

INSTITUTO NACIONAL DE ESTADÍSTICA



Censo de Población

Metodología

Versión provisional

Abril 2024

Índice

1	Introducción	3
2	Objetivos	3
3	Conceptos y definiciones	6
4	El censo de población	7
4.1	La cifra de población censal: el método de signos de vida	7
4.2	Variables demográficas: sexo, edad, nacionalidad, lugar de nacimiento	13
4.3	Variables educativas: nivel de estudios alcanzado, estudios principales en curso, lugar de estudios	14
4.4	Variables laborales: relación con la actividad económica	19
5	Enlaces de registros administrativos con padrón	21
6	Difusión de los resultados	23

1 Introducción

A partir de la publicación de los Censos de Población y Viviendas 2021¹, se produce un cambio de paradigma en las estadísticas demográficas, puesto que la metodología empleada en el censo de 2021, basada por primera vez enteramente en el tratamiento y la explotación de decenas de registros administrativos, permite reproducir cada año el proceso para proporcionar censos de población anuales.

Gracias a esta nueva metodología, se podrá disponer de información actualizada cada año, en el caso de la población, y cada tres o cuatros años en el caso de las viviendas, en lugar de cada diez, como sucedía con los censos decenales que se venían elaborando hasta ahora. Además, estos nuevos censos anuales se elaboran con información recogida en su totalidad a partir de registros administrativos, sin necesidad de encuestar a ningún hogar.

Así, surge como una nueva operación estadística de carácter estructural y de periodicidad anual, el “**Censo de población**”, con código **8892** en el Plan Estadístico Nacional 2021-2024, y con código **30281** en el Inventario de Operaciones Estadísticas. Esta operación será independiente del censo de viviendas, que tendrá menor frecuencia.

Como norma general, a partir de 2024, los primeros resultados de los Censos de población se publicarán en diciembre de cada año, para la fecha de referencia a 1 de enero de ese mismo año.

Sin embargo, en esta primera edición se han publicado conjuntamente los primeros resultados de los censos a 1 de enero de 2022 y de 2023.

Los llamados “Primeros resultados” ofrecen las cifras oficiales de la población residente en España a 1 de enero con un gran detalle territorial (hasta sección censal), junto con sus principales características demográficas.

En los próximos meses se completará esta información demográfica con variables de temática educativa y laboral, entre otras, procedentes de la explotación de ficheros administrativos temáticos proporcionados por distintos organismos.

2 Objetivos

Entre los objetivos que pretende cubrir la operación Censo de población, cabe destacar los tradicionales de un censo:

- a) Recuento de la población. El Censo de Población determina la cifra oficial de población de España y de todas sus divisiones administrativas (las comunidades autónomas, provincias, municipios e incluso distritos y secciones censales) a 1 de enero de cada año.

¹ La publicación del Censo de Población y Viviendas 2021 se completó el 30 de junio de 2023. Se puede acceder a los resultados en

https://www.ine.es/dynqs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176992&menu=resultados&idp=1254735572981#!tabs-1254736195710

- b) Conocimiento de la estructura de la población. El Censo de Población no solo cuenta el número de habitantes sino que investiga, para cada uno de ellos, una serie de características demográficas, económicas y sociales y proporciona una imagen de la estructura de la población.
- c) 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.

Concretamente a nivel europeo, en la fecha de esta primera publicación de resultados, el marco legislativo que regula las estadísticas demográficas está siendo objeto de renovación.

Desde hace varios años los organismos competentes en el sistema estadístico europeo vienen trabajando en un nuevo reglamento que unifique la demanda de estadísticas de stocks y flujos demográficos, que ahora mismo están dispersas en tres reglamentos distintos¹ (uno cubre las estadísticas migratorias anuales, otro los fenómenos demográficos y stocks anuales y un tercero es el de los censos decenales).

En el momento de elaborar este documento metodológico no se dispone aún de una propuesta concluyente pero ya hay un borrador suficientemente maduro, que situaría la mayoría de las actuales demandas del censo con una frecuencia anual. En general, el nuevo reglamento europeo de población, previsiblemente, va a demandar un **mayor detalle territorial** (dando especial importancia a la georreferenciación), una **mayor frecuencia** y una **mayor oportunidad** en la información sobre la población y fenómenos demográficos asociados.

Esta necesidad de información demográfica más frecuente y rápida, que ya existía en la sociedad, se ha visto acrecentada por los acontecimientos de los últimos años (el Brexit, la pandemia de COVID-19, la guerra de Ucrania...). Con los requerimientos de información vigentes se ha comprobado que se tardaba demasiado en conocer el impacto demográfico que tenían.

Los censos anuales permitirán dar respuesta a estos requerimientos más exigentes de datos de población.

Pero más allá de los objetivos tradicionales que cubre la información censal, hay que destacar que esta nueva estrategia basada en registros para construir la información censal supone una renovación global en la información demográfica:

- d) La disponibilidad de información censal anual ha originado un nuevo sistema de estadísticas demográficas, coherentes entre sí, permitiendo eliminar la duplicidad de información demográfica que existía.

Tradicionalmente, tras la publicación de cada censo decenal, se iniciaba una serie estadística de población. Concretamente, a partir de los Censos de Población y Viviendas de 2011, y comenzando el 1 de enero de 2012, se venía publicando la

¹ En el momento actual, las estadísticas demográficas están reguladas por tres reglamentos de Eurostat del Parlamento Europeo y del Consejo: el Reglamento 1260/2013 sobre estadísticas demográficas europeas, el Reglamento 862/2007 sobre las estadísticas comunitarias en el ámbito de la migración y la protección internacional, y el Reglamento 763/2008 relativo a los censos de población y vivienda

serie estadística Cifras de Población, estimada mediante una contabilidad de sucesos demográficos (nacimientos, defunciones, migraciones interiores y exteriores y adquisiciones de nacionalidad española por parte de población extranjera).

Adicionalmente, la publicación cada diez años de un nuevo censo conllevaba una corrección de la serie de datos de población de todo el decenio previo hasta el censo anterior (las tradicionalmente llamadas estimaciones intercensales). Con la disponibilidad de censos anuales, este sistema de correcciones intercensales decenales deja de ser necesario.

Hasta ahora se venían publicando cada año dos juegos de estadísticas de población: las derivadas de la explotación del padrón y las que podemos denominar “estadísticas”, que se construían a partir del censo de 2011. Concretamente, coexistían dos juegos de datos de stocks de población:

- Las Cifras de Población, semestrales (con fechas de referencia a 1 de enero y 1 de julio).
- La Estadística del Padrón Continuo, anualmente (con fecha de referencia a 1 de enero).

A partir de 2023 han desaparecido ambas operaciones, siendo sustituidas por los ya mencionados censos de población anuales y la puesta en marcha, en mayo de 2023 de una nueva estadística de stock de población coyuntural, la **Estadística Continua de Población**, completamente coherente con los censos de población anuales

De esta manera, se dispone no solo de información censal a 1 de enero de cada año con el máximo detalle territorial, sino también de estimaciones trimestrales de población, muy rápidas en el tiempo (con una demora de un máximo de 45 días respecto a la fecha de referencia, frente a los 5 o 6 meses de las antiguas Cifras de Población) y totalmente congruentes con las cifras censales, aunque con menor detalle territorial (hasta provincia /isla).

Esta nueva operación ofrece además información sobre la población residente en viviendas familiares y el número y tamaño de los hogares.

Así se da respuesta a la demanda por parte de los usuarios, tanto externos como internos del INE, de unificar las cifras de población en una sola serie de población consistente a todos los niveles geográficos, evitando las dudas sobre qué lote de cifras utilizar.

La creación de este nuevo sistema de estadísticas demográficas basadas en los censos de población anuales afecta también a las estadísticas de migraciones.

Las antiguas operaciones Estadística de Migraciones y Estadística de Variaciones Residenciales, ambas basadas en el conteo y tratamiento estadístico de ciertos tipos de variaciones residenciales registradas en el padrón continuo, desaparecen.

En su lugar, se tratarán todas las variaciones registradas que puedan tener correspondencia con un cambio de residencia entre dos momentos censales y, a la vez, coherentes con cada dos censos consecutivos. Para responder a estas características, se publicará por primera vez a finales de 2023, la **Estadística de Migraciones y Cambios de Residencia**, que sustituirá a las dos anteriores.

Esta nueva estadística recogerá las migraciones, tanto exteriores como dentro del país, según las definiciones de las recomendaciones internacionales, y desagregadas por un mayor nivel territorial. Además permitirá, en un futuro, el estudio de los cambios de residencia intramunicipales, por ejemplo entre los barrios de una ciudad.

e) La información censal anual permite enriquecer otras operaciones

Los ficheros censales tienen un gran potencial de explotación en sí mismos, por la cantidad de información que contienen para caracterizar la población. Pero además se pueden cruzar a nivel individual con otras operaciones estadísticas. Esto ya se viene haciendo desde hace años en las operaciones del Movimiento natural de la población. Así, a partir de los ficheros censales se añaden características como el nivel de estudios, la ocupación, la relación con la actividad o el estado civil a las madres y padres de los nacidos, o de las defunciones. Este potencial de explotar las características individuales se aprovechará en más operaciones, como la nueva Estadística de Migraciones y Cambios de Residencia y permitirá conocer la naturaleza de los movimientos migratorios de una forma cualitativa, y no solo cuantitativa.

3 Conceptos y definiciones

Población residente: personas, de cualquier nacionalidad, que tienen fijada su "residencia habitual" en el territorio nacional.

Residencia habitual: Lugar donde una persona normalmente pasa los periodos diarios de descanso, sin tener en cuenta las ausencias temporales por viajes de ocio, vacaciones, visitas a familiares y amigos, negocios, tratamiento médico o peregrinación religiosa. No obstante, se ha de puntualizar que sólo se considerarán residentes habituales en una circunscripción territorial:

- Aquellos que, según la definición anterior, hubieran residido habitualmente en la misma por un periodo continuado de al menos doce meses.
- Aquellos que, según la definición anterior, hubieran establecido su residencia habitual en la misma hace menos de doce meses pero tuvieran intención de permanecer en ella al menos un año.

Cuando no se puedan determinar las circunstancias anteriores, se entenderá por "residencia habitual" el lugar de residencia registrada.

Sexo: El sexo se refiere al sexo biológico de la persona.

Edad: La edad en años cumplidos se refiere al número de aniversarios del nacimiento transcurridos en la fecha de referencia, es decir, a la edad en el último cumpleaños.

Nacionalidad: Vínculo jurídico particular entre un individuo y su Estado, adquirido mediante nacimiento, naturalización, declaración, opción, matrimonio y otras fórmulas acordes con la legislación nacional. Se refiere a la nacionalidad en la fecha de referencia, clasificándose como española siempre que el individuo tenga dicha nacionalidad, pudiendo tener además otras.

Lugar de nacimiento: Lugar en que ha nacido una persona, es decir, el lugar de residencia habitual de la madre en el momento del nacimiento, determinado en la fecha de la recogida de datos.

Nivel de estudios alcanzado: El nivel de formación alcanzado por una persona se define como "el máximo nivel que el individuo ha completado" y "se mide a partir del programa de máximo nivel que haya superado, validado, normalmente, por la correspondiente certificación".

Estudios principales en curso: Se proporciona una "Enseñanza principal" que generalmente se corresponde con la enseñanza de Régimen General (E.S.O, Bachillerato, F.P, estudios universitarios etc.) de mayor nivel que está cursando la persona (si cursa más de una).

Tipo de estudios en curso: Campo de estudios universitarios del estudio principal en curso de un estudiante.

Lugar de estudios: Lugar de estudios registrado de un estudiante en relación con el lugar de residencia habitual.

Relación con la actividad económica: Situación actual de la población de 16 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.

4 El censo de población

4.1 LA CIFRA DE POBLACIÓN CENSAL: EL MÉTODO DE SIGNOS DE VIDA

El primer objetivo de los censos es contar el número de habitantes que reside en un territorio en una fecha de referencia, que en el caso de los censos de población anuales es el 1 de enero de cada año.

La cifra de población censal debe ser considerada como la **cifra oficial de población a todos los niveles territoriales**.

El método de estimación de la población residente en los censos anuales es el ya usado para el censo de población de 2021: el método de "signos de vida" o "signos de presencia".

Estos métodos consisten esencialmente en partir de un registro de población, en nuestro caso el Padrón, y evaluar la verosimilitud de la residencia de cada persona a partir de su presencia en otros registros, de manera que para cada individuo se decide si se cuenta como residente o no, de forma más precisa que la contabilización administrativa del Padrón.

Si bien el método de signos de vida supone una mejora de calidad de la cifra de población censal, **el elemento clave del proceso es la existencia de un registro de población, el Padrón**, cuya utilización como "esqueleto" del censo de población supone grandes ventajas respecto a otros posibles registros de población alternativos (registros tributarios, tarjetas sanitarias o ficheros de la D.G. Policía). Todas las

personas residentes tienen la obligación de inscribirse en el padrón del municipio donde viven la mayor parte del año, registrándose incluso población que en principio pudiera parecer de difícil acceso, como personas sin residencia legal o empadronadas en viviendas no convencionales. Además desde la implantación hace más de 25 años del Padrón Continuo, con el INE como organismo coordinador de los más de 8.000 padrones municipales, el Padrón ha madurado mucho como registro de población estableciendo, entre otros, procedimientos de caducidad y comprobación periódica de la residencia de los extranjeros que hacen del Padrón un registro de población actualizado, fiable y preciso.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

- Fuentes relativas al mercado laboral

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

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

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

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

– Otras fuentes complementarias

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

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

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

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

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

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

El método de signos de vida utilizado en los censos de población anuales

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

Además, se aplica a nivel de hogar, de modo que solo se dejan de contar hogares en los que de forma inequívoca todos los miembros del hogar están afectados o bien el hogar resultante no tiene una composición alterada, por ejemplo, los hogares en que no hay relaciones de parentesco entre sus miembros.

El padrón municipal funciona en la actualidad aplicando a la población extranjera unos procedimientos de renovación y comprobación periódica de la residencia que permiten mantener razonablemente actualizado el fichero. Es algo parecido a lo que hacen otros países con registros de población o a cómo funcionan registros como el DNI, que obligan a la renovación cada cierto tiempo.

La población extranjera, a efectos de la gestión padronal, se divide en dos grupos, “ENCARP” y “no ENCARP”. Los extranjeros no comunitarios sin autorización de residencia permanente (abreviados ENCARP), son las personas que residen en España sin tener la nacionalidad española ni la de ningún otro Estado de la Unión

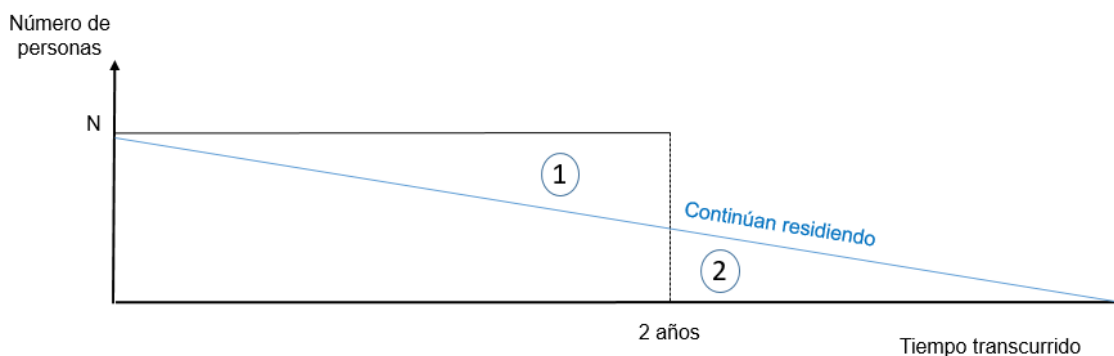
Europea y que no disponen de una autorización de residencia permanente expedida por la Dirección General de la Policía. Desde 2005 estas personas deben acudir a sus respectivos ayuntamientos para renovar su inscripción en el padrón municipal cada dos años. De no hacerlo se pone en marcha un proceso que termina dando lugar a su baja padronal por caducidad (llamada BBC en terminología del padrón).

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

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

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

Esquema del recuento de población extranjera según fecha de caducidad de su inscripción



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

Però la población que inmigra abandona el país de forma más parecida a lo que muestra la línea azul descendente. El triángulo 2 contiene la población que sigue viviendo en España pero cuya inscripción ha caducado. Si cruzamos esta población con los ficheros de signos de vida (laborales, tributarios) encontramos extranjeros que inicialmente

estarían “caducados” pero que por presentar signos suficientes de presencia son “rescatados” y contabilizados en la cifra de población censal.

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

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

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

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

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

La población centenaria

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

errores en la certificación de la defunción que impiden dar de baja a ningún empadronado.

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

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

Para paliar estos defectos, primeramente se comprueban las fechas de nacimiento de la población centenaria cruzando con la base de datos de DNIs y efectuando un pequeño número de depuraciones.

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

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

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

4.2 VARIABLES DEMOGRÁFICAS: SEXO, EDAD, NACIONALIDAD, LUGAR DE NACIMIENTO

En esta primera publicación de resultados se ofrecen solamente las principales variables demográficas de la población.

Esta información se obtiene fundamentalmente a partir de los datos que figuran en el padrón.

Son necesarias pequeñas depuraciones, que afectan a un número residual de registros, en el país de nacimiento y nacionalidad, aplicando aplica un proceso de imputación cuando no se encuentra la información en otros registros administrativos.

También se realiza un proceso de detección de posibles incoherencias en la variable sexo contrastando con fuentes del Ministerio del Interior, así como de la fecha de nacimiento.

Adicionalmente, dado que el lugar de nacimiento se define como el lugar de residencia habitual de la madre en el momento del nacimiento, ha sido necesario buscar esta información en el padrón, quedándonos, cuando fuese posible, con el lugar de residencia del nacido (que debe coincidir con el de la madre) en vez del lugar de nacimiento.

4.3 VARIABLES EDUCATIVAS: NIVEL DE ESTUDIOS ALCANZADO, ESTUDIOS PRINCIPALES EN CURSO, LUGAR DE ESTUDIOS

Posteriormente a la primera publicación del censo de población, se proporcionan las variables de naturaleza educativa para la población de 15 años y más.

Esta información procede de numerosas fuentes administrativas con información personal de temática educativa, que se combinan en el censo para obtener las variables educativas de las personas de 15 años y más.

Detalle de la clasificación de las variables educativas:

El **nivel de estudios alcanzado** se desglosa en base a una adaptación de la CNED-A¹ con un detalle máximo de doce categorías:

- Analfabetos
- Estudios primarios incompletos
- Educación primaria
- Primera etapa de educación secundaria y similar
- Segunda etapa de educación secundaria con orientación general
- Segunda etapa de educación secundaria con orientación profesional
- Educación postsecundaria no superior
- Enseñanzas de formación profesional, artes plásticas y diseño y deportivas de grado superior y equivalentes
- Grados universitarios de hasta 240 créditos ECTS, diplomados universitarios y equivalentes
- Grados universitarios de más de 240 créditos ECTS, licenciados y equivalentes
- Másteres, especialidades en Ciencias de la Salud por el sistema de residencia y similares
- Doctorado universitario

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

¹

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

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

Los **estudios principales en curso** se desglosan en base a una adaptación de la CNED-P¹ con un detalle máximo de doce categorías:

- Educación primaria e inferior
- Primera etapa de educación secundaria y similar
- Educación Secundaria obligatoria (4º curso)
- Enseñanzas de bachillerato y similar
- Enseñanzas de formación profesional, artes plásticas y diseño y deportivas de grado medio y equivalentes
- Enseñanzas de Escuelas Oficiales de Idiomas
- Enseñanzas profesionales de música y danza y similares
- Enseñanzas de formación profesional, artes plásticas y diseño y deportivas de grado superior y equivalentes
- Grados universitarios de hasta 240 créditos ECTS, diplomaturas universitarias y equivalentes
- Grados universitarios de más de 240 créditos ECTS, licenciaturas y equivalentes
- Másteres, especialidades en Ciencias de la Salud por el sistema de residencia y similares
- Doctorado

No se incluyen enseñanzas de educación no formal.

El **tipo de estudios en curso** se desglosa en base a la CNED-F1 con un detalle máximo de diez categorías:

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

1

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

Fuentes de las variables educativas:

El nivel educativo alcanzado se corresponde con **el más alto encontrado de entre todos los registros de la persona**, y se dan a la vez estas dos circunstancias:

- Lo más común es encontrar a una persona en más de una fuente de temática educativa. Esto le ocurre aproximadamente al 85% de las personas de 15 años y más.
- Ningún registro es a la vez completo, actual, y ofrece la información con la desagregación requerida.

Un buen ejemplo de esta dicotomía es el Padrón como fuente de temática educativa. Toda persona que se empadrona debe rellenar un formulario de empadronamiento (hoja padronal) que permite proporcionar los datos según la CNED-A requerida. No obstante, no toda la población figura con ese nivel de detalle.

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

Aproximadamente para un 75% de la población se dispone del nivel detallado y en el restante 25% solo figuran cuatro valores posibles de nivel de estudios agregados. Adicionalmente, esta información puede quedar obsoleta salvo que la persona cambie su empadronamiento y rellene una nueva hoja padronal.

Para actualizar el nivel de estudios, el padrón se viene cruzando desde hace varios años con el registro de Títulos del Ministerio de Educación y Formación Profesional (MEyFP de aquí en adelante), corrigiendo la información de partida en muchos casos. Por tanto, el padrón ofrece una cobertura completa, pero precisa mayor refinamiento.

La estimación del nivel de educativo alcanzado se lleva a cabo mediante la actualización de la información educativa de cada persona. Las fuentes con las que se cuenta son las siguientes:

- **Padrón.** El título escolar o académico es una variable de inscripción padronal que deben recoger los ayuntamientos en su padrón municipal y, por tanto, figura en los formularios de empadronamiento. La incorporación en el padrón está relacionada con la formación del Censo Electoral. Así pues, en la gestión continua del padrón se recoge esta información de los ayuntamientos y se depura con la información que se recibe semestralmente por parte del MEyFP.
- **Títulos del MEyFP.** Se dispone de información de los ficheros de títulos no universitarios (LGE, LOGSE, LOE), universitarios y homologaciones de títulos extranjeros a títulos universitarios y no universitarios, que, de manera continua, el MEyFP envía al Instituto Nacional de Estadística.
- **Censo 2001.** Para los casi 34,5 millones de personas de 16 años o más que respondieron al cuestionario censal en 2001 con fecha de referencia 1 de noviembre de 2001, se tiene información sobre el nivel de estudios alcanzado hasta ese momento.
- **Censo 2011.** Para los casi 3,5 millones de personas de 16 años o más que respondieron al cuestionario censal en 2011 con fecha de referencia 1 de noviembre

de 2011, se tiene información sobre el nivel de estudios alcanzado hasta ese momento.

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

Todas estas fuentes son consideradas para la estimación del nivel educativo. Para las personas de 16 años y más que residían en España a 1 de enero de la referencia censal anterior, se actualiza su nivel educativo.

La cobertura del nivel educativo alcanzado es casi total. Se imputa el nivel educativo a alrededor del 1,5% de las personas de 15 años y más, si bien en la gran mayoría de ellas consta información en uno o varios registros educativos, faltando solo concretar la clasificación detallada deseada (por ejemplo si se trata de un grado universitario de hasta 240 créditos ECTS o más).

En el caso de la información de **estudios en curso**, se utilizan principalmente las dos últimas fuentes referidas a matrículas. También se explota el Registro Central de Extranjeros para detectar nuevos estudiantes a partir de su autorización de residencia por estudios, así como una carga de los Títulos del MEyFP a fecha posterior a la referencia censal para deducir matrículas.

Finalmente, el **lugar de estudios** registrado se obtiene de las mismas fuentes que los estudios en curso, fundamentalmente de los ficheros de matrículas universitarias y matrículas no universitarias del MEyFP. También, el Registro Central de Extranjeros contiene información para conocer la provincia de estudios de la mayoría de los estudiantes así detectados. Para cerca del 92% de los estudiantes ha sido posible extraer el municipio de estudios y para un 0,4% adicional el lugar de estudios se ofrece a nivel de provincia. Esto deja un pequeño residuo cercano al 8% de lugar de estudios no asignado.

Observaciones de la variable de estudios en curso

En la publicación de los censos anuales se ha decidido **depurar los estudios en curso** de un pequeño porcentaje de estudiantes (aproximadamente un 1%), con matrículas demasiado bajas para su nivel educativo alcanzado. Se han permitido matrículas bajas en combinación con estudios alcanzados en casos controlados: por ejemplo tras confirmar que se trata de matrículas en enseñanzas elementales (de música y danza, o enseñanzas de adultos), matrículas en formación profesional básica de alumnos que han accedido a ella a pesar de no haber titulado la primera etapa de educación secundaria, o ciertas matrículas en bachillerato o enseñanzas de adultos.

También se han depurado cerca de mil matrículas que eran demasiado altas para la edad del estudiante.

A estas personas se les ha asignado una categoría específica (“Cursa estudios pero no hay información sobre los mismos”).

Adicionalmente, conviene observar lo siguiente sobre **los estudios en curso de la población más joven**: su asignación se ha resuelto en un porcentaje pequeño de casos mediante imputación y no puede considerarse aún del todo resuelta.

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

Por otro lado es preciso enlazar la información que proviene del sistema educativo con el fichero censal de población y el enlace tampoco es perfecto. Sobre todo para personas extranjeras, no siempre se dispone del identificador (DNI, NIE) que permite mejorar el enlace, sino que las personas figuran en los ficheros de matrículas con su pasaporte. Así, en los ficheros de matriculados en enseñanzas no universitarias se tiene cerca de un 2% de matrículas no enlazadas. Y, por el contrario, en los ficheros del sistema universitario no figuran los nombres y apellidos, que permitirían alcanzar un porcentaje de enlace ligeramente superior al actual, lo que supone que algo más de un 4% de las matrículas universitarias no se consiguen enlazar.

Así mismo, hay que considerar la posibilidad de personas que estén matriculadas en universidades españolas pero residan en el extranjero y cursen formación a distancia. En total, queda un pequeño residuo de entorno al 3% de los ficheros de estudios en curso que no se logran enlazar con ninguna persona del fichero de población, lo cual en parte puede deberse a títulos obtenidos o estudios cursados por no residentes pero en otros casos es el problema típico de falta de enlace por no disponer de buenos identificadores (ausencia de DNI o de nombres y apellidos para muchos registros).

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

Observaciones de la variable “Lugar de estudios”:

Hay que tener en cuenta que el uso principal que se da a las variables “Lugar de trabajo” y “Lugar de estudios” es determinar matrices de movilidad origen-destino. La variable “origen”, que sería el municipio de residencia según el Padrón, en un reducido número de casos puede no reflejar la realidad y originar flujos origen-destino aparentemente extraños. Pero también hay que considerar que el contexto de teletrabajo y formación a distancia, acrecentado durante 2020 por la pandemia, puede aumentar esas situaciones en principio extrañas. Intentar delimitar errores y proceder a correcciones podría resultar aventurado y se ha preferido no introducir correcciones.

4.4 VARIABLES LABORALES: RELACIÓN CON LA ACTIVIDAD ECONÓMICA

Posteriormente a la primera publicación del censo de población, se ofrece la relación con la actividad económica para la población de 16 años y más, basada en la semana anterior a la fecha de referencia del censo.

Esta información procede de numerosas fuentes administrativas con información personal de temática laboral, que se combinan en el censo para obtener la variable estadística.

Fuentes de las variables laborales:

Se han tenido en cuenta las siguientes fuentes:

- Fuentes de tipo laboral:
 - **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 (INSS)** que cobraron una pensión y causa de la misma (jubilación, invalidez, orfandad, viudedad).
- Información educativa sobre estudiantes matriculados en el sistema educativo, tanto para enseñanzas universitarias como no universitarias (ver apartado anterior).
- Información tributaria del modelo 190 de la Agencia Tributaria y modelos similares de las Haciendas Forales.
- 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.

Para cada persona con 16 años y más, se busca información en las fuentes laborales y educativas, en primer lugar. Se busca información adicional de naturaleza tributaria para las personas que no se encuentran en ellas, o para aquellas de las que solo se puede deducir una situación de inactividad distinta a jubilación o incapacidad (por ejemplo perceptores de prestaciones de orfandad o viudedad).

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.

A pesar de ello, y al contrario que en el apartado anterior sobre variables educativas, lo más común es que una persona solo aparezca en una fuente administrativa de temática laboral. O que, si aparece en varias, la información sea la misma. En general, cada fuente administrativa tiene una cobertura casi completa de una situación laboral concreta:

- La mayoría (94,2%) de personas ocupadas en el censo son captadas en el fichero de afiliados de la Tesorería General de la Seguridad Social. El porcentaje aumenta si se tienen en cuenta las mutualidades (98,7%).
- En el caso de los parados, son captados en el fichero de demandantes de empleo del SEPE (96,3%).
- Las personas receptoras de prestaciones de jubilación o incapacidad provienen en su mayoría del Instituto Nacional de la Seguridad Social (90,5% y 95,5% respectivamente). El porcentaje aumenta si se tienen en cuenta las mutualidades (97,3% y 95,9% respectivamente).
- Los estudiantes son en su mayoría (93,7%) matriculados en enseñanzas universitarias y no universitarias.

Observaciones sobre la fecha de referencia:

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

La «situación de actividad actual» es la situación presente de una persona respecto a la actividad económica, basada en un período de referencia de una semana, que puede ser una semana indicada, reciente, fija, de calendario, la semana pasada completa de calendario o los siete días previos al censo.

Dado que la fecha de referencia del censo es el 1 de enero de cada año, se trataría entonces de obtener la situación laboral de la población durante los días previos. Por ello se ha hecho una solicitud específica a la Tesorería General de la Seguridad Social intentando reflejar la situación de afiliación durante la última semana completa previa a la fecha de referencia, y se ha construido la información tomando como ocupados a todas las personas con actividad durante esa semana.

Reglas de decisión de la relación con la actividad:

En caso de que la información de los registros laborales y educativos de una persona no coincida, se aplican unas reglas de decisión determinísticas para clasificarla en una única situación. Esto afecta a cerca de tres millones de personas.

Se han establecido una serie de reglas de prioridad basadas en las recomendaciones de Naciones Unidas y el Reglamento Europeo para los Censos:

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

También se aplican estas reglas en función de la edad:

- Si la persona tiene más de 65 años y no se la encuentra en ninguna fuente administrativa, se la considera inactiva.
- Si la persona tiene 15 años exactos, y no es encontrada en ninguna otra fuente administrativa, se la considera estudiante.
- En el caso de personas de menos de 65 años que no se han encontrado en ninguna fuente administrativa, se les imputa en la categoría de otras situaciones de inactividad. El porcentaje de no asignados se sitúa en aproximadamente el 9,5% de la población de 16 años y más. 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.

No se hacen correcciones, ni por ejemplo depuraciones en función de la edad, sobre la variable resultante de aplicar estas reglas.

5 Enlaces de registros administrativos con padrón

El fundamento de un censo basado en registros administrativos es el enlace o cruce de la información registrada de las personas con los residentes. Por esta razón las peticiones de datos de registros administrativos traen la información personal que identifica a cada persona: el tipo de NIF, en general el DNI para españoles y el NIE para extranjeros, nombre y apellidos (como variables separadas), fecha completa de nacimiento, sexo, e incluso, si están disponibles, la nacionalidad, el lugar de residencia o cualquier variable que, si bien por sí misma no identifica a la persona, sí sirve como parámetro de validación del enlace con padrón.

Hay varios aspectos destacables sobre el enlace de personas con el padrón:

- Hacen falta varias estrategias de enlaces para considerar que una fuente administrativa ha sido enlazada satisfactoriamente con padrón.
- La mayoría de fuentes empleadas en el censo llevan un mínimo de media docena de cruces. En general se han ideado tantas estrategias por fuente como ha sido necesario para alcanzar un porcentaje de enlace cercano al 100%.
- Cada estrategia empleada consta de dos fases:
 - Cruce o enlace de una o varias variables de identificación de la persona (entre la fuente y el padrón).
 - Validación del cruce o comprobación de que el cruce realizado no enlaza por exceso.

Adicionalmente a estas fases, es común que haya una fase de eliminación de duplicados. Por ejemplo, en un cruce por nombre y apellidos, si hay varias personas del padrón que se llaman igual, el enlace ha de hacerse solo con una de ellas. La decisión de con cuál se enlaza se basaría en el resto de variables de identificación de la persona, en este ejemplo podría hacerse en base a la fecha de nacimiento.

- En particular, solo con el DNI / NIE no se satisfaría un porcentaje de enlace cercano al 100%; así como tampoco se podrían validar los cruces, que es otro aspecto crucial del proceso.

En extranjeros, la calidad de los datos personales más allá de su NIE, pasaporte, o documento de identidad, son un buen indicador de la calidad a la hora de enlazar con padrón. Esto se debe a que es posible que distintas fuentes administrativas se hayan actualizado en distintos momentos temporales.

Un ejemplo más genérico de las dificultades que acarrea recibir una fuente con pocas variables identificativas se ha descrito en el apartado 4.3 - *Observaciones de la variable de estudios en curso* (enlace de matrículas universitarias).

Porcentajes de enlace:

A mayor calidad de los identificadores, así como del resto de variables de identificación de la persona, mayor porcentaje de enlace entre registros administrativos y padrón.

Se ha partido de la base de que el porcentaje ha de ser cercano al 100% para poder considerar satisfactorio el enlace de un registro con padrón, pero esta afirmación está llena de matices.

En realidad depende de la naturaleza de la fuente. Por ejemplo, una fuente como el histórico de contratos del SEPE, por contener información de varios años, va a tener un porcentaje de enlace ideal menor. Otro ejemplo es el fichero de afiliados a la Seguridad Social, que a pesar de ser un registro con una calidad de dato excelente, nunca se podrá alcanzar el 100% de afiliados enlazados con padrón. Aunque el proceso fuera perfecto, los afiliados en alta no residentes en España siempre quedarán (correctamente) por enlazar.

Si se excluye el histórico de contratos, el porcentaje de enlace de las fuentes de temática laboral es del 98,5%. Dentro de ellas, las mutualidades de funcionarios rozan el máximo con un 99,6% de enlazados; las fuentes procedentes de la Seguridad Social, que son mucho más numerosas y que contienen una casuística más rica de situaciones laborales y de nacionalidades, enlazan en un 98,9% (Tesorería) y 97,8% (INSS) gracias a tener una gran calidad del dato de identificación.

En cuanto a los ficheros de matrículas universitarias y no universitarias, como se avanzó en el apartado 4.3 - *Observaciones de la variable de estudios en curso*, el porcentaje de enlace es del 97,0%, si bien el 1,7% de ellas son **enlaces probabilísticos**. Se entiende por probabilístico un método de enlace en el que una persona matriculada no es encontrada en el registro de población, pero se le asigna como proxy de otro residente con el que comparte características demográficas (como por ejemplo la edad, sexo, nacionalidad o el lugar de residencia). A efectos estadísticos, dicha matrícula cuenta como enlazada a todos los efectos sobre la estimación censal, si bien es en puridad una estimación que hacemos sobre la persona empadronada, a la que le hacemos corresponder una información de matrícula a través de un donante.

Preparación de los datos y estrategias de enlace:

Con anterioridad al enlace con padrón, se estandariza la información personal que se recibe de los registros. Durante esta fase se adecúan identificadores, fechas de nacimiento y, sobre todo, se normalizan los literales de nombres y apellidos. El hecho de tener dos apellidos o nombres compuestos supone una ventaja a la hora de realizar los enlaces. Esto conlleva, a cambio, un tiempo elevado de preparación y depuración de estas variables de identificación.

Gracias a esta fase, se han encontrado patrones que se aprovechan durante la fase de cruces. Por ejemplo, es común el mismo tipo de errores en la composición de fechas (como intercambiar el mes y el día). También es habitual encontrar desordenados nombre y apellidos. En las validaciones de cada estrategia de cruce se cuenta con herramientas como el cálculo de distancias alfanuméricas o de distancias en fechas de nacimiento que tienen en cuenta este tipo de casuísticas comunes en muchos registros. También se pueden asignar puntuaciones con respecto a lo común o poco común de un apellido, o incluso deducir con cierto margen de error si una persona es o no extranjera en función de su nombre y apellidos.

No obstante, no todas las fuentes proporcionan la información de la misma manera. Por ejemplo, en algunos registros tributarios, se proporcionan nombres y apellidos juntos en una única variable. Esto supone una dificultad añadida en las estrategias de cruce de dichas fuentes, intentando paliarse específicamente en ellas.

Una vez depurada la información registrada recibida, las estrategias de cruce más comunes son:

1. Enlaces por identificador, con y sin dígito de control, con la letra de extranjero; validación por distancias en la fecha de nacimiento y parecidos en nombre y apellidos.
2. Enlaces por igualdad de nombre, apellidos (ordenados o no) y fecha de nacimiento; validación por puntuación de apellidos.
3. Enlaces por igualdad de nombre y apellidos (ordenados o no); validación por distancias en la fecha de nacimiento.
4. Enlaces por identificador (parcial), sin dígitos de control o letra de extranjero, y fecha de nacimiento; validación por parecidos en nombre y apellidos.
5. También se enlaza por identificador (parcial) y se valida por distancias en la fecha de nacimiento y parecidos en nombre y apellidos.
6. Enlaces por parecidos de nombre, apellidos y fecha de nacimiento; validación por igualdad de nombre y apellidos, o por puntuación de apellidos.

Se ha simplificado el uso de variables como el sexo o la nacionalidad para resumir las estrategias más utilizadas, ya que se utiliza más comúnmente la fecha de nacimiento o el nombre y apellidos tanto para enlaces como para validaciones. Por ejemplo, si hay registros sin información de fecha de nacimiento o nombre y apellidos incompletos.

6 Difusión de los resultados

Los primeros resultados de los Censos de población se publican en diciembre de cada año, para la fecha de referencia a 1 de enero de ese mismo año. Ofrecen las cifras oficiales de la población residente en España con un gran detalle territorial (hasta sección censal), junto con sus principales características demográficas.

En la primera edición de los Censos de población, el pasado diciembre de 2023, se han publicado conjuntamente los primeros resultados de los censos a 1 de enero de 2022 y de 2023.

En abril de 2024 se ha completado la información demográfica de la población a 1 de enero de 2022 con variables de temática educativa y laboral, entre otras, procedentes de la explotación de ficheros administrativos temáticos proporcionados por distintos organismos. En septiembre de 2024 se añadirá más información de variables migratorias y temática laboral.

Se publica un amplio conjunto de tablas para los diferentes ámbitos geográficos, llegando hasta sección censal, que contendrán los recuentos de población clasificada por las principales variables demográficas, educativas y laborales.

Se acompañarán de mapas temáticos, principalmente a nivel municipal y de secciones censales, para los principales indicadores.