

INSTITUTO NACIONAL DE ESTADISTICA



# Household Projections

## Methodology

June 2024

# Index

1	Introduction	3
1.1	population and housing censuses	3
1.2	The estimation and projection of households	4
2	Objectives	5
3	Information Sources	5
4	Concepts and definitions	6
5	Methodology	6
6	Calendar and Dissemination Plan	11

## 1 Introduction

The demand for data on the evolution of the number, type and composition of households has been growing gradually in recent years. Traditionally, the population and housing censuses - conducted every ten years - are the main sources of this type of information. The last census, in 2011 revealed not only the extraordinary population increase as a consequence of the massive arrival of foreigners over the previous decade, but fundamentally, that experienced in the number of households and in their typology. Throughout the decade leading up to the last census in 2021, the population and the number of households have continued to grow, even if at a much slower rate than in the decade from 2001 to 2011.

### 1.1 POPULATION AND HOUSING CENSUSES

#### Comparison of dwelling, households and populations en the censuses

Census <sup>(1)</sup>	Population	Dwellings <sup>(2)</sup>	Households <sup>(3)</sup>	Growth since previous census					
				Absolute			Relative		
				Population	Dwellings	Households	Population	Dwellings	Households
1,991	38,872,268	17,220,399	11,852,075						
2,001	40,847,371	20,946,554	14,187,169	1,975,103	3,726,155	2,335,094	5.1%	21.6%	19.7%
2,011	46,815,916	25,208,623	18,083,692	5,968,545	4,262,069	3,896,523	14.6%	20.3%	27.5%
2,021	47,400,798	26,626,315	18,539,223	584,882	1,417,692	455,531	1.2%	5.6%	2.5%

(1) The 1991 census refers to 1 March, the 2001 and 2011 censuses to 1 November and the 2021 census to 1 January.

(2) "Dwellings" refers to total family dwellings, i.e. collective establishments are excluded.

(3) For census purposes, the terms "main dwellings" and "households" are synonymous, since a household is defined as a group of individuals who usually occupy the same dwelling.

Most of the increase in population, number of dwellings and number of households occurred between the 2001 and 2011 censuses.

The 2011 census showed an increase of 14.6% in Spain's population and 27.5% increase in the number of households since the previous census in 2001. This increase was also accompanied by an increase in the number of dwellings (20.3%).

However, it should be noted that the most significant increase was among households, whose number increased by almost 4 million (27.5%), reaching 18,083,692 million. It was significantly higher than had been estimated from surveys (e.g. the Labour Force Survey estimated a total of 17.4 million households in the fourth quarter of 2011).

Spain's population grew by almost 6 million people between the two censuses, with the arrival of foreigners contributing to this extraordinary population increase. In those ten years, the foreign population increased by more than three and a half million, reaching the figure of 5.3 million; this represented 11.2% of the population.

The housing stock also increased between the two censuses. In the 2011 census, the total number of homes in Spain exceeded 25.2 million, having grown by 20.3%.

The most pronounced increase occurred among primary dwellings, whose number increased by nearly 4 million (27.5%), to reach 18.1 million.

Both the population, the number of dwellings and the number of households continue to grow over the following decade, although in a more moderate way. According to the results of the 2021 population and housing census, both population increased by 1.2%, the number of dwellings by 5.6% and the number of households by 2.5%, compared to the previous census.

It can also be seen that while the population has grown over the last three decades, the number of dwellings and households has always grown more dramatically. This larger increase in households than in population has been taking place since the previous decades: thus, between the 1970 and 1980 censuses, households grew by 22.14% while the population grew by 10.88%, and between 1980 and 1991, the increases were 13.37% and 2.98%, respectively.

In addition, there were significant changes in both household composition and size. The average number of household members has decreased considerably since 1970, from almost four people per household to 2.5 in 2021.

In addition, the population of Spain grew by almost 6 million people between 2001 and 2011, with the influx of foreigners contributing to this extraordinary increase in population. During those ten years, the foreign population increased by more than three and a half million, reaching 5.3 million, which represented 11.2% of the population. This was also a reason for the change in the population composition and thus in household dynamics.

#### Population by nationality in the censuses

Census (*)	Population			Growth since previous census						
	Total	Spanish	Foreign	Absolute			Relative			
				Total	Spanish	Foreign	Total	Spanish	Foreign	
1,991	38,872,268	38,518,901	353,367							
2,001	40,847,371	39,275,358	1,572,013	1,975,103	756,457	1,218,646	5.1%	2.0%	344.9%	
2,011	46,815,916	41,563,443	5,252,473	5,968,545	2,288,085	3,680,460	14.6%	5.8%	234.1%	
2,021	47,400,798	41,998,096	5,402,702	584,882	434,653	150,229	1.2%	1.0%	2.9%	

## 1.2 THE ESTIMATION AND PROJECTION OF HOUSEHOLDS

The situation regarding the number and dynamics of households is quite variable, so the statistical system must try to capture, as far as possible, the current situation at any given time, and do so instead of having to wait until ten years have passed and a new census has been carried out.

In the absence of other data, household surveys were frequently used to discover recent developments in the number of households. One of the problems frequently associated with household surveys, however, is the bias that can be introduced into results owing to a lack of response from certain types of households that are more difficult to locate (not at home) or who are unwilling to collaborate. This means that the description of the society and of the forms of coexistence provided by the survey, and even the number of households estimated, may be distorted.

Once the results of the 2011 Population and Housing Census were obtained, which provided all of the precise information on the number and composition of the households that made up the population, the challenge of producing updated data on the number of households was undertaken, which represented a significant advance in the quality of the information produced by the INE. These statistics were processed internally for use

in the calibration of household surveys, and resulted in the launch in 2023 of the Continuous Population Statistics, which publishes households by province and size on a quarterly basis.

Furthermore, as with the population, the recent evolution in the number of households is not the only item of interest. It is also necessary to show, in advance, household creation and dissolution dynamics. That is why the INE proposed the **Household Projection** operation.

This operation was created with the purpose of gradually incorporating as many additional sources of information and methodological improvements into the methodology as deemed necessary to better identify and measure household evolution and composition. These methodological improvements will be incorporated in forthcoming editions, and will be documented in order to inform researchers and users.

---

## 2 Objectives

The Household Projection statistical operation was designed with the objective of supplying every two years a statistical simulation for the next 15 years of the number and size of households, according to the demographic trends and social behaviours currently monitored. The information is available on the households of persons resident in Spain, its Autonomous Communities and its provinces. Furthermore, the results of such operation are, due to the way in which they are constructed, fully consistent with those of the Population Projections also disseminated by the INE.

Since 2014, household projections have been carried out.

---

## 3 Information Sources

The following basic sources of information are used to prepare the Household Projection:

### **Population and Housing Census**

The latest population and housing census, in this case the 2021 census, provides the number of households on the census date, classified by household size and by province, as well as the population residing in family dwellings classified by the size of the household to which this population belongs, age group, sex and province of residence.

### **Population Continuous Statistics (PCS)**

Starting in 2023, after the publication of the Population and Dwellings Census 2021, the Continuous Population Statistics are produced. Includes the entire population series since 1971, as well as an estimation of the resident population in family dwellings and the number of households from 2021. These data correspond to the census data as of 1 January 2021. This produces provisional estimates from the latest available census. Starting in 2021, the Population and Dwellings Census will be separated into Population Census, which will be annual, and Dwellings Census, with less successive frequency.

## Population projections

The Population Projections operation provides the projected population for the next 15 years at the provincial level, a necessary input for the household projection.

---

## 4 Concepts and definitions

**Resident Population:** the resident population in a given geographical area is the persons that, on the reference date, have their regular residence in said area.

**Habitual residence**<sup>1</sup> is the registered place of residence.

**Family Dwelling:** dwelling intended to be inhabited by one or more people who do not constitute a collective, regardless of the ties between them.

**Collective Dwelling:** dwelling intended to be inhabited by a collective, that is, by a group of people subject to a common authority or regime not based on family ties or coexistence. A collective dwelling may occupy only part of a building or - more frequently - the whole of it.

**Household (according to residence criteria):** a human group formed by one or several persons who habitually reside in a family dwelling all or most of the year; family ties do not necessarily exist among the members of the group.

**Household Size:** number of household members.

---

## 5 Methodology

The model chosen to carry out the household projections is based on the so-called propensity method. This method is based on the following article: "Bell, M., Cooper, J., et al (1995) Household and Family Forecasting Models. A review. Canberra. Department of Housing and Regional Development".

The method consists of calculating the propensity (probability) that, based on basic sex and age characteristics, persons of a certain population group will belong to a certain type of household. This propensity is measured by the quotient between the total number of persons belonging to the population group under consideration who reside in households of that type, and the total number of persons in the population group in question. That is, using the formula:

$$PH_{h,s,x}(k) = \frac{P_{h,s,x}(k)}{P_{h,s,x}}$$

Being:

- $PH_{h,s,x}(k)$  the propensity that persons of sex s and age x in province h have of belonging to a type k household.

---

<sup>1</sup> Definition of habitual residence established in *Regulation 763/2008 of the European Parliament and of the Council on Population and Housing Censuses and in Regulation 1260/2013 of the European Parliament and of the Council on European demographic statistics.*

- $P_{h,s,x}(k)$  the population of persons of sex  $s$  and age  $x$  in province  $h$  in type  $k$  households.
- $P_{h,s,x}$  the resident population in family dwellings of persons of sex  $s$  and age  $x$  in province  $h$ .

These propensities are applied to the figures for the population permanently residing in family dwellings; i.e., the population residing in collective dwellings is excluded. With this, an estimate of the number of households of each type consistent with the population in question is obtained.

The household type will be defined based on size, using the following types of households:

1. Household Size 1
2. Household Size 2
3. Household Size 3
4. Household size greater than or equal to 4

Ten population groups have been established for the combinations of sex (man, woman) and age group (under 20 years, 20 to 39, 40 to 59, 60 to 79, over 79 years), bearing in mind that household composition and dissolution dynamics are mainly related to the differences in these variables. At the start, nationality was also considered, but it was finally discarded, since the foreign population is smaller, and is influenced by the processes of acquisition of Spanish nationality.

The household projection consists of simulating the number of households there would be by size and province for the years following a certain start year, on January 1 of each year, so long as the trends of recent years continue.

The first edition of household projections was carried out in 2014 and covered a period of 15 years, from 2014 to 2029. Since then, a new edition is published every two years.

In order to carry out this projection - which will be calculated starting in 2024 and will have a projection horizon of 15 years, that is, until 2039, the calculation method is divided into several steps:

1. Projection of propensities and average sizes of households of four or more people throughout the projective period, by extrapolating the average annual evolution observed in the PCS in the last three years (between 2021 and 2024).
2. Application of projected propensities to projected populations in the Population Projections operation.
3. Obtainment of the number of each size of household in each province for each year of the projection, and application of the projected average sizes of households of four or more people to calculate households of this size.

Thus, the basis for determining the household projection is the population projection, but hypotheses must also be made regarding the evolution of propensities and average sizes.

To project the propensities, an extrapolation is carried out using limited exponential-type formulas, which place a lower cap on decreasing propensities and an upper cap on increasing propensities. The same criterion, with small variations, is used to extrapolate

the mean sizes of households of four or more people. These limits act as asymptotes, toward which average sizes and propensities tend over time.

The use of these formulas ensures that the average size of households of four or more persons could not be less than four, or trigger its growth, and prevents the decreasing propensities from becoming negative and the increasing propensities from reaching values that cause the sum of the propensities of each population group to be much higher than one (1); with one being the theoretical value of said sum.

The formulas used to **calculate the propensities** for each year are the following:

- a) For propensities which, according to the ECP household data, are decreasing between years t-3 and t on 1 January (that is, if  $PHP_{h,s,x}^{t-3} > PHP_{h,s,x}^t$ ), it is calculated:

$$PH_{h,s,x}^{t+n}(k) = 0,0001 + \left[ (PH_{h,s,x}^t(k) - 0,0001) \cdot \left( \frac{PHP_{h,s,x}^t(k)}{PHP_{h,s,x}^{t-3}(k)} \right)^{\frac{n}{3}} \right]$$

Where:

- t is the start year for the projection, which in this case will be t = 2024 and will refer to January 1, as in all the years involved in the projections.
- n is the number of years elapsed since the reference point from which the projection starts (with values n = 1, 2, 3, ..., 15).
- k is the type of household according to size, with values k = 1, 2, 3, 4, 5, with k = 4 referring to households with 4 or more people.
- $PH_{h,s,x}^{t+n}(k)$  is the propensity for a person of sex s and age group x in province h to belong to a household of size k on January 1 of year t+n.
- $PH_{h,s,x}^t(k)$  the propensity for a person of sex s and age group x in province h to belong to a household of size k in the starting year of the projection.
- 0.0001 is the lower cap set for decreasing propensities.
- $PHP_{h,s,x}^{t-3}(k)$ ,  $PHP_{h,s,x}^t(k)$  are the propensities that a person of sex s and age group x in province h belongs to a household of size k on January 1, years t-3 and t respectively.

This formula makes any propensity calculated worth more than 0.0001, which was an arbitrary value chosen based on the values of the minimum propensities calculated from the 2011 census data, and is still valid with the 2021 census data.

- b) For propensities which, according to the ECP household data, are **increasing** between years t-3 and t at 1 January (that is, if  $PHP_{h,s,x}^{t-3} < PHP_{h,s,x}^t$ ), it is calculated:

$$PH_{h,s,x}^{t+n}(k) = \text{LÍMITE} - \left[ (\text{LÍMITE} - PH_{h,s,x}^t(k)) \cdot \left( \frac{PHP_{h,s,x}^{t-3}(k)}{PHP_{h,s,x}^t(k)} \right)^{\frac{n}{3}} \right]$$

Defining all the elements of the formula as in the decreasing case, and considering in this case the LIMIT as the smallest of the following four values

- $PH_{h,s,x}^t(k) + C$ , to prevent the propensities from increasing too much in absolute value.



- $PH_{h,s,x}^t(k) \cdot M$ , to prevent the propensities from increasing too much in relative value. For example, the propensities of belonging to size one households of persons under 20 years of age are very close to 0, but the variability can be very high, and transferring this evolution in the long term can provide values completely unrelated to those of the starting point.
- T, to avoid the propensities being in any case greater than a certain value.
- $PH_{h,s,x}^t(k) + \left[ \left( \frac{PH_{h,s,x}^t(k)}{PH_{h,s,x}^{t-3}(k)} \right)^{\frac{1}{3}} - 1 \right] \cdot S$ , to avoid the exponential formula that calculates the propensities producing very large increases in the long term.

The values set for these formulas are C=0.12, M=2, T=0.9 and S=10, and they have been determined based on a series of tests that vary the possible values and check their operation.

After propensities for the projection period are calculated, they must be adjusted so that the sum of the propensities of all sizes for each population group continues to add up to one - since, when projecting each size separately, it is common for them to stop adding up to one, although they do not deviate too much. This adjustment is made using a recursive procedure, which consists of taking the propensities within each sex and age group, and dividing each of them by the sum of all the propensities in that sex and age group. Should the growth of any these propensities change sign regarding the evolution of the Municipal Register, what remains to be assigned is left constant and is distributed among the propensities of other sizes. The process is repeated until an adjustment is reached that respects the growth sign in all sizes.

The following formulas are used to **calculate the average sizes** of households of four and more people for each year:

- For average household sizes of 4 and more persons which, according to the ECP household data, are **decreasing** between years t-3 and t as of 1 January (that is, if  $TMP_h^{t-3} > TMP_h^t$ ), the formulas used are:

$$TM_h^{t+n}(4) = LIMITE + \left[ (TM_h^t(4) - LIMITE) \cdot \left( \frac{TMP_h^t(4)}{TMP_h^{t-3}(4)} \right)^{\frac{12n}{3}} \right]$$

Where:

- $TM_h^{t+n}(4)$  is the average size of households with four or more members in province h in year t+n.
- $TM_h^t(4)$  is the average size of households with four or more members at the time of reference from which the projection begins, on January 1 of year t, for province h.
- $TMP_h^{t-3}(4)$ ,  $TMP_h^t(4)$  are the average sizes of households with four or more members in province h on January 1 for the years t-3 and t respectively.

The LIMIT is defined as the highest of the following three values:

- $TM_h^t(4) - CPH_{t-1} + C$ , to prevent that the average sizes decrease too much in absolute value.

- $TM_h^t(4) \cdot M PH_{t-1} \times K$ , to prevent that the average sizes decrease too much in relative value.
- T, to prevent the average sizes from in any case being less than a certain value.

The values set for these formulae are  $C = 0,40$ ,  $M = 0,90$  and  $T = 4,1$ , and have been determined on the basis of a series of tests on the basis of known sources and by varying the possible values and checking their performance.

With this formula, the mean size cannot be less than 4.1; this value is an arbitrary and modifiable figure, but it seems reasonable that it should be greater than four, since four would be the extreme case in which there was no household with more than four members, which seems unlikely.

In this formula - unlike the formula used for propensities - the exponent for the power is multiplied by 12. In both cases, an attempt is made to achieve a reasonable evolution of these variables, so that the theoretical series of sizes and propensities have values close to those obtained at the series origin, and evolve similarly to that seen in the Registers for the years considered. Since the magnitude of the two items is different (the propensities less than 1, and the mean sizes greater than 4), the speed of approach to the asymptote marked by the cap is different: if we do not multiply by a constant in the case of mean sizes, the speed is very slow, and the imposition of the cap causes the curve to descend excessively (in the case of the upper cap; analogous to the lower cap). The value 12 was chosen empirically.

- b) For the average household sizes of 4 and more persons which, according to the ECP household data, are **increasing** between years t-3 and t on 1 January (that is, if  $TMP_h^{t-3} < TMP_h^t$ ), the formulas used are:

$$TM_h^{t+n}(4) = LÍMITE + \left[ \left( LÍMITE - TM_h^t(4) \right) \cdot \left( \frac{TMP_h^{t-3}(4)}{TMP_h^t(4)} \right)^{\frac{12n}{3}} \right]$$

With all the formula elements defined as for the decreasing case, and in this case considering the LIMIT as the smallest of the following three values:

- $TM_h^t(4) + C$ , to prevent the average sizes from increasing too much in absolute value.
- $TM_h^t(4) \cdot M$ , to prevent the average sizes from increasing too much in relative value.
- T, to avoid the average sizes being in any case greater than a certain value.

The values set for these formulas are  $C = 0.40$ ,  $M = 1.10$  and  $T = 5.5$ , and they have been determined based on a series of tests based on known sources that vary the possible values and check their operation.

Finally, the projected number of households by size and province on January 1 of each year is obtained from the following formula, which shows all elements previously obtained: that is, propensities, the average sizes of households of four or more members, and the resident population figures in family dwellings, as follows:

$$H_h^{t+n}(k) = \frac{\sum_{s,x} PH_{h,s,x}^{t+n}(k) \cdot P_{h,s,x}^{t+n}}{TM_h^{t+n}(k)}$$

Being:

- $H_h^{t+n}(k)$  the number of households of size  $k$  in the province  $h$  on January 1 of year  $t+n$ , with  $t=2024$  and  $n=1.2...15$ .
- $PH_{h,s,x}^{t+n}(k)$  the propensity that a person of sex  $s$  and age group  $x$  belongs to a household of size  $k$  in province  $h$  on January 1 of year  $t+n$ .
- $P_{h,s,x}^{t+n}$  the population of sex  $s$  and age  $x$  in province  $h$  on January 1 of year  $t+n$  that resides in family dwellings.
- $TM_h^{t+n}(k)$  the average size of households of size  $k$  in province  $h$  on January 1 of year  $t+n$ .

---

## 6 Calendar and Dissemination Plan

The 15-year Household Projection is published every two years, starting in 2014. The last publication took place in June 2024 and contains, in accordance with the Population Projections, information at the provincial level for the years 2024 through 2039.

A single table is published, containing the evolution of the number of households for Spain, its autonomous communities and its provinces on 1 January of each year for the next 15 years, by household size (1, 2, 3, 3, 4, and more members). This table also includes the average household size.