Ministerio de Política Territorial y Función Pública (MUFACE)
- Ministerio de Educación y Formación Profesional
- Ministerio de Defensa (ISFAS)
- Ministerio de Administraciones Públicas (MUFACE)
- Ministerio del Interior

¹ Se está tramitando el convenio con la Diputación Foral de Álava

- Ministerio de Ciencia, Innovación y Universidades

- Antecedentes de la operación:
 - Los censos de Población y Viviendas se elaboran cada diez años. La última edición corresponde a 2011

- Fecha del último dictamen:
 - Los Censos de Población y Viviendas 2011 fueron dictaminados en la reunión de la Comisión Permanente del CSE de 6-10-2010

- Motivación de la solicitud de dictamen:
 - Para la edición de 2021 del proyecto se prevén importantes novedades metodológicas que hacen imprescindible su sometimiento a nuevo dictamen.

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, entendiéndolo 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 que se va a realizar en 2021 será 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 será en el que se haga el censo en un mayor número de países, entre ellos la mayor parte de los europeos.

Como ya ocurriera con la edición de 2011, los censos en la Unión Europea (UE) se realizan 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), asegurará 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 no se ha modificado, 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 geocodificados 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 éstos 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 del Padrón, el Censo realiza ese recuento con fines y metodología distintos del padrón (más información en el apartado 7.1.1).
- 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, culturales, 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) 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 Clase de operación

Resulta complicado clasificar la operación censal, en la que intervienen muy distintos procesos, como se describe a lo largo del proyecto. Esencialmente se trata de una estadística por enumeración exhaustiva, basada en fuentes administrativas.

Desde el punto de vista del producto final, los resultados se presentan agregados en tablas (también se ofrecen muestras de microdatos anonimizados para usuarios avanzados).

5 Contenido

5.1 POBLACIÓN OBJETO DE ESTUDIO

En el Censo de Población se incluyen únicamente las personas, de cualquier nacionalidad, que tienen fijada su “residencia habitual” en el territorio nacional.

El ámbito poblacional del Censo de Viviendas abarca a las viviendas y a los establecimientos colectivos.

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)² la opción de contar como población residente la población registrada (empadronada, en el caso español). Se detalla el ámbito poblacional en los apartados 7.1 y 7.2.

² 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>

5.2 ÁMBITO GEOGRÁFICO O TERRITORIAL

Todo el territorio nacional

5.3 PERIODO/S DE REFERENCIA DE LOS DATOS

Los recuentos estarán referidos a **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).

5.4 VARIABLES DE ESTUDIO Y CLASIFICACIÓN

Las variables objeto de estudio son las personas, los hogares y las viviendas.

En cuanto a las variables de clasificación, se construyen más de dos centenares de variables entre las que son directamente recogidas y variables derivadas. Ambos tipos de variables se describen a lo largo de este documento. En el Anexo figura una lista más detallada.

6 Características del proyecto

6.1 ANTECEDENTES: EL CENSO DE 2011 Y EL CONTEXTO INTERNACIONAL ACTUAL

En 2011 se realizó el último censo de población hasta hoy. Esta edición 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%.

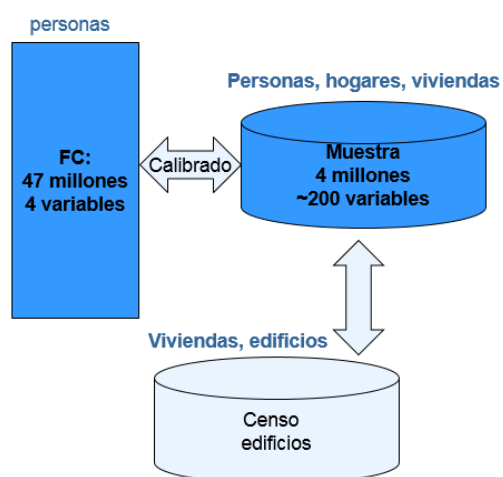
La operación de 2011 dejó de ser un “censo” desde la perspectiva del **proceso** seguido para obtener los datos, si bien sí puede considerarse que el **producto** obtenido era un verdadero censo, al menos para una parte de las variables, en tanto que los datos procedían de un fichero que contenía todas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

El producto censal era el fruto de una combinación de fuentes, 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.

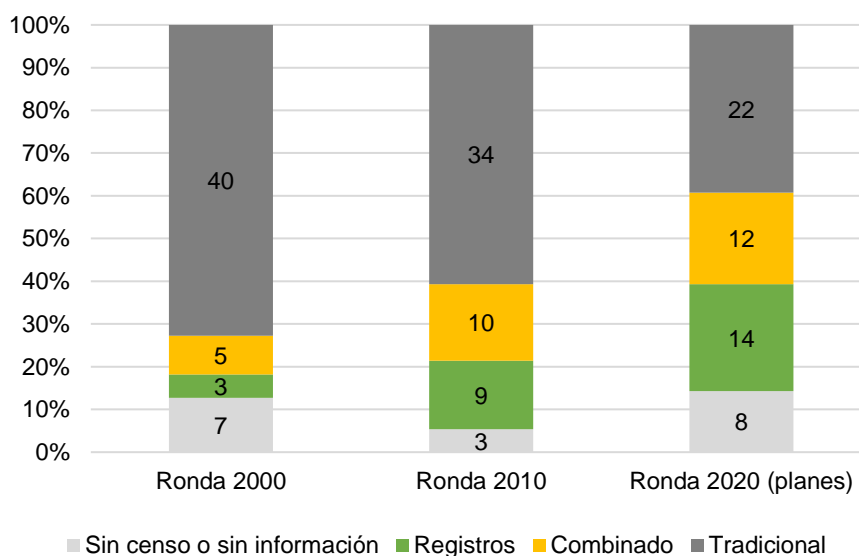
Respecto del panorama internacional, el censo de población es seguramente la operación estadística que se realiza en un mayor número de países del mundo. Según datos de Naciones Unidas, en la ronda 2005-2014, un total de 214 países (que representaban un 93% de la población) realizaron un censo, la mayoría 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.

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.

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 países³, entre los que se encuentran muchos de los más desarrollados del mundo.

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 tienen los siguientes resultados:

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



³ 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.

Como se describe a continuación, España, se sumará en 2021 al reducido grupo de países que realizan el censo basándose en registros administrativos. Además hay ocho países que tienen planificado pasar del Censo tradicional en la ronda de 2010 al Censo combinado en la ronda de 2020. Son los siguientes: Canadá, Chipre, República Checa, Italia, Luxemburgo, Montenegro, Rumanía y Reino Unido.

6.2 EL CENSO DE POBLACIÓN Y VIVIENDAS 2021: PERSPECTIVA GENERAL

En el censo de 2011, si bien se disponía de dos fuentes con datos de personas –el fichero censal y los de la muestra–, el fichero censal solo se utilizó para obtener el recuento de población por municipios (ver apartado 3.1). 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 cuatro variables para cada persona.

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 éste, 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. Cabe destacar las siguientes:

- ✓ El censo de población **es perfectamente factible** basándose exclusivamente en registros administrativos, es decir, no sería 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 Censal⁴ de la UE se cumple 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 descarta 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 implica una decisión estratégica de gran calado. El INE se suma 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 será una operación invisible para la sociedad en general. Esta operación se realizará con un coste cualitativamente inferior, asumiendo muchos menos riesgos de los que supone una operación de recogida masiva de encuestación a los hogares y ofrecerá unas garantías en cuanto a cobertura y calidad muy superiores.

⁴ A lo largo del documento se menciona en muchas ocasiones, por simplificar, el "Reglamento censal". Se trata, en realidad, de un conjunto de reglamentos. Ver Apartado 2.

El censo de 2021 es 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 puede 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.

6.2.1 La elaboración progresiva de ficheros precensales

En el anteproyecto del Censo, publicado en 2016, se anunciaba que la investigación continuaría durante los próximos años con el fin de estudiar detalladamente la viabilidad del modelo. Así, el fichero precensal 2016 (FPC-2016) 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 enlazan decenas de 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.

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, como se explica más adelante, del fichero precensal 2018, que ya aborda un censo de viviendas completo.

A lo largo de 2018 y 2019 se ha venido trabajando en la incorporación de mejoras que se sigue realizando paulatinamente, generándose ficheros precensales anuales. Así, se construyó un FPC-2017 pero que solo se centró en algunas variables.

El siguiente fichero completo es el FPC-2018, que aún no se ha terminado actualmente pero incorpora muchas de las posibles mejoras identificadas en el FPC-2016 y, sobre todo, realiza por primera vez una aproximación completa al Censo de viviendas.

A lo largo del documento se mencionan muchas de las mejoras previstas ya identificadas y que formarán parte de los sucesivos ficheros precensales. De aquí a 2021 se espera construir al menos un FPC-2020 que sirva ya de ensayo general

completo. El Censo de 2021 no será sino una actualización de los procesos ya probados en los sucesivos ficheros precensales por lo que los riesgos de la operación censal 2021 se minimizan.

6.2.2 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 la población se contará 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 viviendo solo, sin adultos.

¿Hasta qué punto los registros administrativos –particularmente el padrón– van a permitir reflejar la realidad del país? Existe una cierta percepción entre los usuarios de que la “realidad demográfica” y la realidad “administrativa” corren caminos separados.

Para empezar, las personas se pueden empadronar, por distintas razones, en lugares distintos de donde residen. Además, es indiscutible que hay cierta información de interés demográfico que no figura en los registros administrativos y que difícilmente va a figurar, al menos en algunos casos. Las relaciones de convivencia o de pareja que no sean declaradas ante una instancia administrativa; o el trabajo informal, no recogido en los ficheros laborales; o el régimen de tenencia de la vivienda en ciertos casos; son varios los ejemplos de información que tradicionalmente se obtiene mediante preguntas directas a los ciudadanos.

En primer lugar, una encuesta (o un censo clásico) se enfrenta siempre a la falta de colaboración. Es un hecho de magnitud creciente y muy preocupante en todos los países de nuestro entorno y está detrás de la búsqueda de otros métodos en muchos casos.

Además, la falta de colaboración no afecta por igual a toda la población, lo que implica que las encuestas sean sesgadas porque la población que contesta cuestionarios no representa correctamente al conjunto de la población.

Cabe hacer también una reflexión sobre hasta qué punto una parte de la realidad que no se registra en ámbitos administrativos llega a ser declarada en un cuestionario estadístico. No siempre será así y no debemos suponer que el cuestionario estadístico se percibe por los ciudadanos como algo tan confidencial como es en realidad. El cuestionario que recoge el INE en un censo clásico seguramente no refleja toda la realidad sino solo aquella que los informantes quieren revelar, incluido el hecho de residir o no en la vivienda.

Con todo, es cierto que una entrevista con un cuestionario estadístico puede aproximarse mejor en ciertos ámbitos a la realidad. Pero para eso precisamente se han creado las encuestas dirigidas a los hogares, que siguen teniendo todo su vigor:

un censo basado en registros administrativos debe complementarse con un buen sistema de encuestas sociales.

Por poner un ejemplo, los registros administrativos no van a ofrecer información correcta sobre convivencia de parejas no casadas, y seguramente subestimaríamos el número de parejas del mismo sexo. Seguramente una encuesta pueda ayudar a suplir esa carencia. Así, la información obtenida en encuestas se erige en un instrumento para ofrecer información, no tan detallada en el territorio, pero sí complementando al censo en temas que no puedan cubrirse desde los registros.

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

- **Se construirá un verdadero “producto censal” referido a 1 de enero de 2021:**

El censo de 2021 estará referido **a la fecha 1 de enero**. La operación censal dará 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 contendrá a todos los residentes en viviendas familiares
- un fichero de establecimientos colectivos que contendrá la población residente en ellos
- un fichero de viviendas, distinguiendo viviendas principales (donde vive la población) y no principales.

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

Así pues, el producto que se pretende construir sería similar al que se dispondría en caso de procesar los cuestionarios de todos los hogares de España mediante un censo clásico.

- **La información provendrá casi exclusivamente de registros administrativos pero se completará 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 subsistirá con total seguridad un residuo de hogares (parejas de hecho, familias reconstituidas...) para los que no sea 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 realizará una encuesta específica para complementar la información sobre algunas variables.**

El proyecto censal se complementará con la realización de una encuesta específica, denominada Encuesta sobre Características Esenciales de la Población y las Viviendas (ECEPOV-2021)⁵. Se desarrollará a primeros de 2021 y permitirá mejorar la imputación de algunas variables censales y proporcionar algunas otras que no se encuentran en registros administrativos. Se trataría 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 información sobre los hijos tenidos desde 1996, primer año desde el cual se puede enlazar con información de nacimientos.

Pero 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 es previsible que se pueda proporcionar individualmente el consumo eléctrico de las viviendas o el nivel del ingresos del hogar o de la persona pero sí se plantea proporcionar el **nivel medio de ingresos de cada sección censal** o municipio o datos sobre **consumo eléctrico** para ámbitos geográficos muy reducidos (normalmente secciones censales): porcentajes de viviendas por tramos de consumo, por ejemplo.

En algunos casos se renunciaría a recopilar la información tal como se publicó en el censo 2011 pero a cambio se ofrecerá información distinta, que sirva para medir el mismo fenómeno y que sí pueda obtenerse de registros. Este es el caso de la información sobre **movilidad cotidiana** (lugar de estudios y trabajo). No se puede dar esta información tal cual se ofrece en censos anteriores pero a cambio se trabaja en dar una aproximación diferente, mucho más rica en cuanto a cantidad y calidad, a partir de datos de localización de teléfonos móviles.

Al tratarse de movimientos de telefonía móvil anónimos, sin identificación del portador del teléfono, se perdería la capacidad de analizar esa movilidad cotidiana clasificada por variables demográficas (sexo, edad, etc.) pero se dispone de mucho mayor detalle en el conocimiento de los movimientos lo que a efectos de estudios de movilidad parece ser prioritario.

⁵ Para más detalles ver apartado 10.2. No obstante, se publicará un proyecto específico.

- **Se dejará de publicar alguna información contenida en censos anteriores.**

En todos los censos se ha dejado de publicar información que se recogía en censos anteriores, normalmente por su falta de actualidad o necesidad. Así, los censos de 2001 se eliminaron preguntas sobre la residencia de hecho y se recogerá solo información de residencia de derecho.

En concreto, a día de hoy, se prevé perder información sobre país de nacimiento de los padres, trabajo no remunerado y sobre equipamiento de viviendas. Esta información sí se recogerá en la ECEPOV-2021 por lo que estrictamente hablando no se pierde sino que se ofrece con menor detalle territorial que el resto.

- **Se ofrecerá una información de mayor calidad y una calidad más contrastable**

La disponibilidad de registros administrativos proporciona, en general, un salto cualitativo en la calidad de la información respecto de la recogida mediante cuestionarios.

Además se proporcionará información sobre la calidad de la información elaborada a partir de las fuentes implicadas en la construcción de cada variable (ver apartado 11.2).

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

El censo de 2021 debería ser el último censo tal como se contempla ahora, en el sentido de ofrecer una foto detallada del territorio una vez cada diez años; se deberá diseñar una estrategia para ofrecer información del mismo nivel de detalle pero de forma más periódica. Seguramente el censo dará paso a una explotación continua de un registro de población actualizado, lo que abrirá la puerta a nuevas estadísticas longitudinales de población y hogares.

7 El censo de población: personas y hogares

7.1 EL PRIMER OBJETIVO: CONTAR LA POBLACIÓN

El objetivo original de los censos es contar el número de habitantes que reside en un territorio. En 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.

¿Cómo se mide mejor el número de habitantes de un país?

Considerar como habitantes de un territorio a la población que figura registrada (empadronada, en el caso español) es solo una de las aproximaciones posibles a medir la población residente. Evidentemente siempre existe la posibilidad de que haya población que reside en España y que no figure empadronada o viceversa, por diversas causas.

De la misma forma podemos argumentar que un censo clásico que intente visitar todos los hogares puede no encontrar a toda la población, de forma que tampoco ese sistema garantiza una contabilidad exacta de la población residente. Además, el propio concepto estadístico de población residente hace que no toda la población que en un momento dado vive en un determinado territorio deba contabilizarse⁶.

Todo método de contabilización de la población puede subestimar o sobreestimar la población residente. Los métodos de recuento clásicos (por visita o envío de cuestionario a todos los hogares) tienden a subestimar población. Lógicamente es difícil que un hogar se deje censar dos veces (aunque la sobreestimación puede provenir por otros motivos); en cambio en todos los censos son muchos los hogares que se quedan sin encuestar por distintas razones, y una de ellas, cada vez más significativa últimamente es la falta de colaboración.

Por muchos esfuerzos que se realicen por entrevistar a la población siempre habrá hogares que no colaboren. Este es un problema creciente en todo el mundo y es una de las mayores preocupaciones de los responsables de los censos de población en los países donde aún se sigue el método clásico de envío de cuestionario a todos los hogares.

Se conoce con relativa precisión cual es el porcentaje de hogares que se niega a contestar en las encuestas del INE, que depende de varios factores (objeto de la encuesta, población objetivo, métodos y periodos de recogida...) pero incluso en el

⁶ Ver definición de residencia habitual en apartado 5.1

más favorable de los casos esa falta de respuesta supera el 10% de los hogares en la actualidad, incluso realizando múltiples intentos de contactar con los hogares.

Eso no significa que al final un censo de población, en general, cuente menos población de la que reside, porque la metodología censal debe tener en cuenta esa falta de respuesta para estimar la población que reside a partir de la población efectivamente entrevistada.

Otro tanto cabe decir de la información recogida en los cuestionarios que, en muchos casos, no figuran completos por lo que hay mucha falta de respuesta parcial. A falta de datos se deben imputar esas variables de los hogares que no contestan.

Un censo basado en registros no se enfrenta a ese problema creciente de la no respuesta o de la información incompleta. Pero puede tener problemas de otro tipo en cuanto a subestimar o sobreestimar la población.

Respecto de lo primero, cabe preguntarse hasta qué punto hay población que reside en España y que no está empadronada. Se puede hacer una primera estimación sobre hasta qué punto una encuesta puede ayudar a encontrar población no empadronada en España usando las propias encuestas del INE, que entrevistan a decenas de miles de hogares cada año.

La Encuesta de Población Activa (EPA) es la mayor encuesta de hogares habitual del INE. Para realizar cada trimestre esta encuesta se selecciona una muestra a partir de un marco de viviendas, independientemente del padrón. Posteriormente se cruzan las personas que han respondido a la encuesta con el padrón, con el fin de complementar la información con datos administrativos.

Por ejemplo, en el cuarto trimestre de 2016, ese cruce da como resultado que de un total de 134.412 personas de 16 o más años recogidos en los cuestionarios de la EPA ha solo hay 179 registros (un 0,13%) no encontrados en la base padronal, que permite buscar también personas con empadronamientos caducados. Esto no quiere decir que esas personas no estén o hayan estado empadronadas sino que, a falta de identificadores tipo DNI, NIE o pasaporte, la búsqueda por otros campos (nombre y apellidos, fecha de nacimiento,...) no ha permitido enlazar unívocamente con el padrón. Evidentemente entre los miles de hogares que bien se niegan a colaborar, bien no se localizan, figurarán también personas no empadronadas pero por razones obvias no es posible estimar ese colectivo.

7.1.1 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.

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”⁷.

⁷ 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.

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 2018 subsiste una diferencia a nivel nacional entre ambas fuentes de unos 65.000 habitantes (un 0,1%).

7.1.2 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.

Se construyó el llamado “fichero precensal” a 1 de noviembre de 2011, que es la fecha de referencia del Censo, a partir del aprovechamiento de los diversos registros administrativos disponibles, tomando al Padrón como elemento básico de su estructura.

El objetivo fundamental del fichero precensal era disponer de información suficiente para realizar el recuento de la población y de su estructura según las principales variables de clasificación (sexo, edad, nacionalidad, país de nacimiento).

Para ello, se partió de la base de datos del Padrón y de la información estadística y de fuentes administrativas (registros del DNI y Tarjetas de Residencia del Ministerio del Interior, nacimientos y defunciones del Movimiento Natural de la Población y del Registro Civil, Seguridad Social, Agencia Tributaria).

Se comprobó la presencia en otras fuentes de la población que figuraba en el padrón y también se comprobó que las defunciones y nacimientos estaban bien incorporados. No solo se cruzaron los registros vivos en padrón sino también registros eliminados por diversas razones, con el fin de hacer el análisis lo más exhaustivo posible.

Como producto de este proceso, en el caso de los españoles, se contabilizó esencialmente a la población que figuraba empadronada, aunque se corrigió por el cruce de defunciones y por la incorporación de nacimientos no contemplados en el padrón.

En el caso de los extranjeros, para la mayoría de las personas que figuraban en el fichero precensal se consideró confirmada su residencia en España ya que se consiguió encontrar esos registros de padrón en otras fuentes: eran afiliados dados de alta como trabajadores o pensionistas o figuraban como contribuyentes en las bases de datos de la Agencia Tributaria, todo ello durante el periodo de referencia (2011).

Los factores de recuento

A estas 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.

Otros registros tendrán un factor de recuento 0, y por tanto esas personas no se cuentan como residentes. Son aquellos que se han encontrado en los ficheros de defunciones y que aún figuran en el padrón, generalmente porque se trata de defunciones recientes aún no procesadas y en algunos casos porque no se ha podido determinar perfectamente el enlace entre persona fallecida y empadronada. Esta es una ventaja de la aproximación con fines estadísticos a los registros: no se puede borrar un registro del padrón al no tener certeza pero sí puede descontarse un residente a efectos estadísticos.

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

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 0 o 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.

Para ello se realizó una partición del fichero de registros dudosos (aproximadamente 1.046.000 personas) en clases o grupos de población definidos por características sociodemográficas: edad, según intervalos de edades agrupadas, la nacionalidad y la provincia de residencia.

Las clases se configuraron de forma que al menos contuvieran 1.000 registros dudosos y que todo registro dudoso perteneciera a una clase. Las clases que resultan con menos de 1.000 dudosos se agruparon con el objetivo de llegar al menos a ese umbral: primero, para una misma nacionalidad y provincia, se agrupaban en mayores intervalos de edad, si esto fuera insuficiente se agruparon para una misma nacionalidad a nivel de la Comunidad Autónoma y si aun así no se conseguía el mínimo se agrupaba para dicha nacionalidad a nivel nacional, por último y residualmente se llegaba a agrupar nacionalidades a nivel nacional. De esta forma el fichero precensal quedó configurado en 724 clases con al menos 1.000 dudosos cada una. La más numerosa fue la de ciudadanos del Reino Unido residentes en la provincia de Alicante con edades entre 65 y 69 años, con casi 6.000 dudosos, de unos 21.000 que figuraban en el fichero precensal.

Se estableció el mínimo de 1.000 dudosos para tener garantías de que se encontrasen representados en la encuesta ya que esta iba dirigida a un 10% de la población. Las características que definen las clases (edad, nacionalidad, provincia de residencia) son características observables al realizar la Encuesta ya que se recogían en el cuestionario

A partir del número teórico de dudosos que se debían haber recogido en la encuesta y del número real de ellos que se encontró (corrigiendo por falta de respuesta) se asignó a cada clase de dudosos un factor de recuento, es decir, una probabilidad de residir en España para las personas de ese grupo.

Así, por ejemplo, para los 6.000 dudosos de Reino Unido entre 65 y 69 años en la provincia de Alicante, resultó un factor de recuento de 0,36 (se encontraron en la muestra el 36% de los esperados)

A cambio, hubo clases con factores de recuento superior a 1: por ejemplo, los pakistaníes entre 25 y 30 años en la Comunidad de Madrid tienen un factor de recuento de 1,67 (se encuentran en la encuesta más de los esperados).

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

7.1.3 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.
- 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 sobre todo a la naturaleza administrativa del padrón, que introduce ciertas restricciones. Con las diversas mejoras puestas en marcha en los últimos años en la contabilización de población esa brecha entre serie estadística y padrón se ha hecho casi insignificante y es de esperar que esa distancia se reduzca. Con todo es previsible que siga habiendo alguna diferencia en el futuro.

Mantener los dos conjuntos de cifras de población no debería producir ninguna alarma si no se separan mucho entre sí y se pueden explicar las diferencias que tienen que ver con los distintos fines de ambas fuentes, padrón y serie estadística.

Si bien el padrón tiene como propósito también cuantificar la población residente, es cierto que por los muchos usos y derechos que se derivan para los ciudadanos a partir de él, no pueda ser siempre inmediata y automática su actualización. Por poner un ejemplo sencillo, una defunción de una persona cuyos datos de identificación no figuren correctamente no generará inmediatamente una baja en el padrón, mientras no se resuelva esa identificación, pero sí se contabilizará en la serie estadística la existencia de esa defunción.

7.2. LA CONSTRUCCIÓN DE UN FICHERO CENSAL DE POBLACIÓN PARA EL CENSO DE 2021: EL MÉTODO DE SIGNOS DE PRESENCIA

De cara al censo de población de 2021 el número de residentes se calculará de una forma que guarda ciertas similitudes con la de 2011: fundamentalmente se partirá de los datos del padrón pero se estimará con métodos distintos la población dudosa.

Si ya en el censo de 2011 esa población dudosa suponía en torno al 3% del total (todos ellos extranjeros) es de esperar que para 2021 sea inferior, dadas las mejoras producidas en los padrones municipales durante estos años. No sería eficiente basar el recuento de una población dudosa tan pequeña en una encuesta dirigida a una muestra de hogares.

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.

El uso de decimales es habitual en las encuestas por muestreo en las cuales la muestra representa a una población a partir de unos factores de ponderación o de elevación que son inherentes al proceso estadístico, pero no lo es en recuentos exhaustivos. Pero los factores de recuento complican la lógica del producto censal y se ha evitado su uso para 2021. Por ejemplo, pensemos que un determinado municipio está habitado por 100 personas con factor de recuento medio de 0,5, es fácil determinar que la población estimada para el municipio sería de 50 habitantes. Pero determinar cuál sería el número de viviendas ocupadas resultaría complejo y con cierto grado de arbitrariedad.

Para 2021, el método que se plantea haría innecesario el uso de factores de recuento y por tanto de decimales. El esquema que se planea es el de la figura 3. El producto censal sería un fichero de microdatos, es decir, contendría aproximadamente 47 millones de registros individuales, uno por cada persona.

Figura 3. Esquema del Fichero Censal de Población

Fichero censal de población											
Número	sexo	edad	nacionalidad	país	lugares de residencia anteriores	estudios en curso y nivel alcanzado	situación en relación con la actividad	ocupación	lugar trabajo, desplazamientos, vacaciones	relaciones con resto de miembros de hogar	
1	H	17	España	Ecuador	AA-BB-CC	YYY	ZZZ	xxx	YY	xxxx	
2	M	21	España	España	AA-BB-CC	YYY	ZZZ	xxx	YY	xxxx	
3	M	13	Marruecos	Marruecos	AA-BB-CC	YYY	ZZZ	xxx	YY	xxxx	
...	
...	
47.000.000	M	89	España	Francia	AA-BB-CC	YYY	ZZZ	xxx	YY	xxxx	

Los países que cuentan la población a partir de registros de población determinan si cada registro se contabiliza o no a partir del cruce con otros registros para determinar la verosimilitud de la residencia. Así, no se tienen registros seguros y dudosos sino que, para cada individuo, a partir del cruce con otros registros, se tiene que decidir,

mediante un modelo matemático, si el registro se considera válido o no. Estos métodos se conocen en la literatura reciente de censos de población como signos de presencia (o más informalmente, como “signos de vida”).

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.

Este trabajo ya se ha realizado de forma todavía no definitiva en el FPC-2016 y en el FPC-2018. Se toma al Padrón como elemento básico de su estructura y a partir de él se elabora un procedimiento para efectuar un recuento de la población.

Para ello, aunque en un principio se parte de la población que está de alta en la base padronal a fecha 1 de enero, se tienen en cuenta todos los movimientos de la base padronal completa, que incluye no sólo estos registros vivos en Padrón en la fecha de referencia censal, sino también todos aquellos registros de personas que en algún momento desde el nacimiento del padrón continuo, en 1996, han figurado empadronados y que se han eliminado por diferentes causas.

En una primera fase se añaden a la base padronal determinadas bajas de oficio previas a la fecha censal que aparecen después de esta fecha ("rebotan") como altas por omisión sin procedencia o con procedencia interior. Consideraremos que estas personas pueden haber estado residiendo todo el tiempo en España y, por ello, se toman como candidatas a ser añadidas a la cifra de población del FPC si se constata que tienen signos de presencia en los registros administrativos.

Por otra parte, la definición de residencia habitual del Reglamento censal indica que una estancia de menos de 12 meses no se considera un cambio de residencia habitual. Por tanto, las personas que residen menos de 12 meses en el extranjero, estando la fecha de referencia censal dentro de ese periodo, deben constar como residentes en España. Así, las personas que emigraron en el año previo pero estuvieron menos de doce meses en el extranjero han de ser contabilizadas en la población censal.

Análogamente se eliminarán de la población censal aquellas personas que, si bien están de alta en la fecha de referencia, han pasado menos de doce meses en España.

Además se comprueba que las defunciones y nacimientos están actualizados a la fecha de referencia censal.

Es decir, la base de población de los ficheros precensales, si bien parte esencialmente de las personas en alta en la base padronal a 1 de enero, se conforma añadiendo y eliminando determinados registros en función de sus movimientos padronales a lo largo del tiempo.

Una vez establecida esta base de población inicial, para la población extranjera se analiza si la inscripción del habitante ha alcanzado su fecha de caducidad, en caso de ser extranjero no comunitario sin autorización de residencia permanente (ENCSARP) o su fecha de comprobación de residencia en el resto de extranjeros

(“No ENCSARP”). Así clasificaremos a cada extranjero de la base precensal según tenga su inscripción caducada o no (“caducados” / “no caducados”).

Después se analizan 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.

- 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 sea 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 del Instituto Nacional de la Seguridad Social.

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...

Todos estos indicadores de presencia en los diferentes ficheros administrativos se calculan a nivel individual y también a nivel de hogar, de manera que se dispone de información sobre cuántos miembros del hogar han tenido actividad en la Agencia Tributaria, cuántos aparecen en el fichero de demandantes, cuántos aparecen como asegurados o beneficiarios de asistencia sanitaria, etc...

Con esta información, vamos a tener extranjeros que inicialmente están “caducados” y que por presentar signos suficientes de presencia son “rescatados” y contabilizados en la cifra de población censal. Y, análogamente, extranjeros que no han cumplido su fecha de caducidad pero no van a ser contabilizados en la población por carecer totalmente de signos de presencia en registros.

Además, tanto para españoles como para extranjeros, se tienen en cuenta los movimientos padronales en los meses inmediatamente posteriores a la fecha de referencia censal. Así, por ejemplo, hay determinados movimientos en los meses posteriores a la fecha de referencia que requieren una intervención directa de la persona o una comprobación de residencia por parte del ayuntamiento, que nos indican que una alta probabilidad de que la persona estaba residiendo en la fecha censal.

Análogamente, otros movimientos nos indican una alta probabilidad de que la persona no estuviera realmente residiendo en España en la fecha de referencia.

Para los menores de 16 años de edad, dado que, en general, no están presentes en registros administrativos, 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.

No obstante, como también se dispone de información sobre los alumnos matriculados en todas las enseñanzas oficiales, se está estudiando usar esa información también como signo de presencia.

La cifra de población estimada por este método a 1 de enero de 2018 asciende a **46.758.311** personas. De ellas, **4.686.520** son extranjeros (el 10,0%).

Si bien globalmente la cifra de población es muy similar a la del padrón, el FPC-18 contabiliza 83.000 españoles más que el Padrón, procedentes fundamentalmente de personas que no están de alta en la fecha de referencia pero son bajas de oficio en fechas anteriores que posteriormente “rebotan” y cuya presencia se comprueba a partir de registros administrativos.

	FPC18	Padrón 1-1-2018	Dif FPC-Padrón
Total	46.758.311	46.722.980	35.331
Españoles	42.071.791	41.988.289	83.502
Extranjeros	4.686.520	4.734.691	-48.171
% Extranjeros	10,02	10,13	-0,11

Próximos pasos y mejoras previstas

En el cálculo de la cifra de población del fichero precensal de 2016 no se pudieron incorporar como fuentes a tener en cuenta para los signos de presencia los datos tributarios de las Haciendas Forales de Álava, Bizkaia, Guipúzcoa y la Comunidad Foral de Navarra. En el FPC de 2018 solo faltaban los datos de Álava.

Se han estudiado pormenorizadamente tanto el fichero de demandantes de empleo (proporcionado por el Servicio Público de Empleo Estatal) como el de perceptores de prestaciones del Instituto Nacional de la Seguridad Social, concluyendo finalmente que la presencia de una persona, con ciertas salvedades, en estos ficheros puede tomarse como signo de presencia en el futuro para el proceso de cálculo de la población.

Además se ha ampliado la petición a la Agencia Tributaria (y Haciendas Forales), para que se incluyan las personas con retenciones por conceptos no contemplados hasta ahora (por ejemplo, prestaciones de incapacidad, renta mínima de inserción, rendimientos de actividades económicas, etc...).

En el caso de los menores de 16 años también se dispone de información sobre los alumnos matriculados en todas las enseñanzas oficiales. Se prevé usar esa información también como signo de presencia para evitar descontar población extranjera cuya inscripción pueda haber caducado.

7.3 EL FICHERO CENSAL DE POBLACIÓN: QUÉ VARIABLES INTERVIENEN

El punto de partida para el análisis de las variables a incorporar es el cuestionario del Censo de Población y Viviendas de 2011 y las obligaciones que impone el Reglamento Censal.

El Censo de Población 2011 contenía 22 preguntas en el cuestionario individual, que podemos agrupar en ocho bloques (datos demográficos básicos, movimientos migratorios anteriores, estado civil y composición del hogar, estudios, cuidados proporcionados, fecundidad, relación con la actividad, movilidad).

Además, el cuestionario de la vivienda contenía siete preguntas sobre instalaciones, características y equipamiento de la vivienda así como sobre el régimen de tenencia.

Las preguntas del cuestionario censal de 2011 se diseñaron con el objetivo, en primer lugar de cumplir con el Reglamento comunitario; además se pretendía dar continuidad a preguntas contenidas en censos anteriores; por último, se intentaba incorporar información demandada por los usuarios. Todo ello siempre con las limitaciones de un cuestionario dirigido a la población en su conjunto que, sobre todo, obliga a que el número de preguntas sea muy limitado.

El análisis completado hasta el momento de escribir este proyecto ha consistido en intentar replicar esas 22 preguntas; es decir, investigar hasta qué punto el censo basado en registros podría al menos contener el mismo detalle temático que el censo de 2011.

Como ya se ha apuntado, el primer ensayo del Censo de 2021, el FPC-2016, se centró en construir el censo de población, es decir, en el *cuestionario individual*. Como primera conclusión del análisis detallado de variables cabe decir que, respecto de la población, puede obtenerse toda la información a la que obliga el Reglamento UE a partir de registros administrativos. Esta información se tiene con una calidad, entendida ésta como número de personas sobre las que se dispone de datos, que sería, como mínimo, similar a la que se tendría en un censo exhaustivo.

Hay otras variables que no son obligatorias pero tienen una gran tradición censal o bien un interés especial entre la comunidad de usuarios del censo; también se ha estudiado su posible obtención y en casi todos los casos es posible hacerlo.

Quedan algunas preguntas, no recogidas en el Reglamento UE, que no pueden encontrarse en registros administrativos. Del cuestionario individual de 2011 parece que la única pregunta que está claramente en ese grupo es la pregunta 12: *“Durante la semana pasada ¿realizó alguna de las siguientes tareas sin recibir dinero a cambio?”*. Las opciones de respuesta eran *“cuidado de menores”*, *“cuidado de personas con problemas de salud”* *“tareas benéficas o de voluntariado”*, etcétera).

Existe, por último, un conjunto de preguntas que no figuraban en el censo 2011 pero que un planteamiento censal como éste de 2021 permite incorporar. A esta cuestión se dedica un capítulo específico en este documento (ver apartado 10).

El análisis de calidad por variables consiste en determinar hasta qué punto se pueden obtener esas preguntas del cuestionario individual utilizando exclusivamente registros administrativos, qué procedimientos se siguen para construir la información,

dónde están las carencias y cuáles son las mejoras que se pueden acometer durante los próximos años con el fin de llegar a 2021 con esas preguntas completamente resueltas a partir de los registros administrativos.

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

- Requisitos
 - o Reglamento UE: Especifica el requisito mínimo para el proyecto
 - o 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: para cada variable se analiza con dos perspectivas:
 - o 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 (se habla de fuente ganadora: la fuente a partir de la cual se asigna a cada persona el valor en concreto para esa variable)
 - o En los resultados: Se cotejan los resultados obtenidos con fuentes de referencia (según el caso son censos anteriores, EPA, Encuesta Continua de Hogares u otras).
- Próximos pasos y mejoras previstas. Dado que el proyecto censal se encuentra en un proceso en continua evolución, en el momento de redactar este proyecto están pendientes diversas mejoras de cara a 2021 pero que aún no se han podido poner en práctica por diversas razones.

7.3.1 Variables demográficas (sexo, edad, nacionalidad, lugar de nacimiento)

En el cuestionario del censo 2011 figuraban estas preguntas:

Cuestionario Individual de la Persona 1

Escriba los siguientes datos para la persona que aparece en primer lugar (Persona nº **1**) en la **Lista de personas**.

Nombre y apellidos:

Fecha de nacimiento:
día mes año

Sexo: Hombre Mujer

País de nacimiento:

España. *Escriba municipio y provincia:*
 Municipio:

 Provincia:
 Otro país:

¿Cuál es su nacionalidad?

Si tiene doble nacionalidad, española y otra, marque ambas opciones y escriba el país correspondiente.
Si tiene doble nacionalidad, pero ninguna es la española, escriba únicamente una de ellas.

Española
 De otro país:

Estas son las variables más sencillas de obtener. El padrón es la fuente para todas estas variables básicas siendo únicamente necesarias pequeñas depuraciones en el país de nacimiento y nacionalidad. De hecho el padrón contiene únicamente las variables sexo, edad, nacionalidad, lugar y fecha de nacimiento, además de los datos de identificación (DNI, NIE o número de pasaporte, necesarias para efectuar los enlaces con otros ficheros).

7.3.2 Variables relacionadas con las migraciones

a) Requisitos

- Reglamento UE:

Además de las variables básicas de cada individuo, el Reglamento exige las siguientes variables relativas a migraciones:

- **Año de llegada a España:**

El año de llegada será el año en que la persona estableció su residencia habitual en el país por última vez (en caso de que lo haya hecho en más de una ocasión). Se declarará el año de la última llegada al país en vez del año de la primera llegada (es decir, el tema «año de llegada al país» no tiene en cuenta las estancias interrumpidas) y se proporcionará con la siguiente desagregación:

- 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 sólo 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 construían 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

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 un año antes del censo” se amplía el número de categorías exigidas en el reglamento a las siguientes:

- 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) de la misma comunidad
 - Distinta comunidad
 - Fuera del país
- No consta.
- No aplicable (niños de menos de 1 año)

En principio, solo se tendrán en cuenta aquellas migraciones interiores producidas entre distintos municipios, a diferencia del Censo 2011, que contabilizaba también los cambios de domicilio dentro del mismo municipio. Sin embargo, se está trabajando en la posibilidad de incluir también los cambios de residencia dentro del mismo municipio, lo que permitiría ampliar las categorías de la variable “Lugar de residencia habitual un año antes del censo” desagregando la categoría:

- Misma región NUTS3 (provincia/isla)

En las categorías:

- Mismo municipio
- Distinto municipio de la misma región NUTS3 (provincia/isla)

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:

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

Los años de llegada a cada región se proporcionarán desglosados año a año.

En caso de incluir los cambios de residencia intramunicipales cabría la posibilidad de plantearse incluir el año de llegada a la vivienda, como se hacía en el Censo 2011.

c) Método y fuentes

Aunque el padrón es muy insuficiente a la hora de proporcionar las variables censales, implícitamente contiene la información necesaria para construir algunas otras variables más allá de las básicas de clasificación del apartado anterior. En concreto, las variables migratorias se pueden obtener casi directamente de la base padronal a partir de los lugares de residencia anteriores, derivados de una explotación longitudinal de la información. También sirve como punto de partida para la composición del hogar como se verá más adelante.

El padrón continuo inicia su historia en el año 1996. En su instante inicial 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.

Durante el año 2015 se completó ya el enlace con el censo de 2001, que fue un censo exhaustivo, de forma que, siempre que se consiga enlazar a un individuo con su cuestionario de 2001 (lo que ocurre para un porcentaje muy elevado de la población) se puede obtener información sobre su historia residencial anterior a 1996.

En cuanto a la variable “País de procedencia”, serán necesarias pequeñas depuraciones similares a las de las variables país de nacimiento y nacionalidad. Está por decidir si se realizará una imputación para aquellos casos en que esté sin cumplimentar.

Las variables relativas a los años de llegada a las distintas regiones pueden obtenerse directamente de la base de datos de Padrón, pero como ya se ha mencionado, dado que el Padrón Continuo empezó a funcionar en 1996, no hay datos anteriores a esa fecha. Por tanto, todos los registros que tienen una fecha de llegada al municipio, provincia, comunidad autónoma o país que coincide con la fecha de inicio del Padrón Continuo, se intentan enlazar con el Censo de 2001, y de ahí se obtendrá también su región de residencia de procedencia.

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.

La variable “Lugar de residencia habitual un año antes del censo” se puede obtener totalmente de la base padronal.

En esencia, los trabajos de diseño de estas variables los podemos considerar como ya finalizados, a falta de posibles detalles a mejorar, por lo que este grupo de preguntas se pueden incorporar con calidad contrastada al fichero censal.

d) Análisis de la calidad de la variable

En el FPC-2018, el porcentaje de registros según la fuente ganadora utilizada para asignar valores a las variables obligatorias “Año de llegada a España” y “Lugar de residencia habitual un año antes del censo” se distribuye como sigue:

Fuente ganadora	Año de llegada a España	Lugar de residencia habitual un año antes del censo
Padrón	89,1%	100,0%
Censo 2001	6,9%	0,0%
Imputación	4,0%	0,0%
	100,0%	100,0%

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.

e) Próximos pasos y mejoras previstas

El diseño de estas variables se considera esencialmente resuelto de cara al censo de 2021 y actualmente están en fase de desarrollo. Como posibles aspectos de mejora se contemplan:

- Tratar de mejorar el enlace de la población con el Censo 2001 para recuperar información en más casos.
- Posibles mejoras en los métodos de imputación.
- Añadir información sobre cambios de residencia dentro del mismo municipio.

7.3.3 Estado civil

a) Requisitos

- Reglamento UE

Se ha de proporcionar el estado civil legal de la persona en la fecha de referencia.

Las categorías obligatorias para esta variable son:

- Soltero/a
- Casado/a o en unión registrada
 - Entre personas de distinto sexo
 - Entre personas del mismo sexo
- Viudo/a
- Divorciado/a o separado/a legalmente

El Reglamento exige esta desagregación como máximo a nacional (por sexo y edad simple). A niveles inferiores no se exige determinar el sexo de los cónyuges.

i. Censos anteriores

En el anterior censo de 2011 el Reglamento europeo (CE) No 1201/2009 exigía diferenciar las categorías divorciado/a y separado/a legalmente.

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

5 ¿Cuál es el estado civil de la Persona 1?

Marque su situación legal, aunque no coincida con la de hecho

<input type="checkbox"/> Soltero/a	<input type="checkbox"/> Separado/a	<input type="checkbox"/> Viudo/a
<input type="checkbox"/> Casado/a	<input type="checkbox"/> Divorciado/a	

b) Planteamiento propuesto para la variable y desglose

A la vista de la información administrativa disponible, las categorías establecidas para la variable Estado civil en el Censo de 2021 coinciden con las exigidas en el Reglamento:

- Soltero/a
- Casado/a o en unión registrada
 - Entre personas de distinto sexo
 - Entre personas del mismo sexo
- Viudo/a
- Divorciado/a o separado/a legalmente

c) Método y fuentes

Dado que el estado civil de la persona no figura recogido en Padrón y que no existe una única fuente que proporcione información de toda la población en su conjunto, la estimación de esta variable se lleva a cabo utilizando diferentes fuentes administrativas:

- Fichero proporcionado por el INSS con las personas que han percibido una pensión de viudedad.
- Estado civil declarado en el modelo 100 del Impuesto sobre la Renta de las Personas Físicas del ejercicio anterior, proporcionado por la Agencia Tributaria y las Haciendas Forales.
- Boletines de matrimonios del Movimiento Natural de Población del INE (desde el año 2005).
- Ficheros de separaciones y divorcios de INFOREG (Registros civiles informatizados) desde 2002.
- Estado civil de los extranjeros inscritos en el Registro Central de Extranjeros
- Estado civil de la persona en los censos de población de 2001 y 2011
- Estado civil de la persona en el modelo 100 del IRPF de ejercicios anteriores.

Teniendo en cuenta que una persona puede figurar en varios registros con un estado civil distinto, se ha establecido un **orden de prioridad** para determinar la “fuente ganadora”

- Consideramos que el fichero de prestaciones de viudedad del INSS (para detectar viudos) y el estado civil que figura en la declaración del IRPF del año anterior tienen la máxima calidad por tratarse de registros de temática económica.
- Tras utilizar estas dos fuentes, las personas que aún no tienen un estado civil asignado se buscan en las tablas de casados construidas a partir de los boletines de matrimonios de MNP y de las personas que figuraban como casadas en los dos censos anteriores y en los ficheros tributarios de ejercicios anteriores. Si estas personas en la fecha de referencia comparten la misma vivienda se las considera casadas, a no ser que figuren en los ficheros de separaciones y divorcios de INFOREG con un año de divorcio posterior o igual al año de matrimonio, en cuyo caso se considerarían divorciadas.
- Las personas que siguen sin tener un estado civil se buscan en el Registro Central de Extranjeros, que resulta especialmente útil para detectar matrimonios de extranjeros fuera de España.
- Por último, para aquellas personas cuyo estado civil no se ha podido determinar en las fuentes anteriores, se aplican una serie de reglas de imputación determinística que tienen en cuenta la edad, si la persona vive sola y el estado civil de los censos anteriores, de modo que :
 - Si la persona tiene menos de 30 años y no tenemos información en ninguna fuente la consideramos soltera.
 - Si la persona es mayor de 60, vive sola y no tenemos información en ninguna otra fuente, le asignamos el estado civil soltero, viudo o divorciado, que tuviera en el último censo.

d) Análisis de la calidad de la variable

- Análisis de calidad en la asignación de la variable

En el ensayo FPC-2016 el número de registros según la fuente ganadora utilizada para asignar la variable se distribuye así:

Fuente ganadora final:

		Soltero	Casado	Viudo	Divorciado / Separado	Total
Fuente ganadora final	AEAT y HF 2018	31,5%	79,3%	15,3%	71,3%	52,7%
	PRESTACIONES	0,0%	0,0%	73,7%	0,0%	5,0%
	MNP y AEAT anteriores	7,6%	17,9%	7,9%	19,9%	12,6%
	INFOREG	0,0%	0,0%	0,0%	5,9%	0,3%
	RCE	6,6%	2,8%	0,9%	2,6%	4,4%
	Imputación	54,3%	0,0%	2,3%	0,3%	25,0%
	Total asignados	100,0%	100,0%	100,0%	100,0%	100,0%

La fuente principal para estimar el estado civil es el fichero del IRPF de la AEAT y Haciendas Forales, del que se obtiene el 53% de las estimaciones. Cada categoría de estado civil se capta, mayoritariamente, en alguna fuente concreta: por ejemplo, el 71% de los divorciados y el 79% de los casados proceden, como fuente principal, de los ficheros tributarios. Los datos sobre pensiones de viudedad proporcionados por el INSS permiten detectar el 74% de los viudos.

Para asignar el estado civil "soltero", una vez contrastada la buena cobertura del resto de estados civiles en los distintos ficheros administrativos, si encontramos que la persona es menor de 30 años y no figura en ningún registro como casada, viuda o divorciada, consideramos razonable imputarla como soltera, por exclusión.

Se tienen 2.536.177 registros, el **5,4%** de la población, a los que en el FPC-2018 no se ha asignado un estado civil a partir de las fuentes disponibles. Hay que tener en cuenta que no se han podido integrar los ficheros tributarios de las Haciendas Forales de Álava ni de Navarra y esta es la fuente principal para estimar la variable estado civil.

En cuanto a la distribución territorial de los registros sin información de estado civil hay que destacar que, al no disponer aún de información tributaria de Álava se tiene un 27% de registros sin estado civil asignado en esta provincia. En el resto del territorio, los registros sin estado civil se concentran en provincias con un alto número de extranjeros como Santa Cruz de Tenerife, Las Palmas, Alicante, Málaga y en las provincias de Ourense y Pontevedra (en torno al 6-8% todas ellas). En las demás, apenas se supera el 4%.

Por otra parte, se observa que para el 11,0% de los extranjeros, 517.712, no se dispone de estado civil, frente al 4,8% de españoles. Cabe destacar que a un 15% de los ingleses y alemanes no se les puede asignar estado civil a partir de registros y se deberán utilizar métodos de imputación.

- Análisis de calidad de los resultados obtenidos

La distribución del estado civil del fichero precensal de 2016 no ofrece diferencias significativas con ninguna de las fuentes con las que se ha comparado:

- Encuesta Continua de Hogares centrada a 1-1-2018 (muestra acumulada de dos años de 100.542 hogares).
- Censo de Población de 2011
- Encuesta de Población Activa centrada a 1 de Enero de 2018 (población de 16 años o más)

e) Próximos pasos y mejoras previstas

Para aquellos registros cuyo estado civil siga sin poder recogerse en ninguna fuente, se está desarrollando un modelo de imputación probabilística basado en las variables con un mayor poder explicativo del estado civil como son la edad, sexo, nacionalidad y composición del hogar y las probabilidades de transición entre los diferentes estados civiles.

Además, no se ha explorado aún la información contenida en los registros de parejas de hecho, que se sumarían a la población casada, según la clasificación censal.

7.3.4 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 Sí, 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	→ N° 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

También se clasifica por “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).

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, por las personas inscritas en la misma hoja padronal.

No obstante, dado que se dispone del fichero precensal de viviendas (más información en apartado 8.2) 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 en el apartado 8, 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 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 padre o madre:

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 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.

Por otra parte, se dispone de los Boletines de Partos del Movimiento Natural de Población, disponibles desde 2002 a 2018, en los que figuran datos identificativos del padre y madre de cada nacido. Se está trabajando en asignar a cada nacido el padre y madre que figuran en estos boletines.

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 y los localizados en los ficheros que la Agencia Tributaria proporciona al INE, quedándonos sólo con el último cónyuge registrado para cada persona. Estas parejas potenciales se dan por válidas en el caso de que continúen compartiendo vivienda en la base de datos padronal.

Tras estos pasos, es evidente que hay aun parejas no casadas que conviven y 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 2016 como con los hogares del Censo 2011, 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%. No obstante, para parejas del mismo sexo, el modelo no funciona tan bien y es necesario continuar la investigación.

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	11.386.750	100,00%	14.231.668	100,00%
Censos anteriores	5.417.755	47,58%	6.968.996	48,97%
Nombre del padre (D.G Policía y MNP)	5.767.263	50,65%	7.140.087	50,17%
Hermanos	4.411	0,04%	4.893	0,03%
Imputado	120.542	1,06%	117.692	0,83%

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	20.125.088	100,00%
Hijo/s en común	12.811.435	63,66%
Censos pasados, MNP y AEAT	5.861.415	29,12%
Imputación	1.452.238	7,22%

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 7,22% 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:

El 61% de las parejas del mismo sexo se han construido a partir del modelo de imputación. En cuanto a las parejas de hecho, el vínculo es muy difícil de determinar a partir de registros (tan sólo se puede utilizar la información de censos pasados o la existencia de hijos en común), por lo que el 44,5% proceden del modelo de imputación.

Fuente ganadora	Parejas de hecho	%	Parejas de derecho	%
Total	1.294.130	100,00%	8.768.414	100,00%
Hijo/s en común	642.820	49,67%	5.762.898	65,72%
Censos pasados, MNP y AEAT	74.855	5,78%	2.855.853	32,57%
Imputación	576.455	44,54%	149.664	1,71%

La distribución de la variable “Estructura de hogar” del fichero precensal de 2016 comparada con la de la Encuesta Continua de Hogares centrada a 1-1-2016 y la del Censo de 2011 muestra que las categorías “Otro tipo de hogar” y “Pareja o padre/madre con hijos y otras personas” están sobreestimadas mientras que el número de “Parejas sin hijos” y “Parejas con hijos menores de 25 años” es inferior al obtenido en las otras dos fuentes mencionadas.

Esta comparación confirma que en algunos tipos de hogar faltan por determinar vínculos, especialmente los conyugales, entre sus miembros.

e) Próximos pasos y mejoras previstas

Respecto a la detección de progenitores, aunque está casi resuelta, se va a añadir a la información de censos anteriores y de los boletines de Partos, los datos de ascendientes y descendientes a cargo consignados en el IRPF.

Se está trabajando en la explotación de los últimos ficheros proporcionados por la Agencia Tributaria, que incluyen como novedad que en una misma declaración del IRPF se permite consignar hasta cuatro progenitores diferentes para los hijos, es decir, se pueden identificar familias en las que no todos los hijos son del mismo padre o madre, enriqueciendo así la tipología de hogares detectados.

En el caso de los extranjeros, se va a incorporar información del Registro Central de Extranjeros relativa a las autorizaciones de residencia a extranjeros por motivos familiares (por ejemplo, reagrupación familiar, familiares de ciudadanos comunitarios, ...). En ellas figura la identificación 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 debe evaluar la mejora en la detección de vínculos familiares que se obtiene explotando esta información.

También se están diseñando rutinas específicas para la búsqueda de padres y madres que tengan en cuenta la nacionalidad, ya que el procedimiento de asignación del primer y segundo apellido de los niños es diferente en cada país. Para los españoles hay que incorporar a los algoritmos de búsqueda los últimos cambios legislativos, que permiten elegir el orden de los apellidos de los hijos.

En cuanto a la detección de parejas, se incorporarán los datos de cónyuges de las Haciendas Forales de País Vasco y la Comunidad Foral de Navarra. Esto, junto con la acumulación de los datos tributarios de ejercicios de renta siguientes de la AEAT y de los boletines de matrimonio de MNP incrementará anualmente la base de datos de potenciales parejas, facilitando la detección de cónyuges.

No obstante, el principal reto es tratar de paliar la falta información en registros para detectar parejas de hecho sin hijos. En la línea de otros países con censos basados

en registros, se seguirá trabajando en procesos de imputación sencillos, calibrando los resultados a los de las encuestas a hogares y haciendo especial esfuerzo en detectar lo mejor posible las parejas del mismo sexo. En este sentido se aprovechará al máximo la información procedente de la ECEPOV.

Cabe destacar, como se comentará en el apartado 12, que en la difusión de microdatos del Censo 2021 está previsto proporcionar indicadores de calidad a nivel de registro y variable, de manera que habrá una fácil identificación de los hogares cuyas relaciones se han construido a partir de imputación. Es decir, si un hogar es clasificado como pareja sin hijos, el usuario especializado va a saber si esta información ha salido de algún registro administrativo o del modelo de imputación.

7.3.5 Estudios en curso y nivel de estudios alcanzado

a) Requisitos

i. 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

ii. 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)
- Para personas de 16 años o más: el máximo nivel de estudios alcanzado (pregunta 10), el campo al que pertenecen dichos estudios (pregunta 11) e información sobre los estudios en curso (pregunta 18).

- Nivel CINE 3 – Educación secundaria alta
- Nivel CINE 4 – Educación postsecundaria no terciaria Educación terciaria
- Nivel CINE 5 – Educación terciaria de ciclo corto
- Nivel CINE 6 – Grado en educación terciaria o nivel equivalente
- Nivel CINE 7 – Nivel de maestría, especialización o equivalente
- Nivel CINE 8 – Nivel de doctorado o equivalente

Para las personas que estén cursando algún estudio, en los niveles más altos, se proporcionará información también de la naturaleza de los estudios.

c) Método y fuentes

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:

- Información sobre nivel educativo contenida en Padrón a 1 de enero de cada año (CNES).
- Registro histórico de Títulos del Ministerio de Educación y Formación Profesional (MEyFP)
- Registro histórico de graduados en enseñanzas no universitarias del MEyFP
- Registro histórico de egresados en universidades del Consejo de Universidades
- Registro de certificados de profesionalidad del Servicio Público de Empleo Estatal (SEPE)
- Registro de demandantes de empleo del SEPE.
- Información de los censos previos de 2011 y 2001.
- Alumnos matriculados en enseñanzas no universitarias del MEyFP. Cursos 2014-2015 en adelante
- Registro histórico de matrículas en enseñanzas universitarias del Consejo de Universidades.

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

El padrón es la fuente de partida que garantiza disponer para toda la población de un valor de nivel educativo alcanzado pero esa información es, en principio, 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:

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

No obstante, no toda la población figura con ese nivel de detalle. Por diversas razones, sobre todo en empadronamientos antiguos, figuran solo cuatro categorías. Aproximadamente en un 60% de la población se dispone del nivel detallado y en el restante 40% solo con esos cuatro valores posibles de nivel de estudios. Pero además, esta información queda obsoleta salvo que la persona cambie su empadronamiento y rellene una nueva hoja padronal. No obstante, el Padrón se cruza con el registro de Títulos del MEYFP, corrigiendo la información de partida en muchos casos.

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.

Para aquellos casos donde no se encuentra información, o son muy jóvenes o se tiene una primera aproximación en forma de intervalo, se cruza con la información de matrículas. Para cada persona se toma su matrícula más reciente a la fecha del FPC y la de mayor nivel CINE, de tal modo que en algunos casos, se puede imputar el nivel educativo alcanzado por ser un requisito para la inscripción en determinados cursos.

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.

Para la asignación de los estudios en curso, para cada persona, se selecciona la matrícula de mayor nivel CINE y más cercana a la fecha de referencia (1 de enero de 2016). Por ejemplo: si una persona tiene una matrícula de doctorado y otra de grado se tomaría la información del doctorado.

Para las personas que cursen dobles titulaciones universitarias se tomará únicamente una de ellas, puesto que no existe un criterio internacional ni nacional

para la elección entre una de ellas ya que conducen a titulaciones de igual categoría. No obstante el total de dobles grados recibidos no supera apenas los 20.000 casos.

d) Análisis de la calidad de la variable

La estimación del nivel educativo 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.

Fuente ganadora para nivel educativo en FPC-2016 (15 años y más):

Fuente ganadora	Registros	%
Total de registros	39.599.330	100,0%
Censo 2001	15.808.825	39,9%
Nivel de Padrón	10.410.577	26,3%
Nivel de Padrón cruzado con títulos	6.662.921	16,8%
Demandantes del SEPE	3.048.314	7,7%
Egresados universitarios	1.259.012	3,2%
Matrículas del MECD	1.043.959	2,6%
Censo 2011	656.007	1,7%
Graduados del MECD	361.370	0,9%
Títulos del MECD	238.163	0,6%
Valor imputado	78.789	0,2%
Registro Central de Extranjeros	20.200	0,1%
Matrículas de Universidades	11.193	0,0%

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 ésta.

El hecho de que el registro de Títulos del MEyFP proporcione información de manera residual se explica porque ya Padrón cruza con esta fuente, e incorpora un nivel corregido que sí aparece en tercer lugar como fuente ganadora.

Solo el 0,2% de registros resultan imputados mediante las distribuciones observadas en el Censo 2011.

Fuente ganadora para estudios en curso:

Fuente ganadora	Registros	%
Total matriculados	3.481.373	100,0%
Matriculas de Universidades	1.404.502	40,4%
Matriculas del MEyFP	2.076.871	59,6%

Comparando con otras fuentes (Censo 2011, EPA), se observa que el FPC-2016 presenta resultados muy similares. Existe todavía una sobreestimación en las categorías bajas de estudios, lo que es debido a la ausencia de registros administrativos para el grupo de personas muy jóvenes que nos permitan afirmar cual

es el nivel real. Esta carencia de datos se debe solucionar para posteriores ediciones del FPC.

e) Próximos pasos y mejoras previstas

Se han hecho 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. No obstante, la cobertura de estos registros no es completa aun y se sigue trabajando de forma conjunta con el MEyFP para su mejora.

La cobertura por enseñanzas tampoco es completa, en el sentido de que algunas comunidades aún tienen que incorporar estudios de FP a sus bases de datos.

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.

La estimación del nivel educativo es más difícil para determinados tipos de población, como son los extranjeros o aquellos más jóvenes cuya edad no alcanza la mínima para obtener un título oficial y no están en los ficheros de matrículas. En este sentido, se va a incorporar una nueva fuente de datos procedente del SEPE que proporciona el nivel educativo de los contratados (y del puesto de trabajo) para todos los contratos formalizados desde el 2001 hasta hoy, que se supone que va a mejorar la estimación del nivel educativo especialmente en la población extranjera.

Se está trabajando en corregir un sesgo a la baja en niveles educativos bajos, ya que para determinadas personas jóvenes se puede estar subestimando su nivel educativo por ausencia de registros educativos, lo cual es normal debido a que las fuentes no recogen todas las matrículas existentes en el sistema y el primer título oficial educativo que obtiene una persona es el de Educación Secundaria.

Con las mejoras previstas en la cobertura de ficheros de matriculados en enseñanzas no universitarias se espera disponer de una información bastante completa para construir estas variables.

7.3.6 Relación con la actividad económica

a) Requisitos

i. 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”).

ii. Censos anteriores

En el anterior censo de 2011 el reglamento europeo (CE) No. 1201/2009 exigía desglosar a las personas desempleadas según si habían tenido un empleo previo o no, y se incluía 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 15)

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 autoclasificación del informante, con una pregunta como que figura nº 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 proponen estas siete categorías, es decir, las del Reglamento más una, que se daría para cualquier desagregación (municipal o inferior):

- Población económicamente activa
 - Empleado
 - Desempleado
- Población económicamente inactiva:
 - Jubilado, prejubilado, pensionista o rentista
 - Estudiante
 - Persona con 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.

De momento, 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 (curso 2015/2016) como no universitarias (curso 2014/2015).
- Información tributaria del modelo 190 relativa al ejercicio económico de 2015 proporcionada por la Agencia Tributaria.

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.

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. Por ejemplo:

- Si la persona se encuentra como empleada en alguna fuente, se la considera empleada.
- Si la persona está desempleada, matriculada en algún curso, y cobra una prestación por invalidez, prevalece la situación de paro.
- Si la persona percibe varias prestaciones, la de jubilación prevalece frente al resto.
- 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 inactiva, estudiante.

d) Análisis de la calidad de la variable

Con datos del FPC-2016, 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:

Fuente ganadora final:

Fuente ganadora	Registros	%
Total personas 15 años y más	39.599.330	100,0%
TGSS	15.973.022	40,3%
INSS	8.351.855	21,1%
Sin información	4.131.100	10,4%
SEPE	4.105.434	10,4%
Fichero matrículas MECD	1.762.893	4,5%
Imputación determinística	1.540.904	3,9%
AEAT	1.483.625	3,7%
Fichero matrículas universitarias	853.757	2,2%
MUFACE	614.385	1,6%
CLASES PASIVAS	418.353	1,1%
ISFAS	316.223	0,8%
MUGEJU	47.779	0,1%

Se tienen, por tanto, 4.131.100 no asignados que se reparten así por sexo y grupo de edad (quinquenal):

Grupo de edad \ Sexo	TOTAL	HOMBRES	MUJERES
Total registros sin relación asignada	4.131.100	1.534.547	2.596.553
De 16 a 19 años	476.262	237.201	239.061
De 20 a 24 años	357.411	179.891	177.520
De 25 a 29 años	379.085	178.011	201.074
De 30 a 34 años	383.159	174.302	208.857
De 35 a 39 años	422.560	181.717	240.843
De 40 a 44 años	433.233	168.083	265.150
De 45 a 49 años	412.607	138.712	273.895
De 50 a 54 años	404.821	112.987	291.834
De 55 a 59 años	407.950	87.608	320.342
De 60 a 64 años	454.012	76.035	377.977

El porcentaje de no asignados se sitúa en el 10,4% de la población de 15 y más para el FPC 2016. 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.,

Provisionalmente 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 1-1-2016. Se ofrece una comparativa de censos anteriores y del resultado provisional del FPC2016 con la EPA de ese momento.

Comparativa de estimaciones de EPA y Censos y Diferencias en Tasas de Variación. Años 2001, 2011 y prueba 2016.

	EPA 4º trim 2001	Censos 2001	Dif	EPA 4º trim 2011	Censos 2011	Dif	EPA(1-1-2016)	FPC2016	Dif
Ocupados	16.348,2	16.329,7	-0,1%	17.807,5	17.514,6	-1,6%	18.118,4	18.056,1	-0,3%
Parados	2.283,9	2.692,8	17,9%	5.716,8	7.377,3	29,0%	4.749,1	4.198,3	-11,6%
Inactivos	15.931,6	15.201,4	-4,6%	15.427,0	14.191,8	-8,0%	15.626,0	16.893,6	8,1%

El sesgo al alza por la autoclasificación 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: una cierta sobreestimación de inactivos e infraestimación de parados. Las diferencias se concentran en el colectivo de personas jóvenes y en extranjeros.

El método seguido hasta hoy en el FPC-2016 solo considera como parados a los demandantes de empleo registrados por lo que no sigue del todo la recomendación internacional de intentar medir "paro OIT"

Todos los registros que aún faltan por enlazar han sido provisionalmente considerados como inactivos llegando a un 7,8% de diferencia. No obstante con la mejora en cobertura de registros, se prevé encontrar a parte de ellos en alguna otra situación administrativa de actividad, reduciendo en un futuro las diferencias de parados e inactivos con la EPA.

No obstante esta es una limitación inherente a un censo basado en registros administrativos que solo puede subsanarse mediante imputación. La condición de

parado se obtiene principalmente de encontrar a la persona registrada como tal 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 ajustado a lo que nos da una encuesta como la EPA al que recoge un censo basado en cuestionarios que clasifica esa variable usando una única pregunta (la denominada "autoclasiicación"), como puede verse en la tabla anterior.

e) Próximos pasos y mejoras previstas

En el momento de redactar este documento no se ha completado el análisis del FPC-2018 para esta variable. Los datos solicitados a distintas instituciones para el FPC2016 fueron stocks relativos a 1 de enero de 2016, lo que no permitía obtener, estrictamente hablando, información precisa sobre la semana de referencia (la última del año). Esto además provocó el problema de encontrar personas cuyo contrato terminaba el 31 de diciembre de 2015 y que ya no figuraban, por tanto, como ocupados, pero tampoco lo estaban en el fichero de demandantes de empleo a 1-1-2016. Esta limitación se subsanó para siguientes ediciones del FPC dado que se solicitaron también datos de ocupados y demandantes de empleo para la semana de referencia (la última del año). En las pruebas realizadas hasta hoy con el FPC2018 parece que se mejora la estimación de parados e inactivos notablemente.

Otra limitación del fichero de 2016 es que no se disponía de información exhaustiva aún de alumnos matriculados en todas las enseñanzas, lo que hace que para edades 16-20 se tenga un porcentaje aún demasiado elevado de no asignados. Se ha ido mejorando la información de cobertura de personas en el sistema educativo sustancialmente durante 2018 y 2019.

7.3.7 Variables relativas a los ocupados

a) Requisitos

i. 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 CNO11 a un dígito:

- 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 NACE rev. 2 (21 categorías).

La situación en el empleo tiene las siguientes categorías obligatorias:

- Trabajador por cuenta ajena
- Trabajador por cuenta propia
 - Emplea personal
 - No emplea personal
- 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:

- 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)

ii. 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:*
 [Grid for occupation details]

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 [Grid] *Si no ha encontrado su actividad o tiene dudas, escribala a continuación:*
 [Grid for activity details]

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: [Grid]
 Provincia: [Grid]
 En otro país: [Grid]
 Código postal de su centro de trabajo o de estudios: [Grid]

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 propone llegar a una desagregación de la ocupación como mínimo a dos dígitos de la CNO2011, aunque intentando llegar a tres dígitos (lo cual parece factible en muchos casos), sector de actividad a tres dígitos de la CNAE-09 (son 272 categorías: mucho más detallado de lo exigido), ubicación del lugar de trabajo a nivel de provincia y, para la variable situación profesional, proporcionar la información obligatoria incluyendo las subcategorías de empresarios según si emplean o no personal.

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.

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 una adicional:

- Trabajadores que cotizan a la Seguridad Social que estaban de alta (Tesorería General de la Seguridad Social).
- 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.
- Información relativa al último contrato vigente de cada persona a 1 de enero de 2016 en el SEPE.
- Información de censos previos 2011 y 2001.

Se distinguen dos situaciones: si la persona es funcionaria o si está afiliada a la Seguridad Social.

Si bien un funcionario, podría estar dado alta ejerciendo otra actividad económica residual, se toma como empleo principal el recogido en el fichero de su mutualidad, prescindiendo de información suya que figure en la Tesorería o el fichero de contratos del SEPE.

El fichero de contratos del SEPE es realmente la principal fuente para determinar la ocupación. Para ello 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, por tanto datos de más de 23 millones de 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 del SEPE contiene la ocupación a tres dígitos de la CNO-11.

Este fichero tiene una limitación de cara a su uso para determinar la ocupación actual de las personas. Naturalmente, para aquellas personas que firmaran contratos hace ya varios o incluso muchos 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. Esto tiene implicaciones que se analizan más adelante.

En caso de no encontrarse información, se acude a información de los censos previos.

En el caso de los funcionarios mutualistas, por tanto, se acude únicamente a la información de los ficheros de MUFACE, MUGEJU o ISFAS que corresponda para estimar estas cuatro variables.

La estimación de la actividad económica según la CNAE-09 viene dada por información contenida en los propios ficheros. Se dispone de la ubicación del puesto de destino a nivel de provincia y la ocupación se estima en función del cuerpo de

pertenencia dentro de la Administración. La situación profesional se asigna de manera directa como “empleados por cuenta ajena”. En el caso de funcionarios de ISFAS se dispone de información de la CNO-11 en el fichero para la mayoría de casos salvo los funcionarios civiles.

Para las personas empleadas según la Tesorería la estimación de estas cuatro variables se realiza con la información de la Tesorería, del fichero de contratos del SEPE y censos previos.

La actividad económica de los afiliados viene recogida en el fichero de altas de la Seguridad Social, 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 y la situación profesional viene marcada por el régimen de cotización, tipo de contrato y tipo de relación laboral y código de Cuenta de Cotización.

d) Análisis de la calidad de las variables

En principio, la estimación de algunas de estas variables no supone mucha dificultad. La variable más complicada de todas es la ocupación.

La distribución de fuentes ganadoras para actividad económica, lugar de trabajo y situación profesional es idéntica a la distribución de la relación con la actividad económica.

La ocupación se asigna de forma ligeramente diferente a las variables anteriores ya que para algunas personas encontradas como empleadas en la Tesorería no se encontró información en el fichero de contratos del SEPE. Se encuentra el 94,3% de los casos, aunque se espera que la distribución de fuentes y calidad mejore. De hecho, en las primeras pruebas con FPC-2017 y FPC-2018 los enlaces con los ficheros de contratos del SEPE han mejorado.

En cuanto a los resultados que se obtienen, comparando con los datos de la EPA y el censo de 2011 a dos dígitos se observa que, en general, el número de ocupados asignado por el método es consistente con esas fuentes de forma muy razonable en muchos casos. No obstante se observan sesgos que pueden corregirse de cara a desarrollos futuros.

En cuanto a la rama de actividad, si comparamos con los datos de la EPA se observa bastante consistencia con los datos del FPC-2016.

En cuanto a la situación profesional la comparativa con la EPA produce resultados muy similares. La mayor diferencia aparece en el colectivo de asalariados. La EPA realiza una corrección de falsos autónomos, de modo que finalmente son considerados como asalariados. De momento en los ficheros precensales no se ha llevado a cabo esta corrección, considerando que si una persona aparece registrada en la Seguridad Social como autónomo, es autónomo a efectos censales.

Para la variable de lugar de trabajo se tabula el porcentaje de personas que residen en cada provincia y no trabajan en ella y se comparan con los resultados del Censo 2011. Esta comparación no es directa, 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.

e) Próximos pasos y mejoras previstas

La variable ocupación es quizás la más compleja de este grupo, dado el posible sesgo a la baja del registro de contratos, en el cual parece claro que están sobrerrepresentados los puestos de ocupación elemental en detrimento de altos puestos. La antigüedad de los contratos y la falta de actualización de la información una vez se firma el contrato, pueden explicar este sesgo.

Para mejorar la asignación de la ocupación, se ha trabajado con los microdatos de más de 250.000 personas procedentes de encuestas de mercado laboral (EPA y Encuesta de Estructura Salarial). Para estos registros se conoce su ocupación actual, bien por la EPA bien por la ESS, y la ocupación estimada según el FPC. Comparando ambas fuentes se ha confirmado que las discrepancias se concentran en un número reducido de categorías de la CNO: el FPC sobreestima las ocupaciones elementales e infraestima las más altas. Además se sobreestiman grupos concretos en los que las ocupaciones no están completamente definidas, como el 430 “Otros empleados administrativos sin tareas de atención al público”.

A partir de las variables disponibles en ambas encuestas y en los registros administrativos utilizados, se han seleccionado, mediante un análisis de correspondencias múltiples, las que presentan mayor correlación con las distintas ocupaciones “problemáticas”. Estas variables han resultado ser el nivel educativo, el grupo de cotización y la actividad. A partir de la combinación de las posibles categorías que toman estas tres variables se han construido 363 perfiles diferentes de individuos. Para cada perfil se ha estudiado la distribución de las ocupaciones que les asignan la EPA y la EES. Esta información será la que se tome como base para corregir los sesgos en la ocupación asignada a partir del registro de contratos del SEPE.

Otra mejora que se está implementando para determinar la situación en el empleo, es incluir información sobre retenciones por rendimientos de actividades económicas, procedentes de fuentes tributarias, para detectar trabajadores por cuenta propia.

7.3.8 Estudio de las variables relacionadas con la movilidad

La información relacionada con las preguntas de movilidad fue de la más demandada del Censo 2011. Tanto los sistemas de peticiones automáticos como los de peticiones a medida reflejaron que un gran porcentaje de los usuarios estaban interesados en este tipo de información.

En el Censo de 2011 se preguntó a las personas por la ubicación de su lugar de trabajo/estudio (pregunta 19) y también por si pasaba 14 noches o más (por motivos laborales o de veraneo) en un municipio diferente del de residencia (pregunta 4). Con estas dos preguntas se calculó la denominada *población vinculada*, incluida por primera vez en el Censo 2001, que permite disponer de una mejor estimación de la carga poblacional real que soporta un municipio.

Adicionalmente se incluyeron otras tres preguntas relacionadas con el número de viajes diarios, el medio de transporte y el tiempo transcurrido para llegar al lugar de trabajo/estudio.

4 En los últimos doce meses, ¿ha pasado más de 14 noches (aunque no fueran seguidas) en otro municipio de España o en otro país?

Puede ser por razones de trabajo, estudio, fines de semana, vacaciones o porque reside en más de un municipio

NO → (Pase a la pregunta 5)

SÍ → Indique el lugar donde ha pasado más noches y el número aproximado de noches que ha pasado allí:

Otro municipio:

Provincia:

Otro país:

Nº de noches:

¿Dispone en este lugar de una segunda vivienda (ya sea en propiedad, alquiler o cedida gratis)?

SÍ NO

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

20 ¿Cuántos viajes de ida y vuelta realiza normalmente cada día desde esta vivienda hasta el lugar de trabajo o estudio?

Ninguno (porque tiene una segunda residencia desde la que se desplaza a trabajar o estudiar)

Uno diario (es decir, uno de ida y otro de vuelta)

Dos o más diarios

Ya ha terminado el Cuestionario Individual de la Persona 1

21 ¿Cómo va normalmente desde su casa hasta el lugar de trabajo o estudio?

Si utiliza varios medios de transporte, marque los dos con los que cubre más distancia

<input type="checkbox"/> En coche o furgoneta como conductor	<input type="checkbox"/> En moto
<input type="checkbox"/> En coche o furgoneta como pasajero	<input type="checkbox"/> Andando
<input type="checkbox"/> En autobús, autocar, minibús	<input type="checkbox"/> En tren
<input type="checkbox"/> En metro	<input type="checkbox"/> En bicicleta
	<input type="checkbox"/> Otros medios

22 ¿Cuánto tiempo tarda normalmente desde su casa hasta el lugar de trabajo o estudio?

<input type="checkbox"/> Menos de 10 minutos	<input type="checkbox"/> Entre 45 y 59 minutos
<input type="checkbox"/> Entre 10 y 19 minutos	<input type="checkbox"/> Entre 1 hora y hora y media
<input type="checkbox"/> Entre 20 y 29 minutos	<input type="checkbox"/> Más de hora y media
<input type="checkbox"/> Entre 30 y 44 minutos	

Conseguir toda esta información a partir de fuentes administrativas no resulta posible.

En la TGSS se dispone de información de la cuenta de cotización que, indirectamente y posiblemente con algunos problemas de calidad, puede proporcionar información sobre el lugar de trabajo. No obstante la ubicación geográfica de la cuenta de cotización no siempre coincide con el lugar físico del puesto de trabajo, lo que limita el análisis.

Existe una nueva y prometedora fuente que podría dar respuesta a gran parte de la información proporcionada en 2011: el uso de la **información de coordenadas espacio-temporales de los teléfonos móviles**.

El INE se encuentra en la actualidad estudiando las posibles formas de acceder a esta información que obra en poder de las compañías operadores de telefonía móvil que cuentan con red propia de antenas y es prematuro anunciar la cantidad y calidad de la información de la que se dispondrá para el censo de 2021. En cualquier caso, lograr disponer de información de esa naturaleza se sitúa como una de las líneas prioritarias de investigación dentro del proyecto censal.

Si se pudiera analizar la información proporcionada por la señal de los teléfonos móviles, bajo ciertas hipótesis, como por ejemplo asumir que el lugar de residencia de una persona es aquel en donde se encuentra habitualmente por la noche o que su lugar de trabajo es aquel en el que está a ciertas horas del día, se puede llegar a reconstruir información bastante parecida a la que se ofrece para el Censo 2011.

Adicionalmente esta fuente abre la puerta a nuevos productos como, por ejemplo, poder conocer diferencias entre la población diurna y nocturna de una ciudad o poder analizar día a día la evolución de la población de un determinado lugar geográfico.

Por otro lado, el estudio de los medios de transporte utilizados, el tiempo empleado, el número diarios de trayectos, formará parte de la investigación complementaria al censo, la encuesta ECEPOV-2021 ya mencionada.

8 Censo de viviendas

8.1 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.

Como excepción, no se consideran viviendas los recintos que, a pesar de estar concebidos inicialmente para habitación humana, en el momento censal están dedicados totalmente a otros fines (por ejemplo, los que estén siendo usados exclusivamente como locales).

Tipos de vivienda: A continuación podemos observar un pequeño esquema de los distintos tipos de viviendas usados en el Censo de 2011.

- Viviendas familiares
 - Viviendas convencionales
 - Viviendas principales
 - Viviendas secundarias
 - Viviendas vacías
 - 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 (por ejemplo durante las vacaciones) se denomina vivienda **secundaria**. Cuando permanece sin ser ocupada se denomina **vacía**.

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). Entran en este grupo las cuevas, chabolas, caravanas...

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.

En el caso de las viviendas familiares, el fichero que se construyó para el censo de 2011 incluía datos para toda la muestra de viviendas relativos a la propia vivienda y al edificio que la contenía.

Análogamente a lo realizado con el fichero de personas, de cara al próximo censo de 2021 está previsto construir un fichero completo de viviendas familiares. Para ello es necesaria la integración de diversas fuentes administrativas.

Este capítulo se dedica al análisis de las viviendas familiares. El capítulo siguiente se dedica a los establecimientos colectivos.

8.2 EL FICHERO CENSAL DE VIVIENDAS (FCV)

De forma análoga a lo establecido para la población, se trabaja en construir un fichero que contenga todas las viviendas familiares del territorio nacional. De nuevo, el problema se puede contemplar en dos dimensiones: conocer cuántas viviendas hay dentro de cada ámbito de población y por otro lado conocer sus características a efectos censales (las variables que se precisa obtener para cada vivienda).

El producto que se pretende obtener, el *Fichero Censal de Viviendas* (FCV), sería un fichero que contendría un conjunto mínimo de datos para todas las viviendas familiares.

Además de este fichero, para las viviendas principales (aquellas donde reside población) se tendría un conjunto adicional de variables tales como el régimen de tenencia o datos de equipamiento. Este es el *Fichero Censal de Viviendas Principales* (FCVP).

La información a construir figura en el siguiente esquema (tomando como valores aproximados 26 millones de inmuebles totales y los 18 millones de viviendas principales de 2011):

Figura 4. Esquema de ficheros censales sobre viviendas para 2021

Fichero censal de viviendas

Número (identificador de vivienda)	Datos de localización (hasta sección censal) y coordenadas xy			...	características del edificio
	superficie				
1	xxx	45		...	zzzz
2	xxx	67		...	zzzz
3	xxx	120		...	zzzz
...			
...			
...			
26.000.000	...	30		...	zzzz

Fichero censal de viviendas principales

Número (identificador de vivienda)	Otras		
	régimen de tenencia	características de la vivienda	Equipamiento de la vivienda
1	yyy	xxxxxx	xxxxxx
3	yyyy	xxxxxx	xxxxxx
6	yyyy	xxxxxx	xxxxxx
...	...		
...	...		
18.000.000	yyyy	xxxxxx	xxxxxx

La principal decisión a tomar es cuál es, si existe, el registro administrativo que sirva de esqueleto de partida a partir del cual construir el FCV. En el caso de la población

resulta fácil la elección: el Padrón; en el caso de las viviendas la elección de ese fichero de partida no es tan evidente.

Tras un análisis detallado se ha considerado que la fuente administrativa a partir de la cual se elabora el Censo de Viviendas debe ser 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 catastral, además, ha mejorado sustancialmente durante los últimos años, como se ha podido constatar en los múltiples análisis realizados por el INE durante 2016-2018 para poder elegirlo como fuente.

El principal inconveniente para usar el catastro en el censo de viviendas es que no está enlazado con la población residente: no se sabe a priori qué vivienda del padrón corresponde a qué vivienda de catastro, es decir, quien vive en cada referencia catastral.

El Catastro no va a ser la única fuente para confeccionar el FCV. El trabajo se va a completar con el uso en algunos casos de una segunda fuente, el “marco de direcciones georreferenciadas” (MDG) del INE. Por último, son necesarias ciertas tareas de depuración manual que ya se están acometiendo y que se describen en detalle más adelante.

El MDG es un fichero elaborado por el INE desde noviembre de 2011. Consiste en la relación de vías, tramos, aproximaciones postales (direcciones postales hasta número, bloque/portal/escalera) y huecos en ellas (detalle hasta piso y puerta) existentes en la totalidad del territorio. Este MDG se construye sobre todo a partir del Censo de Edificios de 2011 y se va actualizando en campo de forma continua.

Este marco está relacionado con la población puesto que está cruzado con los ficheros padronales a través de las direcciones en que están registradas las personas en el Padrón y este cruce se revisa anualmente desde 2017, a la vez que se incorporan las modificaciones producidas en el callejero, comunicadas por la Oficina del Censo Electoral.

También contiene información derivada del cruce con los ficheros catastrales y del Instituto Geográfico Nacional (IGN) por lo que respecta a las coordenadas geográficas o a las referencias catastrales de los inmuebles. Aproximadamente se dispone tanto de las coordenadas como de las referencias catastrales de las aproximaciones postales en un 83% del fichero. En conjunto este fichero tiene un 90% de las viviendas con coordenadas geográficas.

Para tomarlo como base del FCV, el Catastro proporciona una información más completa que el MDG en tanto que garantiza mejor la exhaustividad (las nuevas viviendas no aparecen en el MDG salvo que alguien se empadrona en ellas), dispone de un identificador único y estable en el tiempo, la referencia catastral, que nos

permite enlazarlo con otros ficheros, muy particularmente con la AEAT, y ofrece coordenadas geográficas para el 100% de las viviendas, no solo para el 90% de ellas, con lo que resuelve por completo el problema de la georreferenciación.

Además, el MDG puede contener viviendas duplicadas. Dado que la forma de caracterizar una vivienda es su dirección postal y ésta información no está normalizada, una misma vivienda puede aparecer en dos momentos distintos referida de dos formas y en el MDG podrá aparecer como dos viviendas distintas. Este problema es objeto de mucha atención en el INE y hay múltiples procesos para evitarlo pero en algunos casos, sobre todo cuando los edificios tienen estructuras complejas, por ejemplo con más de una escalera dentro del mismo portal, se llega a hacer imposible la detección de duplicados.

El catastro está, en cambio, libre de ese problema, aunque tiene otros, como en algunos casos la falta de división horizontal que hace que muchas viviendas compartan una única referencia catastral o determinados bienes inmuebles, cuyo destino es distinto a vivienda, pero realmente se usan como tal.

La mejor solución está, por tanto, en la construcción de un marco de viviendas a partir de la información catastral pero combinando el uso del MDG y acudiendo a tareas de depuración manual.

8.2.1 Construcción del FCV a partir del Catastro

La Dirección General del Catastro⁸ elabora un registro administrativo, el Catastro Inmobiliario, en el que se describen los bienes inmuebles rústicos, urbanos y de características especiales. La inscripción en el Catastro es obligatoria y gratuita, características que lo diferencian del Registro de la Propiedad.

El Catastro contiene los denominados “bienes inmuebles” 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).

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.

El proceso de desentrañar esta información ha sido iterativo durante estos últimos años y aún no se puede considerar finalizado pero en el momento de redactar este proyecto ya se ha alcanzado el punto de madurez suficiente y se conoce con bastante

⁸ Se describe el Catastro de territorio fiscal común. Los catastros forales, con algunas peculiaridades, esencialmente tienen la misma naturaleza.

precisión donde están los problemas pendientes de subsanar para disponer de un FCV completo para 2021.

Inicialmente se construye un recuento de viviendas a partir del Catastro. Para ello, se parte de los ficheros de Construcciones y Bienes Inmuebles. Ambos se cruzan y se van identificando los distintos casos. A grandes rasgos: viviendas con una construcción (unifamiliares de una planta y pisos en bloques de edificios), viviendas de varias construcciones (dúplex en bloques de edificios, unifamiliares de varias plantas), viviendas en bloques o urbanizaciones sin división horizontal (todas forman un único bien inmueble).

Estas viviendas, dado que están perfectamente georreferenciadas, se pueden visualizar en el territorio. Para ello se usa una herramienta interna, la Infraestructura de datos espaciales del INE (IDE-INE).

En la figura 5 se observa un detalle de zona urbana. Las parcelas en amarillo son aquellas que contienen viviendas⁹.

Figura 5. Detalle de manzanas de Barcelona según la IDE-INE



Dado que todas las viviendas de catastro están georreferenciadas se les puede asignar la sección censal¹⁰ en la que se encuentran para poder efectuar los cálculos y comprobaciones a este nivel de detalle.

⁹ En la figura, para cada portal figuran dos números que reflejan el posible número de viviendas según dos recuentos a partir de los ficheros de Catastro. En general no hay divergencias, pero sí las hay en algunos casos. Este sistema de usar varios recuentos posibles y analizar divergencias ha sido un elemento central del análisis de los ficheros de catastro durante 2016 a 2018.

¹⁰ La **sección censal** es una unidad básica en el proyecto del Censo de Población y viviendas. Una sección censal es una subdivisión del municipio, usado con fines estadísticos y electorales. Con excepciones, es un territorio continuo habitado por un mínimo de 500 habitantes y un máximo de 2.500

El primer elemento de análisis consiste en comprobar el número de viviendas totales obtenidas para cada sección censal. Este número se compara con el número de viviendas obtenido en el recorrido del Censo de Edificios de 2011. También se compara con el número de hogares que contiene el fichero de población en cada sección censal. En la actualidad se contrasta con los ficheros precensales de 2016 y 2018. En 2021 se contrastará con el fichero censal de personas definitivo (el FCP).

Contar el número de hogares es lo mismo que contar las viviendas principales y es, en principio, imposible que el número de viviendas principales en una sección censal sea superior al número de viviendas totales.

El análisis numérico a nivel de sección censal proporciona una primera aproximación a la construcción del FCV. El primer recuento de viviendas totales a partir de catastro a nivel de sección censal se realizó en 2016 y se constató que solo unas 900 secciones censales (de unas 34.000 en territorio fiscal común, un 2,6%) contenían un número demasiado bajo de viviendas. En más del 95% de los casos los recuentos por sección censal resultan muy consistentes, también con el Censo de 2011.

En la IDE-INE se pueden encontrar casi siempre de forma visual cuales son los orígenes de esas anomalías. Se pueden observar las construcciones según ortofotografías, Google Street View, o coordenadas de edificios tomadas en el Censo de Edificios de 2011.

Las principales fuentes de tales anomalías son:

Algunas de las parcelas catastrales se extienden por zonas superiores a una sección censal, de forma que todas las viviendas de una parcela se asignan solo a una de esas secciones, la que contenga una mayor proporción de la superficie de la finca, quedando las otras sin viviendas asignadas.

En algunos casos no hay división horizontal de forma que una parcela catastral puede contener muchas viviendas (según padrón) pero solo una según Catastro, lo que da lugar a que $V_{principales} > V_{totales}$

También hay casos de discrepancia entre la sección asignada a una parcela catastral y la sección asignada para dicha parcela o edificio en el FPC.

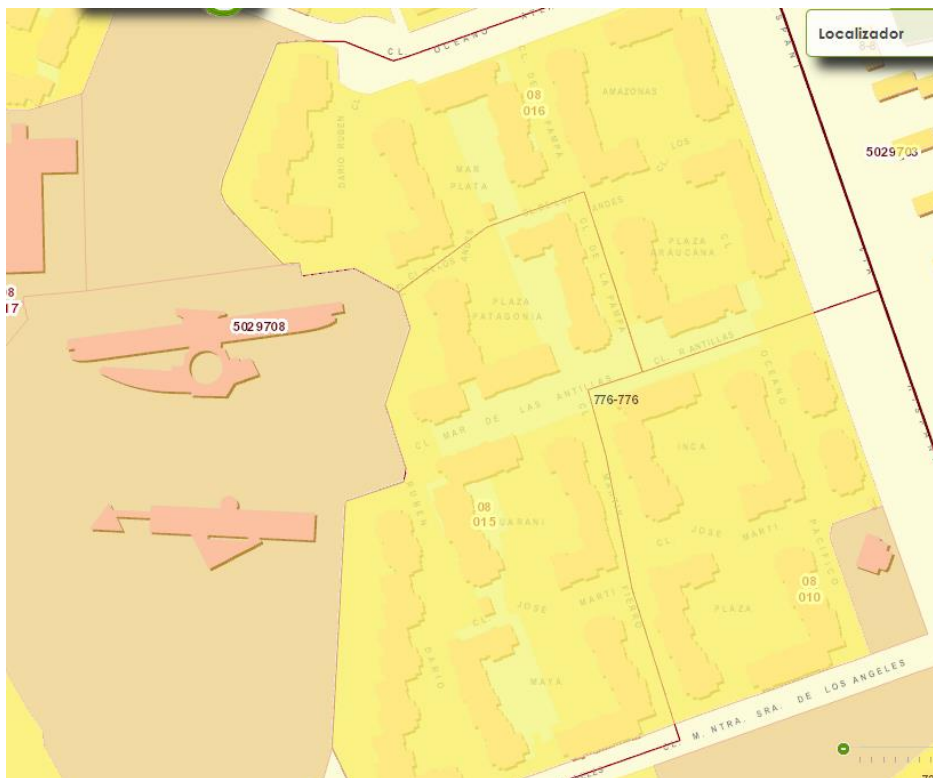
En general esas anomalías se van resolviendo manualmente, de forma que el fichero que se va construyendo no es el producto directo de explotar las fuentes catastrales sino que es producto también de esta depuración.

Por ejemplo, se observa en la figura 6 como una parcela catastral (recinto amarillo) contiene 776 viviendas y se extiende por las secciones 08016, 08015 y 08010 de Zaragoza, dando como producto que todas esas viviendas se cuentan en la sección que contiene mayor proporción de la superficie de la parcela, que está en la sección 08015.

personas (realmente se delimita en función del número de electores). Es un recinto cerrado, perfectamente delimitado por accidentes del territorio, geográficos o urbanísticos. Es una unidad mucho más homogénea que el municipio dado que rara vez supera los 3.000 habitantes y, salvo en municipios pequeños, tampoco suele bajar de los 1.000.

La sección censal será la unidad de referencia principal para la difusión de datos del censo 2021; también juega un papel fundamental como elemento de análisis, como en este caso.

Figura 6. Detalle de parcela que extiende por varias secciones censales (IDE-INE)



La corrección de anomalías a nivel de sección censal implica solucionar problemas de diversos tipos. En algunos casos aparecen problemas de trazado de las líneas que delimitan la sección censal, pero sobre todo la solución de problemas pasa por la partición de parcelas catastrales y por la asignación de una nueva referencia a las viviendas dada la insuficiencia de la referencia catastral (RC).

La RC es un identificador, de veinte posiciones, en principio único para cada bien inmueble (una vivienda, una plaza de garaje, etc.) asignado por Catastro y usado con fines tributarios. Las 14 primeras posiciones de la RC delimitan la parcela, luego un identificador usado también en el Censo es el RC14, el identificador de parcela catastral.

El análisis realizado para el Censo 2021 nos permite observar que, en algunos casos, un edificio completo tiene decenas de 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. Este identificador consiste en la RC, 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 figura 7 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 7. Detalle de una sección censal con parcelas sin división horizontal (IDE-INE)



Otro tipo de anomalía detectada, a nivel de sección censal, consiste en que hay coordenadas geográficas situadas sobre parcelas que no se corresponden con viviendas. En la figura 8 se observa el caso de una parcela de Loeches (Madrid) donde figura población empadronada.

Figura 8. Sección censal con población empadronada en parcelas no destinadas a viviendas (IDE-INE).



Como se puede observar en la imagen, los puntos rojos y rosas son coordenadas geográficas de población empadronada sobre edificaciones que no son viviendas (no están en amarillo). Al consultar estas parcelas en los ficheros de Catastro, vemos que están declaradas como oficinas. No obstante, al observar la parcela desde Google Street View, se constata claramente que son viviendas.

Se está investigando un procedimiento para detectar todos estos casos, ya que, aunque en los ficheros de Catastro no estén definidos como viviendas, si tendrían que considerarse como tales. Mientras tanto, se irán corrigiendo de forma manual, añadiendo esas referencias catastrales al directorio de viviendas.

También se han encontrado casos similares de personas que viven en apartahoteles en muchas zonas de Canarias. Cabe la duda de si incluir esa población, al menos en algunos casos, como residente en establecimientos colectivos. Esta cuestión está siendo objeto de análisis actualmente y no hay aún una decisión definitiva.

En el apartado dedicado a los alojamientos se vuelve sobre este asunto para el caso específico de viviendas no catastradas.

8.2.2 Enlace entre FCV y viviendas principales del FCP

Subsanados los problemas encontrados en los recuentos a nivel de sección censal se procede a la siguiente fase, en la que de nuevo aparecen inconsistencias que se deben subsanar manualmente: se trata de enlazar viviendas del FCV construido en la fase anterior con las correspondientes viviendas principales del fichero de personas, el FCP (actualmente se contrasta con los ficheros precensales). Dicho de otra forma, se trata de asignar, a cada hogar (=vivienda principal) del FCP cuál es la referencia catastral de esa vivienda.

El objetivo es conseguir enlazar al 100%, es decir, que todas las viviendas principales del FCP (aproximadamente había 18 millones de hogares en el censo de 2011) tengan asignado el identificador RC24. Al hacer esto estaremos 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.

Para este proceso se ha ideado un algoritmo en varias fases, que ha dado hasta hoy resultados excelentes, pero que está siendo objeto de mejora continua por lo que se espera llegar a la fecha censal con niveles de enlace superiores a los que se mencionan aquí.

El método seguido actualmente es el siguiente:

- Paso 1. Se toma la información de la 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

misma¹¹. 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 el mismo municipio se asignará al hogar del fichero FCP la RC de la vivienda según el FCV. Esto permite que a nivel nacional (salvo País Vasco y Navarra), un 61,9% de hogares tengan asignada una RC.

Además, con el FCV distinguiremos qué referencias catastrales se refieren 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 (se asignará solo si hay una única RC en la misma provincia) y los hogares para los que no se dispone de ninguna RC en el fichero de la AEAT.

- Paso 2. Para las viviendas que no se enlazaron (el tercer caso antes citado) se realiza el cruce del FCP con el fichero FIN46 de titulares de catastro. Si una persona se encuentra en el FIN46 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 10,3% a nivel nacional.
- Paso 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 16,8% de enlace a nivel nacional. Para las que no se consiga asignar se pasa a la siguiente etapa.
- Paso 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 8,0%. Para las que no se consiga asignar se pasa a la siguiente etapa.
- Paso 5. Se completa con nuevos cruces que podrían establecerse a la vista de los resultados que se vayan obteniendo, recurriendo en última opción a listados de comprobación y depuración manual.

A día de hoy este procedimiento permite ofrecer RC a 14 o a 20 posiciones a la gran mayoría de viviendas del fichero precensal FPC-2018. Los resultados a julio de 2019 se observan en la siguiente tabla:

¹¹ En los modelos tributarios correspondientes del País Vasco y Navarra se usa otro identificador (número fijo) pero figura peor informado que en el modelo 100 de la AEAT.

Porcentajes de asignación de RC a viviendas del FPC (territorio fiscal común)

Provincias	% Asignadas	% REF20	% REF14	% Pendientes
Albacete	97,6%	77,6%	20,0%	2,4%
Alicante/Alacant	96,9%	75,1%	21,9%	3,1%
Almería	93,5%	66,9%	26,6%	6,5%
Ávila	95,2%	73,5%	21,7%	4,8%
Badajoz	97,2%	78,0%	19,2%	2,8%
Balears, Illes	93,8%	67,8%	26,0%	6,2%
Barcelona	99,1%	88,2%	11,0%	0,9%
Burgos	96,7%	81,5%	15,2%	3,3%
Cáceres	95,9%	73,3%	22,6%	4,1%
Cádiz	96,3%	75,8%	20,5%	3,7%
Castellón/Castelló	98,1%	79,8%	18,3%	1,9%
Ciudad Real	97,7%	79,0%	18,7%	2,3%
Córdoba	97,9%	80,3%	17,6%	2,1%
Coruña, A	94,3%	77,4%	17,0%	5,7%
Cuenca	96,0%	73,3%	22,7%	4,0%
Girona	97,5%	76,9%	20,6%	2,5%
Granada	96,1%	73,8%	22,4%	3,9%
Guadalajara	96,3%	80,5%	15,8%	3,7%
Huelva	95,3%	72,8%	22,5%	4,7%
Huesca	96,6%	78,0%	18,6%	3,4%
Jaén	97,4%	78,3%	19,1%	2,6%
León	95,7%	78,4%	17,3%	4,3%
Lleida	95,7%	77,5%	18,2%	4,3%
Rioja, La	98,9%	88,6%	10,3%	1,1%
Lugo	90,1%	75,7%	14,4%	9,9%
Madrid	99,4%	91,6%	7,8%	0,6%
Málaga	95,3%	71,3%	24,0%	4,7%
Murcia	95,6%	73,4%	22,3%	4,4%
Ourense	90,2%	68,2%	22,0%	9,8%
Asturias	95,7%	84,7%	11,0%	4,3%
Palencia	97,4%	82,8%	14,6%	2,6%
Palmas, Las	94,8%	61,4%	33,4%	5,2%
Pontevedra	93,4%	75,5%	17,9%	6,6%
Salamanca	97,0%	80,9%	16,1%	3,0%
Santa Cruz de Tenerife	94,1%	59,5%	34,5%	5,9%
Cantabria	97,3%	84,5%	12,8%	2,7%
Segovia	97,3%	77,7%	19,6%	2,7%
Sevilla	97,8%	82,2%	15,6%	2,2%
Soria	94,3%	78,7%	15,7%	5,7%
Tarragona	97,3%	80,6%	16,7%	2,7%
Teruel	97,0%	75,5%	21,5%	3,0%
Toledo	97,5%	78,0%	19,5%	2,5%
Valencia/València	98,7%	88,8%	9,9%	1,3%
Valladolid	98,9%	89,4%	9,5%	1,1%
Zamora	95,7%	76,6%	19,0%	4,3%
Zaragoza	98,6%	86,6%	12,0%	1,4%
Ceuta	91,9%	61,8%	30,1%	8,1%
Melilla	94,8%	68,4%	26,4%	5,2%

En aquellos casos en que varias viviendas principales compartan la misma RC20 (ya se ha mencionado antes el problema de la falta de división horizontal) se asignará de forma secuencial un RC24 distinto a cada una de ellas, generando nuevas viviendas también en el FCV.

Cuando no se pueda delimitar la RC completa sino solo la RC14 de las viviendas, de nuevo ésta se asignará aleatoriamente dentro de las no asignadas de cada parcela.

Por último, subsistirá siempre un reducto de viviendas principales para las que no se pueda asociar la vivienda correspondiente del FCV. Se asignarán de forma aleatoria dentro de las secciones censales entre las viviendas no asignadas usando la información disponible (nombres de calle, número...)

Como resultado de todo este proceso, se dispondrá, con un grado de imputación muy bajo, de un identificador estadístico (RC24) para el 100% de las viviendas principales¹².

¹² Ver apartado de *Alojamientos* para el tratamiento de viviendas no catastradas

8.3 DETERMINACIÓN DEL TIPO DE VIVIENDA SEGÚN USO

El método clásico, usado hasta el censo de 2011, para la determinación de la clase de vivienda convencional según su uso (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.

Para 2021 se propone una clasificación diferente y se proporcionará otro tipo de información sobre viviendas según uso.

8.3.1 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.

Como se ha mencionado en el apartado 8.2.2. Este paso está prácticamente resuelto de forma que todas las viviendas del fichero de personas (salvo problemas menores) tienen RC24 asignadas.

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.

De cara al censo de 2021 cabría plantearse de nuevo esas mismas dos categorías (secundaria, vacía) e incluso quizá una tercera para otros usos, como podría ser el de pisos turísticos, de creciente implantación en muchas zonas.

Pero los registros administrativos en la actualidad no permiten llegar tan lejos. Así pues, se opta como primera aproximación, por proporcionar para todas las viviendas solo dos categorías: principales y no principales. Pero no se renuncia a proporcionar otra información que permita a los usuarios distinguir entre zonas donde mayoritariamente hay viviendas vacías y secundarias, como se detalla a continuación.

8.3.2 Clasificación de viviendas según consumo eléctrico

Para el censo 2021 se pretende establecer una nueva clasificación alternativa de las viviendas según su consumo eléctrico. Es una clasificación distinta a la del uso, basada en hechos más objetivos pero no va a ser posible identificar, para cada vivienda concreta del FCV, a qué categoría pertenece (secundaria, vacía).

Esta nueva clasificación estará basada en la cantidad anual de consumo eléctrico (kwh) de cada vivienda. Está pendiente de decidir el número de rangos de consumo para clasificar las viviendas

El fichero administrativo de partida es la información anual de consumo eléctrico, a partir de fuentes tributarias, individualizada de más de 40 millones de contadores eléctricos de todo el territorio nacional. Se trata del modelo 159 de la AEAT (y similares de haciendas forales). Esta información corresponde no solo a viviendas sino también a locales, comunidades de vecinos o empresas de todo tipo. En definitiva, la unidad es el contador o Código Único de Punto de Suministro (CUPS).

Tras un estudio detallado de este fichero surge un problema de difícil solución de cara al censo de 2021. No siempre se consigue asociar esos CUPS a las viviendas del FCV para poder así determinar en concreto qué consumo tiene cada vivienda del FCV. Esto obliga a incorporar la información de consumo eléctrico con un planteamiento alternativo.

La primera dificultad a la hora de cruzar esos datos de consumo eléctrico con el FCV reside precisamente en determinar cuáles de esos registros (CUPS) corresponden a viviendas. En la actualidad podemos determinar, con cierto margen de error, que entre 22 y 24 millones de esos CUPS corresponden a viviendas, lo cual sería consistente con la existencia de unos 25-27 millones de viviendas, dado que algunas viviendas vacías no tendrán siquiera suministro eléctrico.

Una segunda dificultad se encuentra en el enlace entre ficheros. El identificador para enlazar FCV y consumo eléctrico podría ser la referencia catastral pero su uso todavía no está generalizado en el modelo tributario 159, de manera que el fichero de consumo eléctrico apenas contiene ese campo relleno en un 51% de los registros.

A falta de un identificador único de vivienda los cruces se tienen que realizar a partir de la dirección postal. Estos procesos de cruce a partir de direcciones postales siempre son más complicados y dan peores resultados que los de personas a partir de nombres y apellidos. Actualmente se logra enlazar en torno a un 70% de las viviendas lo cual no permite determinar, para cada vivienda del fichero censal, su consumo eléctrico.

Se está trabajando en una mejora de los cruces utilizando como apoyo el NIF de los titulares de los contratos del que se dispone en el fichero de consumo eléctrico y el NIF de los residentes según padrón del que se dispone en el marco, así como del NIF de los titulares de los bienes inmuebles de que se dispone en catastro. No obstante hay que tomar con mucha precaución esta información y garantizar que el enlace se produce no solo a través del NIF sino que simultáneamente se produce o bien por referencia catastral o bien por dirección postal.

Además, los propios trabajos de mejora del FCV van a permitir normalizar direcciones postales y van a producir mejoras en el enlace. Pero no es razonable esperar que el porcentaje de CUPS a las que se logre enlazar una RC superior al 80%. Sin duda subsistirá un número de secciones donde el cruce no arroje porcentajes suficientemente altos.

Eso no invalida la construcción de indicadores para el censo a partir del consumo eléctrico pero obliga a un planteamiento diferente. Un primer objetivo que parece alcanzable es determinar la sección censal en un porcentaje elevado de los registros que en el fichero de consumo eléctrico se han identificado como viviendas.

Si se dispone de tal clasificación, una aproximación realista aunque quizá demasiado prudente consistiría en proporcionar resultados renunciando al enlace entre CUPS y RC y utilizar las cifras agregadas que se obtienen explotando únicamente el fichero de consumo eléctrico a nivel de sección censal (al menos de municipio) para construir indicadores globales de viviendas según consumo.

Este tipo de información, que podemos denominar, “**de contexto**” no se tendría para cada vivienda en particular sino de forma agregada por secciones censales, complementando la información que se ofrezca de los productos principales, el FCP y el FCV.

Es decir, a partir del FCV se podría proporcionar información para cada sección censal del total de viviendas y de las viviendas principales; por otro lado, a partir del fichero de consumo eléctrico, sin cruzar con ninguna otra fuente, se podrían proporcionar indicadores para cada sección censal sobre porcentajes de viviendas por rango de consumo.

Obviamente esto podrá arrojar información inconsistente, como secciones censales en las que hay más o menos viviendas principales de las que tienen un consumo eléctrico que podríamos considerar compatible con estar ocupadas.

Otro tipo de información agregada que podría darse para cada sección censal usando solo datos del fichero de consumo eléctrico podrían ser estadísticos sobre consumo medio o mediano, restringido (o no) a viviendas que tienen consumo, u otros sobre distribución del consumo (percentiles u otros) que, con seguridad, ayudan a dar una imagen agregada de las secciones censales en cuanto a su consumo eléctrico, e indirectamente, a su ocupación (incluso a su nivel de vida).

8.3.3 Locales

Como en ocasiones anteriores, en el Censo de Población y Viviendas de 2011 se realizó un Censo de Edificios. Este no fue exhaustivo pero se dirigió al 80% de las secciones del territorio nacional y éstas se recorrieron de forma completa. El 20% restante de secciones se completó con información administrativa.

El trabajo consistió en enumerar y georreferenciar todos los edificios que tuvieran alguna vivienda y determinar sus características mediante un cuestionario de edificio. Por tanto, no todos los edificios de cada sección se censaron.

Dentro de cada edificio se seleccionaban los huecos¹³, los cuales a su vez se clasificaban en viviendas y locales.

A efectos censales, un local es un recinto estructuralmente separado e independiente (en el mismo sentido que en la definición de vivienda) que no está exclusivamente dedicado a vivienda familiar y en el que se llevan o se pueden llevar a cabo actividades económicas dependientes de una empresa o institución. El recinto debe estar situado en un edificio, ocupándolo total o parcialmente.

En censos anteriores se investigaban estos locales. De hecho en 1991 se publicó un “censo de locales”. Esta información de carácter más bien económico se fue diluyendo en los censos demográficos. Así, en 2001 apenas se proporciona una mínima clasificación de los 2,5 millones de locales según su actividad y en 2011 ni siquiera se censan los locales “a pie de calle”, los que están fuera de los portales.

Pero dentro de la estructura de un portal, no solo hay viviendas, al menos si hablamos desde el punto de vista de un censo de población. Una vivienda en un tercer piso, concebida como tal, pero utilizada en la práctica exclusivamente como una consulta médica o como un despacho de abogados, no es una vivienda sino que debe contarse como un local.

Así, en el censo de 2011 se contabiliza (dentro de los portales) un total de 25,849.338 “inmuebles” pero de ellos solo 25.218.536 son consideradas como viviendas en tanto que se usan como tales.

Los datos de viviendas 2021 no serán del todo comparables a las viviendas de 2011 sino más bien a los inmuebles. Con la información catastral no puede determinarse si algo que está declarado como vivienda se está usando como local. Se puede conocer parcialmente, cuando está declarado así.

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 cualquier caso, el censo de 2021 (como el de 2011) no tiene como objetivo proporcionar información sobre locales sino solo sobre viviendas, por lo que ya no cabe hablar de “huecos”. Será vivienda todo aquello que figure en Catastro como tal, independientemente de su uso real, que no podrá determinarse a partir de registros administrativos.

¹³ En terminología de censos anteriores, el hueco es una unidad operativa usada en el cuaderno de recorrido del agente censal, que corresponde, según su uso, a una vivienda o un local.

8.4 VARIABLES PARA TODAS LAS VIVIENDAS

Análogamente a lo expuesto para la población, una vez determinado el conjunto total de viviendas que forman el FCV se trata de determinar las características de estas viviendas.

Hay que distinguir, en primer lugar entre el total de viviendas y las viviendas principales porque para este subconjunto se precisa de más información.

La siguiente lista muestra las variables contenidas en el censo de 2011 y marcadas con asterisco, las requeridas en el Reglamento UE:

Datos de la vivienda

- Tipo de vivienda (*)
- Régimen de tenencia (*)
- Calefacción (*)
- Cuanto de aseo con inodoro (*)
- Baño o ducha (*)
- Acceso a Internet
- Sistema de suministro de agua (*)
- Superficie útil (*)
- Número de habitaciones (*)

Datos del edificio

- Número de plantas sobre rasante
- Número de plantas bajo rasante
- Tipo de edificio (*)
- Número de inmuebles
- Año de construcción (*)
- Estado del edificio
- Ascensor en el edificio
- Accesibilidad del edificio
- Garaje
- Número de plazas de garaje
- Gas
- Tendido telefónico
- Agua caliente central
- Evacuación de aguas residuales

El reglamento UE establece unas demandas sobre las variables de las viviendas bastantes limitadas. En concreto, para todas las viviendas, se exige conocer:

- Número de ocupantes (que será cero en las no principales)
- Tipo de edificio en el que encuentran (residencial: con una, dos, tres o más viviendas, no residencial)
- Antigüedad del edificio

Las dos primeras se proporcionarán a nivel municipal y la última sólo a nivel provincial.

Para las viviendas principales se requieren además estas otras variables, también a nivel NUTS-3 (provincias e islas) excepto las 3 últimas para las que solo se exige el nivel de detalle NUTS-2 (CCAA).

- Régimen de tenencia
- Superficie útil (o, en su defecto, número de habitaciones)
- Estándar de densidad (que se calcula a partir de las dos anteriores)
- Sistema de suministro de agua (*)
- Instalación de baños y retretes (*)
- Tipo de calefacción (*)

(*) En el Reglamento se reconoce que, en el caso de algunos Estados miembros en los que haya pruebas (basadas en censos anteriores, en fuentes de datos administrativos o en datos de encuestas muestrales) de que prácticamente todas las viviendas convencionales dispongan de estos servicios o instalaciones, estas características se puede asumir que están presentes en el 100% de ellas y codificarse todas por igual. Cuando los Estados miembros elijan esta opción, deberán certificar esta hipótesis y explicar este aspecto en los metadatos. En el caso de España esto afectaría a las dos primeras de estas tres, no así al tipo de calefacción.

No obstante, esta información se recogerá mediante la encuesta ECEPOV-2021, tal como se expone más adelante.

Algunas otras de las variables relativas a viviendas o edificios pueden obtenerse a partir de la información contenida en los ficheros de los distintos catastros. El catastro incorpora información sobre el año de construcción/reforma del edificio y sobre la superficie de los bienes inmuebles. Esta información se incorporará al FCV luego se dispondrá de ella para todas las viviendas.

Otras variables como el número de plantas sobre o bajo rasante puede obtenerse también de los ficheros de Catastro.

8.5 VARIABLES DE LAS VIVIENDAS PRINCIPALES

Como ya se ha dicho, algunas variables solo se requieren para las viviendas principales. Se describe la situación de éstas siguiendo el esquema usado para las de la población.

8.5.1 Régimen de tenencia

a) Requisitos

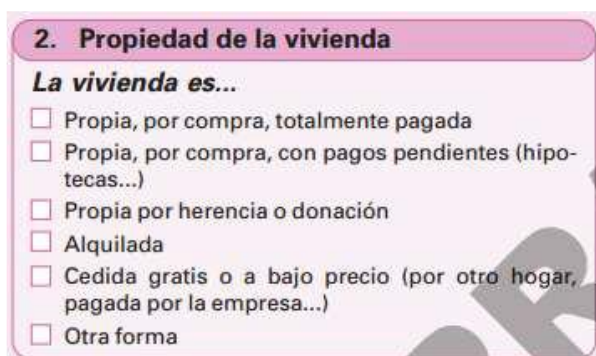
i. Reglamento UE

El reglamento solo obliga a proporcionar datos a nivel provincial, como mucho, para las viviendas principales. Las categorías obligatorias para esta variable son solo cuatro:

- **Vivienda en propiedad** (al menos un ocupante de la vivienda es propietario de la misma)
- **Vivienda en alquiler** (al menos un ocupante de la vivienda es arrendatario de la misma y ningún otro ocupante es propietario)
- Vivienda con **otro régimen de tenencia**
- **No consta**

ii. Censos anteriores

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



2. Propiedad de la vivienda

La vivienda es...

- Propia, por compra, totalmente pagada
- Propia, por compra, con pagos pendientes (hipotecas...)
- Propia por herencia o donación
- Alquilada
- Cedida gratis o a bajo precio (por otro hogar, pagada por la empresa...)
- Otra forma

b) Planteamiento propuesto para la variable 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 proponen estas tres categorías, coincidentes con las especificadas en el actual Reglamento. Estas categorías se intentarán establecer para todas las viviendas principales:

- 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)

c) Método y fuentes

En los ficheros precensales la estimación de esta variable se está llevando a cabo mediante el cruce de las siguientes fuentes:

- Información tributaria del modelo 100 relativa al ejercicio económico del año anterior proporcionada por la Agencia Tributaria, que incluye para cada persona declarante las referencias catastrales de su vivienda habitual más garajes y anexos, pero sin diferenciar unos inmuebles de otros, así como la titularidad sobre los mismos (propiedad, alquiler, otro).
- Información catastral referida a 1 de enero de 2016 proporcionada por la Dirección General del Catastro que proporciona información sobre la titularidad de los inmuebles.

Solamente con esta información tributaria se logra ya conocer el régimen de tenencia de casi un 80% de los hogares. Después de estos cruces queda una parte de las viviendas sin información. En principio, para el FPC-2016 se ha procedido a imputar, para lo cual se recurre a la utilización de la Encuesta Continua de Hogares de los años 2015 y 2016, con una muestra efectiva total de unas 107.000 viviendas.

Se ha ideado un primer método de imputación probabilística, todavía provisional, usando las características de los miembros que componen el hogar para determinar la probabilidad de que los hogares no asignados figuren en régimen de alquiler o de propiedad. No obstante, el método para la construcción de esta variable todavía será objeto de análisis futuros más detallados.

d) Análisis de la calidad de la variable

Considerando el conjunto de viviendas del FPC-2016 el número de registros según la fuente utilizada a la hora de asignar la variable régimen de tenencia (RT) se distribuye así:

Fuente ganadora:

Fuente de información	Porcentaje
Total	100,0%
AEAT	64,3%
CATASTRO	13,7%
Sin información	21,9%

Todavía se observa un porcentaje muy elevado de no asignados en Navarra y en las tres provincias vascas debido a que aún no se disponía en el FPC-2016 de información tributaria de las haciendas forales.

Con la imputación provisional de los no asignados a nivel provincial utilizando como referencia la ECH, la distribución resultante es muy parecida a la de la propia ECH:

Comparación entre los resultados obtenidos en el Fichero Precensal y el Censo de 2011

	FPC2016 %	Censo 2011 %	Diferencia	ECH elevada %
Total viviendas familiares	100,0%	100,0%	0,0%	100,0%
Total Propiedad	76,3%	78,9%	-2,6%	77,4%
Total Alquiler	17,0%	13,5%	3,5%	17,4%
Total Otros Casos	6,7%	7,6%	-0,9%	5,2%

En resumen, se consigue asignar régimen de tenencia a un 78% de las viviendas, lo que supone casi un 84% si se excluyen País Vasco y Navarra. El resto se imputan probabilísticamente en función de las características del hogar.

e) Próximos pasos y mejoras previstas

Son varias las mejoras pendientes que van a permitir aumentar el porcentaje de viviendas a las que se asigna régimen de tenencia:

- En esta primera aproximación, no se ha tenido en cuenta la información que aporta el fichero de la Agencia tributaria sobre alquileres desde el punto de vista del arrendador. Utilizando esta información se podría asignar régimen de tenencia 'alquiler' a viviendas en las que residen otras personas que no son sus titulares pero de las cuales el titular informa que están en régimen de alquiler.
- También está pendiente de analizar en detalle la información que han aportado las Haciendas Forales, tanto administrativa como tributaria.

Con la mejora de los enlaces entre FCV y FCP se podrá mejorar la información sobre propietarios de viviendas, dado que en el procedimiento actual dicha identificación se ha realizado solo a través de los NIFs de los residentes. Es razonable suponer que el régimen de tenencia se va a poder asignar en más de un 85% de casos. Sobre todo se va a poder identificar casi exhaustivamente las viviendas en régimen de propiedad, por lo que la imputación va a ir encaminada casi siempre a asignar el alquiler como régimen de tenencia.

8.5.2 Otras variables sobre las viviendas principales

De la lista expuesta en el punto 8.4 algunas variables se precisan para todas las viviendas; otras solo para viviendas principales. Descontando el régimen de tenencia ya expuesto, las demás variables que se precisaban para las viviendas principales eran la superficie útil (o, en su defecto, número de habitaciones), el estándar de densidad (que se calcula a partir de las dos anteriores), sistema de suministro de agua e instalación de baños y retretes.

Dado que no es necesario proporcionar esta información de forma exhaustiva – el Reglamento solo exige datos por NUTS-2 o NUTS-3 – esta información se va a incorporar en el cuestionario de la encuesta complementaria al censo, la ECEPOV-2021, que tendrá tamaño muestral suficiente para garantizar la elaboración de esta información a nivel NUTS-3.

8.6 ALOJAMIENTOS Y PERSONAS SIN HOGAR

En España apenas hay población residiendo en alojamientos. En el censo de 2001 se cuentan 3.143 y en 2011 son 2.097 personas. Estas cifras probablemente infraestimen el número real de alojamientos, pero es difícil censar todos ellos en una operación de campo en el territorio.

En cualquier caso, es uno de los elementos que actualmente es más difícil de extraer de los registros administrativos. De acuerdo con el actual Reglamento se pide clasificar los “locales de habitación”, es decir, las unidades donde puede residir población en:

- Viviendas convencionales ocupadas
- Otras unidades de vivienda
- Establecimientos colectivos

Las “otras unidades de vivienda” pueden ser chozas, cabinas, cabañas, chabolas, caravanas, casas flotantes, graneros, molinos, cuevas o cualquier otro refugio utilizado para la habitación humana en el momento del censo, con independencia de si estaba concebido para ello. Esta información se solicita a nivel municipal.

También se debe especificar cuánta población reside en cada uno de ellos. En este caso la población se clasificaría así:

- Ocupantes de vivienda convencional
- Ocupantes de establecimiento colectivo
- Ocupantes de otras unidades de vivienda y personas sin hogar

Por tanto, a efectos del Reglamento, las personas que viven en alojamientos se contarían conjuntamente con las personas sin hogar, sin que sea necesario separarlas. Además, esta información solo se debe entregar a nivel de NUTS-3 (provincia, isla).

Pendiente de un análisis aún no abordado en profundidad, el número de alojamientos a nivel municipal se puede considerar nulo en general, en línea con lo encontrado en censos anteriores, si bien cabe realizar investigaciones específicas en algunos municipios concretos. En el padrón es difícil identificar los alojamientos pero se conocen tratamientos específicos en algunos ayuntamientos, que identifican chabolas, por ejemplo. Todavía se encuentra en fase de estudio la posible incorporación de fuentes locales o de otros ámbitos.

Por otro lado, no está previsto realizar ninguna investigación sobre personas sin hogar para el Censo 2021, toda vez que el INE ya dispone de una encuesta específica sobre la materia. De todas formas, la investigación sobre establecimientos colectivos puede ofrecer cierta información relacionada con personas sin hogar dado en general las personas sin hogar en España se encuentran empadronadas, en el 90% de los casos, y lo están en cierto tipo de establecimientos colectivos¹⁴.

Pero independientemente de la información sobre alojamientos, aunque en cierto modo relacionada con ellas, se puede realizar una investigación de otro tipo, que

¹⁴ Para más información, se pueden consultar los datos de la última Encuesta a las Personas sin Hogar (INE): https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176817&menu=ultiDatos&idp=1254735976608

puede arrojar información útil sobre asentamientos o viviendas no recogidas en Catastro.

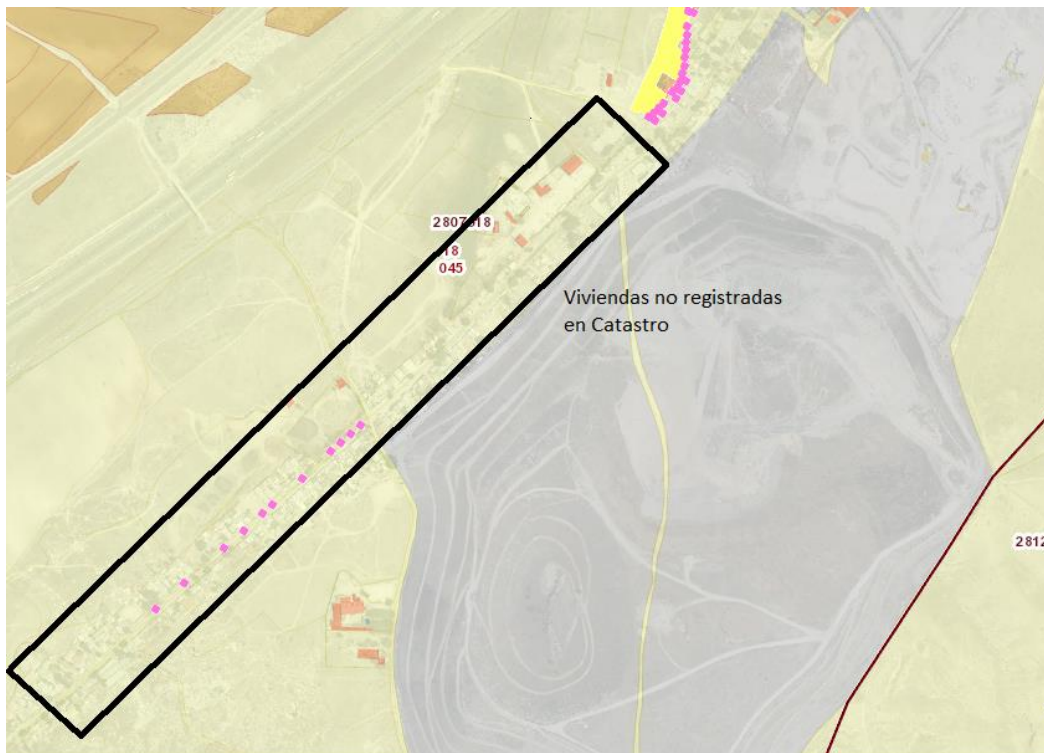
Si bien es necesario completar todavía esta investigación, en líneas generales se puede avanzar que el censo de 2021, propiciada por el uso de registros administrativos, se plantea proporcionar una nueva información sobre población residente en zonas no catastradas.

El enlace entre fichero de población y fichero de viviendas, detallado en el apartado 8.2.2 permite identificar algunas situaciones especiales como aquellas en las que reside población en viviendas no registradas por el Catastro.

Una de las anomalías detectadas a nivel de sección censal o inferior consiste en encontrar más viviendas principales (es decir, hogares del FCP) que viviendas en su conjunto o direcciones postales en el FCP para las cuales no se tienen viviendas en el FCV. Este tipo de anomalías es relativamente común por la diferente nomenclatura de calles y vías usadas en ficheros de Catastro y Padrón pero tras los enlaces entre ficheros antes mencionados y con el uso de herramientas visuales se están detectando claramente algunas zonas donde el problema reside en la existencia de población empadronada en viviendas ilegales o, en todo caso, no registradas en Catastro. Algunos ejemplos muy relevantes se detectan al analizar a nivel de sección censal. Por ejemplo, en la figura 9 se observa una sección censal de Madrid.

En esta sección, en concreto, se contabilizarían en el FPC-2016 un total de 993 viviendas principales pero apenas se cuentan 49 en el FCV.

Figura 9. Sección censal con presencia de viviendas no registradas (IDE-INE)



Este es el caso más extremo observado en los análisis realizados hasta hoy. Se puede observar como una población superior a 3.000 personas en esa sección censal figura empadronada en viviendas no catastradas (las zonas catastradas son las parcelas en amarillo).

Además, los puntos de color rosa de la figura corresponden a las coordenadas geográficas tomadas por el agente censal que recorrió la zona en 2011. Esta zona no se pudo censar tampoco con calidad (según el censo de 2011 se estimaban 320 viviendas principales). Observando esa zona con Google Street View (figura 10) no se observa que se trate de alojamientos sino más bien de zonas marginales con viviendas no registradas en Catastro.

Figura 10. Viviendas no registradas en sección censal 2807918045 de Madrid.
(Fuente: Google Street View. Fecha de la captura: agosto 2019).



Lo que se plantea es una generación manual, en estas zonas, de polígonos que contendrían población empadronada en viviendas no catastradas. En estos casos concretos, se podría trazar (manualmente) un polígono como el de la figura 9, sería una “pseudoparcela” y asignar a ella, con ayuda de herramientas como google Street View y sistemas de información geográfica, a la población recogida en el FCP en direcciones postales que caen dentro de ella. Esta tarea ya se ha iniciado y se han identificado varias zonas en algunas ciudades (Móstoles, Almería, Sevilla...).

9 Establecimientos Colectivos

9.1 LA POBLACIÓN RESIDENTE EN ESTABLECIMIENTOS COLECTIVOS EN EL CENSO DE 2011

No hay que olvidar que, si bien la mayor parte de la población (el 99% en el último censo), reside en viviendas familiares, hay una pequeña parte de la población que reside en las denominadas viviendas o establecimientos colectivos. Y también es preciso cuantificarlas y obtener las características sociodemográficas de estas personas.

En el Censo de 2011, la población residente en establecimientos colectivos se investigó mediante una operación de campo independiente, la *Encuesta de Colectivos*, que permitió determinar el número y tipo de establecimientos colectivos y las características demográficas básicas de las personas que residían en ellos.

La Encuesta de Colectivos utilizó un marco obtenido de distintas fuentes, aprovechando información disponible en el INE, tanto del censo anterior, como de diversas encuestas (por ejemplo la Encuesta de Alojamientos Turísticos o la Encuesta de Discapacidad EDAD-08) y de directorios facilitados por organismos externos, como el Ministerio de Defensa.

Sobre el directorio previo de establecimientos colectivos se hizo una investigación exhaustiva para todos los tipos de establecimiento excepto los hoteles, pensiones y similares, que se investigaron mediante una muestra. Dentro de cada colectivo se entrevistaba a una muestra de personas.

La encuesta permitió determinar el número de establecimientos colectivos en la fecha de referencia censal y su tipo, así como el número y las características demográficas básicas de las personas que residían en ellos. Estas variables y su nivel de desagregación, menor que en el caso de la población en viviendas familiares, están recogidas en la reglamentación desarrollada por la Unión Europea a excepción del tipo de establecimiento, que se incluyó atendiendo exclusivamente a necesidades de usuarios nacionales.

Hay que mencionar que en el censo de 2011 se introdujo una importante novedad a la hora de censar la población residente en establecimientos colectivos. Se consideró residente en un establecimiento colectivo aquella persona que lleva residiendo o espera residir un año en él, independientemente de su situación de empadronamiento, pero también se les preguntó a las personas por esta circunstancia. Así, se obtuvo que solo el 54% de las personas residentes en colectivos estaban empadronadas en ellos, destacando las prisiones en las que apenas un 10% de la población residente está empadronada.

Dado que el censo de población se construyó a partir de un tratamiento de ficheros que a su vez derivaba del padrón, las personas residentes en viviendas familiares se contaron donde figuraban empadronadas. Así, las personas que residiendo en establecimientos colectivos no estaban empadronadas en ellos, se contaron como residentes en las viviendas familiares donde figuran empadronados para evitar contarlos doblemente.

9.2 PLANTEAMIENTO PARA EL CENSO 2021

El planteamiento para el censo de 2021 parte de considerar como población residente en establecimientos colectivos a las personas que figuran en el FCP en direcciones que corresponden a éstos, es decir, personas empadronadas en ellos. Se podrá, por tanto, proporcionar la misma información para cada persona que la obtenida para residentes en viviendas familiares, así como las tipologías de los establecimientos en que residen.

Esto supone, por un lado una mejora considerable respecto a la cantidad de información obtenida en censos anteriores, donde el cuestionario solo se dirigía a viviendas familiares y por tanto la investigación censal era más limitada para la población institucionalizada.

Por otro lado, considerar como residentes en establecimientos colectivos solo a los empadronados en ellos es obviamente una limitación. Particularmente en el caso de las residencias de personas mayores y en el de las prisiones el criterio de residencia usado puede dar lugar a cifras muy dispares.

La principal dificultad reside en conocer cuándo una persona del FCP está residiendo en un establecimiento colectivo para lo cual lo primero es disponer de un directorio de establecimientos colectivos.

La primera fase consiste en obtener un registro completo y actualizado de candidatos a ser establecimientos colectivos.

Para ello se ha partido, por una parte, de la información ya disponible en el INE procedente tanto de la Encuesta de Colectivos del Censo 2011 como de la Encuesta de Discapacidad 2008, así como los ficheros de instituciones religiosas, militares y prisiones proporcionados por organismos externos, conformando un primer directorio de **8.359** centros:

Total	Residencias mayores	Residencias discapacidad	Instituciones religiosas	Militares y penitenciarias	Otros
8.359	4.221	1.090	1.195	290	1.563

Dado que casi un 40% de las direcciones postales de los establecimientos de este directorio no enlazan inicialmente con las del fichero precensal de personas, se ha depurado manualmente tomando como referencia la información actualizada de residencias de mayores de portal de envejecimiento del CSIC y de centros para personas con discapacidad del IMSERSO.

Mediante las búsquedas manuales realizadas hasta el momento se han investigado todas las direcciones del directorio inicial correspondientes a residencias de mayores y discapacidad que no enlazaban con direcciones del fichero precensal y se han modificado o completado 1.890 direcciones, consiguiendo encontrar 37.661 personas empadronadas en ellas.

También se está añadiendo un número considerable de establecimientos colectivos nuevos, de todas las tipologías, pero especialmente del tipo “viviendas tuteladas”, aunque por sus características son más difíciles de distinguir de las viviendas familiares convencionales.

Se están investigando otras tipologías de centros, como residencias universitarias, centros de menores o residencias de trabajadores en los que figuran personas empadronadas en ellos.

Por otra parte, se dispone de un directorio obtenido de Catastro que proporciona las direcciones postales, situadas en el territorio fiscal común, con un uso catastral diferente del residencial (prescindiendo de otros usos en los que, a priori, se descarta que pueda haber establecimientos colectivos). Consta de 8.304 direcciones que incluyen, además de residencias de mayores y discapacidad, instituciones religiosas, hospitales de larga estancia, cuarteles, prisiones,... así como otro tipo de establecimientos fuera de los objetivos del censo como albergues, residencias de estudiantes, colegios mayores, centros de día...

Total	Residencias (mayores, discapacitados)	Instituciones religiosas	Militares/pen itenciarías	Otros
8.304	4.621	1.988	776	919

A partir de él se recuperarán otros posibles establecimientos colectivos. Según las primeras pruebas realizadas, será especialmente útil para localizar instituciones religiosas e instituciones militares no recogidas en otras fuentes.

Una vez investigado este fichero de posibles establecimientos colectivos surgen algunos inconvenientes que se están corrigiendo actualmente. Por ejemplo se han descartado usos como “*facultades, colegios y escuelas*” que no corresponden al concepto de establecimiento colectivo. Sin embargo, revisando ejemplos concretos en los ficheros de Catastro parece que algunas residencias de estudiantes están incluidas en esa tipología.

Una vez integrados ambos directorios, se dispondrá de un registro completo de establecimientos colectivos.

Las direcciones postales de estos establecimientos se cruzarán en una segunda fase con los ficheros precensales de viviendas y personas, de modo que se pueda identificar qué hogares del fichero censal corresponden a viviendas colectivas y el número y características de las personas empadronada en ellas.

Por último, queda pendiente un análisis más detallado de los hogares del fichero precensal que, no estando marcados como hogares colectivos en la fase anterior, presentan un número de miembros y/o edades anómalas para ser un hogar familiar.

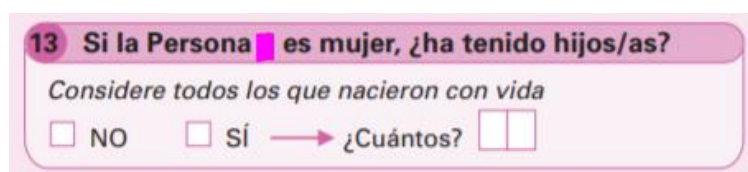
10 Incorporación de nuevas variables no investigadas en censos anteriores

Un censo de población y viviendas clásico se basa en el cuestionario como instrumento de recogida de información. El hecho de que los cuestionarios se envíen a todos los hogares y que por tanto se deban cumplimentar por los informantes sin ayuda de un entrevistador limita enormemente el contenido de la investigación. El cuestionario censal debe ser cuanto más corto y más fácil de entender mejor, porque con ello se conseguirá una mejor calidad de respuesta.

En el caso de un censo que, como éste, se construye a partir de registros administrativos, esta limitación desaparece; por tanto, la investigación puede abarcar conceptos difíciles de incluir en un cuestionario censal y en muchos casos muy demandados.

En la actualidad los trabajos del censo se encaminan prioritariamente a las variables obligatorias o de mayor tradición, mencionadas en apartados anteriores. Pero está prevista la incorporación de alguna información adicional. En la medida de lo posible, según el avance de los trabajos, podría ampliarse esta lista.

La primera de estas variables que se incorporaría como novedad es la de la fecundidad alcanzada. De hecho, esta información no es del todo novedosa. En el cuestionario de 2011 se incluía esta pregunta para cada miembro del hogar:



13 Si la Persona es mujer, ¿ha tenido hijos/as?
Considere todos los que nacieron con vida
 NO SÍ → ¿Cuántos?

Esta información permitía analizar la fecundidad alcanzada por mujeres de todas las generaciones.

En el censo de 2021 se va a incorporar información más detallada sobre fecundidad alcanzada. Dado que se dispone de datos de nacimientos con identificador del padre y de la madre (DNI, NIE) desde 1996 se pueden enlazar todos los nacimientos habidos desde entonces a su correspondiente padre y madre. Hay que tener en cuenta la limitación de que no siempre enlaza correctamente la información, que se contemplan solo los nacimientos recogidos en los boletines del INE (ocurridos en España, por tanto) y que solo se tienen desde ese momento, por lo que no se recoge la fecundidad alcanzada por todas las generaciones. Pero aun así se ofrece información mucho más completa que una simple pregunta sobre hijos tenidos.

10.1 NUEVAS VARIABLES DE CONTEXTO

Para el censo 2021 se contempla también la incorporación de información que, por distintas razones no se recoge a nivel individual, sino agregada. El ejemplo más sobresaliente es el de los ingresos.

No se puede obtener para el censo información individualizada del nivel de ingresos a partir de fuentes tributarias pero sí es posible incorporar información sobre el nivel de ingresos medios u otros indicadores de ingresos (como pueden ser datos de desigualdad) a nivel más agregado, para preservar mejor el secreto estadístico.

Se ha acordado con las fuentes tributarias proporcionar esa información a nivel de sección censal. Esta información “de contexto” sobre la población puede ser extremadamente útil como aproximación a la condición socioeconómica, algo muy demandado por los usuarios de censos anteriores.

Otro caso es el del consumo eléctrico de las viviendas. Tal como se recoge en el apartado 8.3.2 no es posible asociar a cada vivienda su consumo eléctrico individual pero sí se puede proporcionar esta información a nivel de sección censal.

Otro elemento muy novedoso que se está analizando de cara al censo 2021 es la posible incorporación de información sobre movilidad a través de la posición de los teléfonos móviles, como se menciona en el apartado 7.3.8

También cabe la posibilidad de incorporar otra información relativa al entorno físico dado que la población se encuentra completamente georreferenciada, por lo que se puede asociar a información geográfica de otras fuentes, como el SIOSE¹⁵. La cantidad y tipo de indicadores a incorporar no se ha decidido todavía.

Toda esta información (ingresos, consumo eléctrico, movilidad) se podrá proporcionar a niveles de sección censal y superior (municipios...) y se podrá incorporar también en los ficheros de microdatos asociados a cada persona como información de contexto o relativa a su entorno, es decir, no relativos a la persona o al hogar concreto, sino a la sección censal donde reside.

¹⁵ **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** a escala de referencia **1:25.000**, integrando la información disponible de las Comunidades Autónomas y la Administración General del Estado. Más información: www.siose.es

10.2 LA ENCUESTA COMPLEMENTARIA (ECEPOV-2021)

Como ya se ha mencionado a lo largo de este documento, para completar la información ofrecida por el censo, se ha diseñado una encuesta (se publicará un proyecto específico) denominada Encuesta sobre Características Esenciales de la Población y Viviendas, ECEPOV-2021. La encuesta tendrá un tamaño muestral efectivo en torno a 200.000 hogares lo que permitiría dar estimaciones no solamente a nivel provincial sino también para ámbitos más reducidos y en particular para ciudades a partir de determinado umbral en torno a 50.000 habitantes.

Hay algunas variables contenidas en censos anteriores que siguen siendo muy demandadas por los usuarios que no se pueden recoger bien a partir de registros.

Entre ellas cabe destacar:

- País de nacimiento de los padres, que permite estudiar las segundas generaciones de los inmigrantes
- Uso de lenguas cooficiales o de otras lenguas
- Equipamiento de las viviendas (calefacción, WC,...), información necesaria para el Reglamento a nivel provincial.

Por otra parte cabe señalar que si bien el tipo de hogar se va a construir de forma bastante aceptable en general, algunos tipos de hogar menos frecuentes pero de mucho interés demográfico no se van a poder identificar correctamente usando solo registros (parejas del mismo sexo, hogares reconstituidos,...). La ECEPOV-2021 se va a usar también para mejorar la imputación en ciertas variables del censo como la composición del hogar y el régimen de tenencia.

La ECEPOV-2021 se publicará como producto independiente del Censo 2021 y sustituiría ese año a la encuesta continua de hogares (ECH).

La información fundamental se tendría al máximo nivel de detalle geográfico con el producto principal, el propio Censo 2021; la que requiere más elaboración (segundas generaciones de inmigrantes, equipamiento de las viviendas, uso de lenguas cooficiales, disponibilidad y uso de segunda vivienda, relaciones de pareja, composición detallada del hogar y nuevos tipos de hogar) la proporcionaría esta encuesta a un también nivel bastante detallado dado que se entrevistará a más del 1% de la población española.

11 Evaluación de la calidad.

11.1 LA CALIDAD UN CENSO BASADO EN REGISTROS

Existe unanimidad entre la comunidad de productores de información censal en considerar a los países nórdicos como los más avanzados en este ámbito. Todos ellos realizan desde ya varias ediciones un censo basado en registros y cada vez son más los países que dan el salto a este modelo.

En España, la opción de un censo que se apoya en registros parece a día de hoy mejor que cualquiera de las alternativas basadas en la recogida directa por diversos motivos.

Hay que tener en cuenta que, a partir de los distintos registros, se dispone de información de casi todas las variables que se estudiaron en el Censo de 2011. Además, la gran mayoría de los registros disponen de identificadores de calidad que permiten la posibilidad de realizar enlaces entre las distintas fuentes con poco margen de error.

Para algunas variables se obtiene un nivel de detalle muy grande con un coste muy pequeño si lo comparamos con una metodología tradicional.

El censo que utiliza en todo lo posible registros elimina, o al menos atenúa, algunas de las desventajas más importantes de un censo basado en la toma directa de datos por parte de los ciudadanos, como son:

- El informante del cuestionario no siempre conoce todas las respuestas. Muchas veces es otra persona la que responde por él, esto se conoce como un *proxy* o informante indirecto. Es posible que esta persona no conozca de forma detallada algunas preguntas (como por ejemplo las relativas a variables económicas).
- Los resultados basados en auto-clasificación (ocupación, relación con la actividad) muchas veces no reflejan del todo bien la realidad.
- El cuestionario resulta difícil de rellenar para muchos informantes, lo que propicia errores de todo tipo o información incompleta.

Además, la realización de un censo tradicional o basado en encuestas se enfrenta a otros problemas que no están presentes en un censo como éste, o lo están en menor medida entre los que cabe destacar la dificultad sistemática para encontrar ciertos tipos de personas. Por ejemplo, una parte de la población extranjera siempre es más difícil de censar.

El censo basado en registros pasa a ser una labor transparente para la sociedad. El INE frecuentemente recibe quejas de los ciudadanos seleccionados para realizar encuestas en las que se le reclama que no se debería preguntar por cosas que "*la Administración ya conoce*". El censo así diseñado es también, la plasmación del esfuerzo por no molestar a los ciudadanos con información que la Administración ya tiene, aunque no resulte nada fácil de tratar en muchos casos por dificultades técnicas de todo tipo.

Por último, entre las ventajas que proporciona esta estrategia destaca la posibilidad de obtener información con una frecuencia mayor que la decenal pudiendo llegar a ser anual en determinadas situaciones. El esfuerzo de disponer de datos

administrativos puede dar frutos casi con la frecuencia que se desee, y casi con mismo coste, mientras que un censo tradicional apenas cabe plantearlo cada diez años (algunos países lo hacen cada cinco).

En fin, la conclusión a la que se llega por parte del INE es que el censo permitirá encontrar casi toda la información que figura en los cuestionarios, que lo hará con un nivel de calidad que a día de hoy es superior y que abrirá un enorme campo a la estadística pública. No debemos olvidar que, además de la información que se busca recoger para 2021, los registros pueden proporcionar en el futuro mucha información difícil de recoger en cuestionarios (salud, ingresos, discapacidad, ocio)

11.2 INDICADORES DE CALIDAD DEL FICHERO CENSAL

El fichero censal se construye asignando para cada individuo y cada variable un valor que se obtiene a partir de las distintas fuentes implicadas. Así, para el individuo i , el valor que toma la variable j lo denominamos V_{ij} . Pero para cada individuo y cada variable está previsto también asignar unos valores C_{ij} (que denominamos *indicador de calidad*) con la que se ha construido el valor V_{ij} .

Al hacer eso, el fichero censal, por tanto, estará compuesto por una dupla (V_{ij} , C_{ij}) para cada individuo y cada variable, como figura en el esquema siguiente.

Figura 10. El fichero censal de personas. Duplas de valores.

		Variables						
		1	2	j	n
Personas	1							
	2							
	...							
	...							
	.							
	i					V_{ij} / C_{ij}		
	m							

La forma en la que se construye este indicador de calidad tiene que ver con la fiabilidad que se asigna a cada fuente para cada variable. Por ejemplo, la variable *“nivel educativo alcanzado”* se le asigna a cada individuo a partir de un algoritmo en el que intervienen distintas fuentes.

Si para la persona i el nivel educativo alcanzado se ha construido a partir de una fuente directa del sistema educativo (por ejemplo, porque dispone de un título universitario) podemos asignar a esa variable para esa persona en concreto un valor $C_{ij}=1$

En caso de que para una persona el valor se haya asignado a partir de una fuente indirecta o de menor calidad, como por ejemplo, porque la única información que disponemos para esa persona es que refiere un determinado nivel educativo en el censo de 2001 o en el formulario de empadronamiento, podríamos asignar en ese caso $C_{ij}=2$.

Si no se dispone de ninguna referencia en ninguna fuente, el dato se habrá construido a partir de imputación, y en ese caso el indicador de calidad sería $C_{ij}=3$.

En resumen, para cada individuo y para cada variable, podemos tener una dupla formada por el valor que toma la variable en concreto y el indicador de calidad correspondiente, que podría tomar los valores 1,2 y 3¹⁶.

Un análisis de la matriz de indicadores de calidad nos da una información muy útil sobre la calidad del fichero censal en su conjunto. Si analizamos por filas (personas) esta matriz podemos ver, para cada individuo, cuál es la calidad de la información construida para él; si analizamos la información por columnas podemos ver, para cada variable, con qué calidad (y a partir de qué fuentes) se ha construido.

Se publicarán informes de calidad a partir del análisis, tanto vertical como horizontal, que ayuden a los usuarios a analizar la información censal.

Así mismo está previsto que en los microdatos que se publiquen (ver apartado 12 sobre difusión de la información) se proporcionarán los indicadores de calidad individual, lo que podría permitir al usuario especializado incluso deshacer imputaciones realizadas por el INE y sustituirlas por las suyas propias.

¹⁶ Todavía no se ha decidido por completo la escala de valores que se usará para los indicadores de calidad.

11.3 INFORMES DE CALIDAD

Como ya ocurriera con el Censo de 2011, Eurostat ha promovido la aprobación de legislación específica sobre la calidad del Censo de 2021. El reglamento (UE) N° 881/2017 trata aspectos relativos a la estructura que tienen que tener los informes que se realicen sobre la calidad de los datos y al formato técnico para la transmisión de dichos datos.

Esto permite contar con un marco de referencia para el seguimiento de la calidad de la operación censal a nivel europeo.

12 Difusión de resultados

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

- a. 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

España está a la cabeza en la cantidad de información que pone a disposición de los usuarios en materia censal, con la salvedad obligada por el carácter muestral de la operación de 2011. El contenido de la difusión del Censo 2021 será muy superior al del Censo 2011 debido a que la recuperación del carácter exhaustivo permitirá proporcionar más datos de ámbitos pequeños.

En cuanto a los productos que se prevé diseñar, se aprovechará la experiencia previa –del INE como productor y de los usuarios en la explotación de la información censal– para difundir los resultados mediante herramientas de tipo **data warehouse**, como ocurre actualmente con los Censos de 1991, 2001 y 2011.

En estos sistemas, con las restricciones necesarias para garantizar la confidencialidad de la información, los usuarios podrán diseñar las consultas “a la carta” que mejor se ajusten a sus necesidades.

Se dispondrá de servicios web que permitan realizar explotaciones automatizadas de la información para facilitar la reutilización de la información censal en otras aplicaciones.

Dada la disponibilidad de coordenadas geográficas para toda la población y la existencia de nuevas variables de entorno o de contexto a niveles territoriales detallados, está previsto potenciar considerablemente el sistema de difusión para navegación por mapas. Se prevén los siguientes tipos de acceso:

- **Mapas temáticos:** Las consultas de los usuarios en las que se incluyan cualquier nivel geográfico podrán ser representadas mediante mapas temáticos.
- **Rejilla censal (*Census grid*):** Eurostat está promoviendo ya desde 2011 la difusión de un conjunto de datos europeos armonizados basados en una red uniforme y atendiendo 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 la celda es constante y se conoce la posición en coordenadas del centro de cada celda, se puede decir que todas están georreferenciadas. Además de ser continuas para todo el territorio, se dispone de la gran ventaja de su forma invariable con lo cual permite un fácil estudio de una variable a lo largo del tiempo. Para ello es absolutamente necesario disponer de

los datos de los Censos también georeferenciados que se asocian por sus coordenadas a la celda correspondiente de la malla elegida.

Una novedad respecto del Censo de 2011 es la existencia de reglamentación específica para la información a difundir mediante esta rejilla¹⁷

- **Análisis por áreas definidas por el usuario:** En el censo de 2011 se ensayó una nueva forma de difusión de forma más libre que la rejilla de 1 km²¹⁸ mediante rejillas de tamaño variable en función de la cantidad de observaciones (hay que tener en cuenta que los datos provenían de una muestra).

Para 2021 se plantea generalizar el concepto de malla censal a cualquier otro polígono cerrado siempre que se superen umbrales de población contenida en ellos.

El sistema de difusión censal podrá permitir responder a preguntas como cuáles son las características de la población en un radio de 500 metros en torno a un determinado punto (si en ese círculo se supera ese umbral de población) o a una distancia dada de un tramo de carretera o de una playa.

Esta forma de difusión -no necesariamente basada en divisiones administrativas- puede realizarse dado que cada hogar tiene asignadas unas coordenadas GPS aproximadas (las del edificio donde habitan).

En particular, es especialmente interesante para el análisis de la población en grandes aglomeraciones urbanas donde el usuario de la información podrá escoger sus propias agregaciones de población, siempre condicionado por esos umbrales poblacionales antes mencionados de forma que se garantice la protección de la confidencialidad.

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.

¹⁷ 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

¹⁸ https://www.ine.es/censos2011_datos/rejilla_web.pdf

13 Calendario de tareas

Las tareas a acometer están condicionadas por el grado de éxito en la fase actual de cruce de registros. De forma tentativa se pueden establecer los siguientes hitos en el proyecto:

- Año 2020:

- Abril: publicación del proyecto definitivo en www.ine.es
- En el primer semestre de 2020 se dispondrá de una versión completa del FPC-2018.
- Se inician los trabajos del FPC-2020. Se deben identificar las posibles fuentes adicionales y decidir el ámbito final del censo en cuanto a variables a estudiar.

- Año 2021:

- Primer semestre: encuesta ECEPOV dirigida a la población
- Octubre 2021-abril 2022. Cruces con registros administrativos, aplicación del método de signos de residencia.

- Año 2022. Incorporación de resto de fuentes. Elaboración de variables censales.

- 4º trimestre 2022. Resultados preliminares del Censo

- Año 2023 (primer semestre):

- Publicación de resultados completos el censo tanto para fines nacionales como para satisfacer las demandas de información internacional.

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

El método mediante el cual se está construyendo 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 censos cada diez años cuando se dispone de información tan rica y detallada 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. Incluso cabe preguntarse si la propia denominación de “censos” para este tipo de información sociodemográfica tiene sentido. Si observamos lo que han hecho otros países que han realizado ya la transición a basar el censo en registros administrativos, se puede observar que, de hecho, el término “censo” solo se usa para referirse a los censos históricos. Por ejemplo, en el caso de Suecia, el último censo que mencionan en su página web es el de 1991. A partir de ahí los productos que se ofrecen son distintas estadísticas de población, viviendas y hogares, derivados de registros administrativos.

Estos países (ocurre también en el caso de Holanda, Dinamarca y otros) generan un producto estadístico llamado “censo de población y viviendas” solo para su entrega a Eurostat dado que hay un reglamento que así lo exige, pero no muestran a sus usuarios en sus ámbitos nacionales la información bajo tal denominación.

Cabe pensar que en España se pueda seguir la misma aproximación. No habría así, un censo 2031 sino que a partir de 2021 se podrían diseñar nuevos productos sociodemográficos basados en registros administrativos que proporcionen información de detalle temático y geográfico con una frecuencia mucho mayor.

Esa estrategia no es fácil de diseñar. Por poner un ejemplo, 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.

14.1 EL MARCO EUROPEO

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 cual 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.

Se creó una Task Force en el seno del grupo de trabajo de Censos en Eurostat en la que específicamente se ha iniciado el debate sobre el futuro más allá de 2021.

Se preparó un primer documento de visión, en el que se señalaba estos dos elementos:

- Los usuarios claramente valoran como un punto débil de los censos de 2011 en la UE que los datos solo estén disponibles de manera decenal y que se publiquen tres años después de la fecha de referencia. Las consultas con usuarios mostraron la necesidad de una mayor frecuencia y puntualidad en la publicación de datos similares a los del censo, especialmente en temas de población y migración.
- La gran mayoría de los países europeos han realizado, o están planeando o considerando, cambios importantes en las fuentes de datos y métodos utilizados para el censo; el mayor uso de datos de fuentes administrativas y el alejamiento de la enumeración tradicional de censos puerta a puerta. Esto significa que es factible avanzar hacia una producción de datos más rápida y frecuente.

Los Estados Miembros y Eurostat acordaron, en la reunión del grupo de Directores de Estadísticas Sociales (DSS) de octubre de 2016, y pidieron a Eurostat preparar una estrategia que proporcionara más detalles sobre la puesta en marcha de esta visión, teniendo en cuenta la necesidad de un nuevo marco legal.

El memorándum de Budapest¹⁹ (septiembre de 2017) es un documento de estrategia firmado por los responsables máximos de los institutos de estadística europeos y Eurostat. En él se establece: *“Como parte de la estrategia del censo posterior a 2021, desarrollar y promover una compilación anual creciente de datos de tipo censal como fuente de información sobre las características demográficas, sociales y económicas de los migrantes y de la población en general, teniendo en cuenta la disponibilidad de fuentes de datos adecuadas y garantizar sinergias y coherencia con las estadísticas demográficas y de migración existentes.”*

Desde entonces se han iniciado los trabajos para definir qué tipo de información sería necesario proporcionar cada año, o con otra periodicidad, y con qué plazos en cada caso.

19

<https://ec.europa.eu/eurostat/documents/7330775/7339365/FINAL+Budapest+memorandum.pdf/96a6db89-1395-44a5-8a46-85e8c49d576c>

La nueva recopilación de datos anual debería contemplar las actuales recopilaciones de datos que no forman parte del censo (demografía, migración y estadísticas de ciudades), además de agregar nuevos desgloses cuando sea necesario a la luz de la evolución estadística y las prioridades cambiantes.

A día de hoy, los Estados Miembros (EE.MM.) publican y envían a Eurostat información anual sobre demografía (stocks de población, eventos demográficos como nacimientos defunciones...), sobre migraciones, etc. Se pretende que las nuevas recopilaciones de datos tengan en cuenta las recopilaciones de datos relacionadas con la población existentes.

En particular, dado que la migración sigue siendo una alta prioridad para los encargados de formular políticas, se prevé un aumento de las estadísticas sobre migrantes y asuntos relacionados con la migración.

En el momento de elaborar este proyecto de Censos 2021 no se dispone de una propuesta concluyente de información estadística clasificada por tipo y frecuencia más allá de algunas cuestiones generales que parecen ya acordadas como que el nuevo reglamento debería entrar en vigor con el fin de que la primera entrega de información por parte de los EE.MM. se produzca en 2024 y que se proporcionará información cada año, cada cinco años y cada diez años.

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 cabe ir más allá en la producción de información sociodemográfica a partir de las nuevas fuentes.

De forma tentativa se mencionan las líneas estratégicas que se contemplan en la actualidad para las estadísticas demográficas del INE.

En primer lugar, dado que el censo de 2021 es, en definitiva, un registro estadístico de población y viviendas con fecha 1 de enero de 2021, el primer objetivo es mantener esa actividad mediante la puesta en marcha de **una nueva operación estadística que se podría denominar “Registro Estadístico de Población y Viviendas” (denominación provisional: REPV)**. Este registro sería la fuente a partir de la cual se tendrían estadísticas anuales de población (las actuales operaciones Cifras de Población y Estadística de Migraciones, que seguirían existiendo).

Este REPV se iniciaría en 2022 y podría dar lugar a una **explotación estadística** más allá de las dos ya mencionadas. El Registro contendría también información agregada a niveles de sección censal o superiores sobre distribución de renta o consumo eléctrico de las viviendas. Esta información podría dar lugar a publicaciones específicas separadas del estilo de la actual estadística experimental **Atlas de Distribución de Renta de los hogares** más allá del censo 2021.

Se va a proponer la **desaparición de la actual Encuesta Continua de Hogares**, que se dejaría de recoger en diciembre de 2020 y su sustitución por una nueva estadística, no basada en trabajos de campo sino en la explotación longitudinal del Registro de Población antes mencionado. Esta nueva estadística se podría iniciar en 2023 con referencia a 2022. Contendría parte de la información de la actual ECH pero añadiría una perspectiva longitudinal afrontando temas nuevos como la formación, modificación y disolución de hogares (emancipación, duración de las parejas, formas de convivencia...). Se trataría de una estadística anual basada, en principio, completamente en el REPV.

La puesta en marcha del Registro Estadístico de Población y Viviendas podría mejorar a las actuales estadísticas en varias direcciones:

- Por un lado, al disponer en ese Registro de toda la población perfectamente georreferenciada, se abre la puerta a proporcionar datos de estadísticas como cifras de población, migraciones, o movimiento natural de la población, no solo por divisiones administrativas (municipios, provincias...) sino también usando rejillas de 1km² u otras. Además se espera que haya demanda de información en ese sentido por parte de Eurostat
- Se va a poder incorporar información longitudinal en las estadísticas demográficas. Por ejemplo, duración de los episodios migratorios, tiempo transcurrido entre eventos. Esto permitirá también hacer seguimientos de cohortes para la realización de estadísticas experimentales (por ejemplo, seguimiento de la muestra de una encuesta para ver a lo largo del tiempo su comportamiento)

- Se podrá incorporar a otras estadísticas información de variables contenidas en el REPV, tales como nivel educativo, relación con la actividad o movimientos migratorios pasados.

La existencia de nuevas fuentes administrativas no debería hacernos olvidar la importancia de la recogida de datos en operaciones de campo, necesaria para complementar y contrastar a los registros. Se considera necesario mantener un sistema de encuestas demográficas (fecundidad, migraciones) con cierta periodicidad.

15 Estimación de costes

En tanto que la operación censal no requiere de la realización de ninguna operación de campo, los costes se refieren fundamentalmente a recursos humanos empleados en los Servicios Centrales del INE y a la infraestructura informática para el desarrollo y la difusión.

Los costes totales de la operación para el periodo 2020-2023 se estiman en **3,2 millones de euros**. (507,000 euros anuales de costes de personal y 1,2 millones de euros de infraestructura y desarrollo TIC).

Esta estimación no incluye los costes de la encuesta ECEPOV-2021, que será objeto de proyecto específico.

Anexo. Lista de variables de los Censos de Población y Viviendas 2021

A continuación se relaciona, agrupado por categorías, un listado no exhaustivo de las principales variables que formarán parte del Censo 2021. El número final de variables no está del todo determinado y será superior al que aquí se detalla.

Con efectos de simplificar este documento, no se han incluido todas las variables derivadas (por ejemplo, existirían muchas variables derivadas con relación a la composición de los hogares y los núcleos familiares), sino solamente aquellas que o bien son obligatorias según la Reglamentación Comunitaria o bien destacan por su importancia.

Grupo 1-Características de las personas (28 variables: 15 obligatorias)

1.A-VARIABLES DEMOGRÁFICAS BÁSICAS (9 variables: 7 obligatorias)

<i>Nombre de la variable</i>	Obligatoria Eurostat	<i>Fuente principal</i>
1. Sexo	SI	Registros
2. Fecha de nacimiento	NO	Registros
3. Edad	SI	Registros
4. Estado civil legal	SI	Registros
5. País/Lugar de nacimiento	SI	Registros
6. País de nacionalidad	SI	Registros
7. Lugar de residencia	SI	Registros
8. Población en función del tamaño de los núcleos	SI	Registros
9. Número de hijos nacidos vivos	NO	Registros

1.B-VARIABLES RELATIVAS A MIGRACIONES (14 variables: 2 obligatorias)

<i>Nombre de la variable</i>	Obligatoria Eurostat	<i>Fuente principal</i>
1. Año de llegada a España	SI	Registros
2. País de procedencia	NO	Registros
3. Año de llegada a la comunidad autónoma	NO	Registros
4. Comunidad autónoma de procedencia	NO	Registros
5. Año de llegada a la provincia	NO	Registros
6. Provincia de procedencia	NO	Registros
7. Año de llegada al municipio	NO	Registros
8. Municipio de procedencia	NO	Registros
9. Lugar de residencia hace 1 año	NO	Registros
10. Relación residencia actual y hace 1 año	SI	Registros
11. Lugar de residencia hace 10 años	NO	Registros
12. Relación residencia actual y hace 10 años	NO	Registros
13. Lugar de residencia anterior	NO	Registros
14. Relación residencia anterior y actual	NO	Registros

1.C-VARIABLES RELATIVAS A LA EDUCACIÓN (3 VARIABLES: 1 OBLIGATORIA)

<i>Nombre de la variable</i>	<i>Obligatoria Eurostat</i>	<i>Fuente principal</i>
1. Nivel de estudios alcanzado	SI	Registros
2. Tipo de estudios completados	NO	Registros
3. Estudios en curso	NO	Registros

1.D-CARACTERÍSTICAS RELATIVAS A LA ACTIVIDAD ECONÓMICA (5 VARIABLES; 5 OBLIGATORIAS)

<i>Nombre de la variable</i>	<i>Obligatoria Eurostat</i>	<i>Fuente principal</i>
1. Relación actual con la actividad	SI	Registros
2. Sector de ocupación	SI	Registros
3. Situación de actividad actual	SI	Registros
4. Situación en el empleo	SI	Registros
5. Ubicación del lugar de trabajo	SI	Registros

GRUPO 2-VARIABLES RELACIONADAS CON LOS NÚCLEOS FAMILIARES Y LA ESTRUCTURA DEL HOGAR (13 VARIABLES: 6 OBLIGATORIAS)

<i>Nombre de la variable</i>	<i>Obligatoria Eurostat</i>	<i>Fuente principal</i>
1. Situación en cuanto a la familia	SI	Registros
2. Tipo de núcleo familiar	SI	Registros
3. Tamaño del núcleo familiar	SI	Registros
4. Número de núcleos en el hogar	NO	Registros
5. Situación en cuanto al hogar	SI	Registros
6. Tipo de hogar privado	SI	Registros
7. Tamaño del hogar privado	SI	Registros
8. Número de familias en el hogar	NO	Registros
9. Número de generaciones en el hogar	NO	Registros
10. Convivencia con padre	NO	Registros
11. Convivencia con madre	NO	Registros
12. Convivencia con cónyuge/pareja	NO	Registros
13. Convivencia con otro pariente	NO	Registros

Grupo 3-Características de las viviendas (16 variables: 10 obligatorias)

<i>Nombre de la variable</i>	<i>Obligatoria Eurostat</i>	<i>Fuente principal</i>
1. Régimen de tenencia de la vivienda	SI	Registros ECEPOV
2. Tipo de vivienda donde se reside	SI	Registros
3. Tipo de vivienda familiar convencional	SI	Registros
4. Número de ocupantes de la vivienda	SI	Registros
5. Superficie útil	SI	Registros
6. Superficie media por residente	SI	Registros
7. Suministro de agua	SI	ECEPOV
8. Instalación de retretes	SI	ECEPOV
9. Instalaciones de baños	SI	ECEPOV
10. Tipo de calefacción	SI	ECEPOV
11. Servicio de acceso a Internet	NO	ECEPOV
12. Problemas de aislamiento	NO	ECEPOV
13. Separación de residuos	NO	ECEPOV
14. Problemas zona donde está ubicada la vivienda	NO	ECEPOV
15. Infraestructuras en la zona donde está ubicada la vivienda	NO	ECEPOV
16. Consumo eléctrico medio de la zona en donde está ubicada la vivienda	NO	Registros

Grupo 4- Características de los edificios (15 variables: 2 obligatorias)

4.1 Datos del edificio (7 variables: 2 obligatorias)

<i>Nombre de la variable</i>	Obligatoria Eurostat	<i>Fuente principal</i>
1. Viviendas según tipo de edificio	SI	Registros
2. Período de construcción del edificio	SI	Registros
3. Número de plantas del edificio (sobre rasante)	NO	Registros
4. Número de plantas bajo rasante	NO	Registros
5. Número de inmuebles del edificio	NO	Registros
6. Estado del edificio	NO	Registros/ ECEPOV
7. Coordenadas del edificio	NO	Registros

4.2 Instalaciones del edificio (8 variables: 0 obligatorias)

Nombre de la variable	Obligatoria Eurostat	<i>Fuente principal</i>
1. Accesibilidad del edificio	NO	ECEPOV
2. Ascensor	NO	ECEPOV
3. Disponibilidad de garaje	NO	Registros ECEPOV
4. Número de plazas de garaje	NO	Registros ECEPOV
5. Gas	NO	ECEPOV
6. Dispositivo de energía renovable	NO	ECEPOV
7. Agua caliente central	NO	ECEPOV
8. Tipo de sistema de evacuación de aguas residuales	NO	ECEPOV