



Economic and Social Council

Distr.: General
2 August 2023

Original: English

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Population and Housing Censuses

Twenty-fifth Meeting

Geneva, 20–22 September 2023

Item 2 of the provisional agenda

Lessons learned from censuses of the 2020 round

General overview of 2021 register-based census in Spain

Note by Instituto Nacional de Estadística (INE, National Institute of Statistics) Spain*¹

Summary

The 2021 Census was the first fully administrative register-based census in Spain. In this paper we address the conditions we believe were essential for the success of this methodological challenge: The *Padrón* as the population register 'backbone'; the existence of thematic administrative registers for each domain, such as education, civil status, economic activity, household, etc.; as well as legal access to those registers.

We exemplify the general approach we undertook for some of the more complex-to-build population variables.

* Prepared by Cristina Casaseca Polo and Julia Bermejo Sanz.

Note: The designations employed in this document do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

¹ This document was submitted late due to delayed submission of the document by the National Institute of Statistics, Spain.



I. Introduction

1. Traditionally, censuses have been the most important operation in the statistical offices of all countries. In the collective imagination of society as a whole, the census year was the period in which statistical institutes mobilized a significant amount of human and material resources in order to count and survey us all. The arduous census work was planned well in advance and with great precision, so that even today, for the most veteran, the feeling persists that the experience of a statistician is measured according to the number of censuses he or she has experienced during his or her professional career.

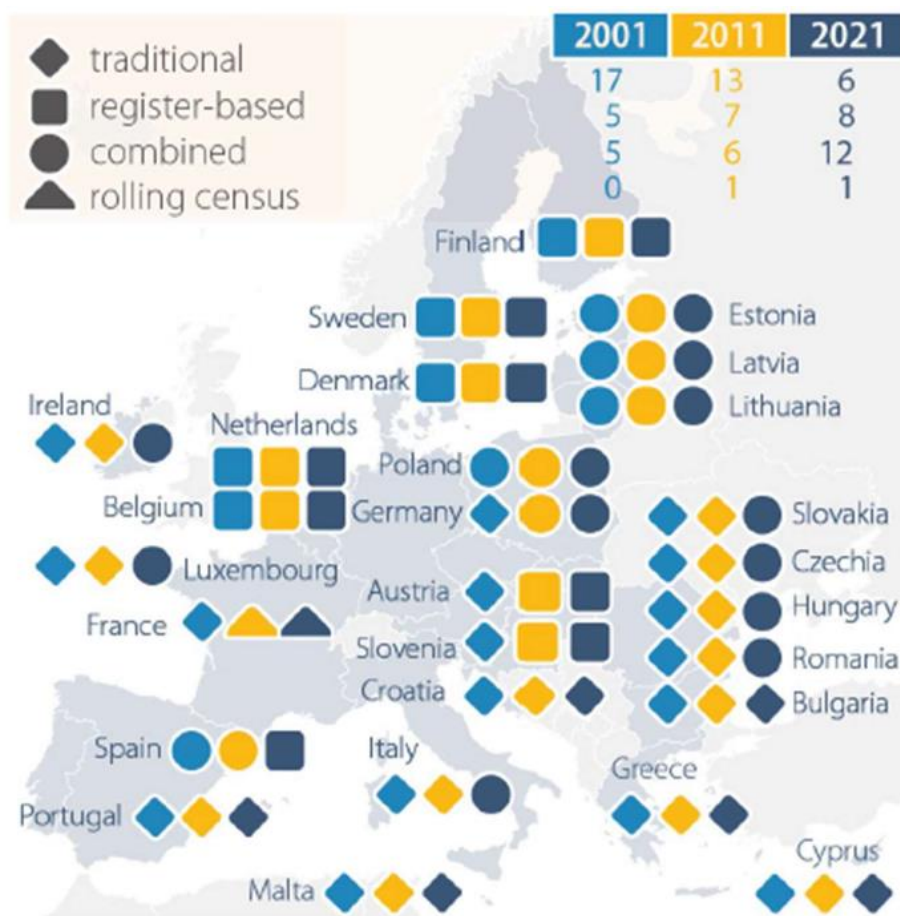
2. For the 2001 Census, the last traditional-exhaustive census carried out in Spain, around 45,000 people were hired for four months and collected information from more than 40 million people, at a cost of approximately 250 million euros.

3. Two decades have passed since then and the panorama of the current 2021 Population and Housing Census is completely different: the census work goes almost unnoticed by society and even within the INE, except for the approximately fifteen people directly involved in the project as well as the horizontal units that collaborate in it. The 2021 census is the first census in Spain based entirely on the exploitation of administrative records. This methodological change will allow the census process to be carried out every year and thus produce annual censuses, making the 2021 census the last decennial census in our country.

4. Henceforth we join the small group of countries that base their censuses entirely on administrative sources: only fifteen countries in the world currently construct their censuses with administrative data, while in the European Union there are eight, the largest of which is Spain.

Figure 1

Collection methods of the Population Censuses in the EU-27 (2001, 2011, 2021)



Source: European Parliamentary Research Service (EPRS) based on data from UNECE and the NSIs of EU Member States.

5. The COVID-19 pandemic has accelerated the plans that many developed countries already had to orient their population censuses towards a register-based methodology, as this method is unanimously considered to be the most efficient, of higher quality and coverage, and offering less risk compared to a mass collection of questionnaires from households.

6. In fact, the results of the survey carried out by the United Nations Statistics Division on the impact of COVID-19 on censuses revealed that the pandemic had significantly affected the global round of censuses. Many countries had to delay the operation and only fifteen of the 130 countries that responded, including Spain, stated that the pandemic did not alter the census plans because they were based on administrative records.

II. The road to the Spanish register-based census

7. The trajectory in these twenty years until reaching the current census goes through the realization in 2011 of a combined census, with a methodology that used, on the one hand, the administrative information contained in registers for 47 million people and, on the other hand, a survey of a sample of 9 per cent of households. Although it represented a step forward towards the use of administrative records, the reception of the new census methodology by users was not entirely positive, as it was interpreted, in some areas, as a ‘cheap’ census that showed a loss of territorial detail with respect to previous censuses.

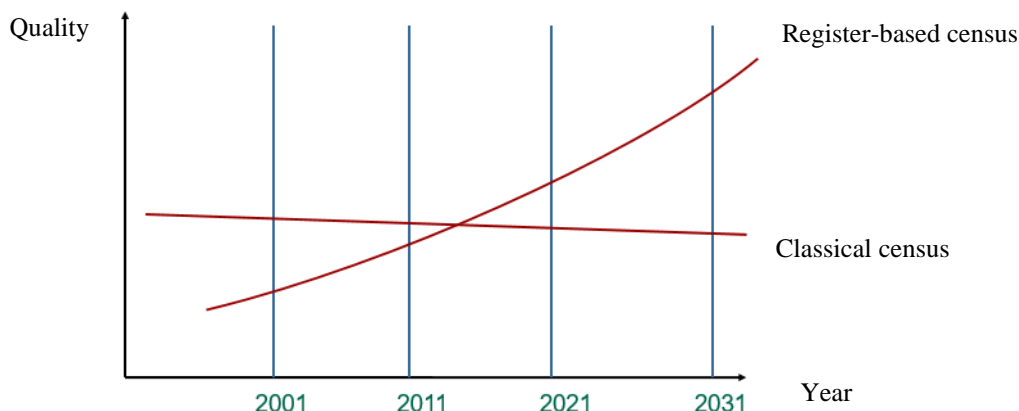
8. At the end of the 2011 census work, the information contained in the administrative files of various different agencies was analyzed, preparing each year the so-called ‘pre-census files’, as partial dress rehearsals of the future 2021 census. Finally, in 2018, it was determined that the quality of the available information would make it possible to dispense completely with census questionnaires and to construct an exhaustive census in 2021, without sampling, but based entirely on administrative sources. Part of the information from these pre-census files had already been incorporated for years as classification variables (educational level, occupation) of deaths, births and marriages.

9. In Spain, this leap from a classical or combined census to a census based on registers has been possible because a series of conditions have been met. Firstly, Law 12/1989, of May 9, 1989, on the Public Statistical Function, together with the European census regulations, allows INE to collect data for statistical purposes from the various different administrations. In addition, there is an extensive system of administrative records, with a high level of quality, whose content makes it possible to cover the many topics of the census variables (labour, education, migration, housing, etc. ...).

10. But undoubtedly the key element of the process is the existence of a population register, the *Padrón*, whose use as the ‘backbone’ or skeleton of the population census offers great advantages over other possible alternative population registers (tax registers, health cards or files of the Directorate General of the police.). All residents are obliged to register in the *Padrón* of the municipality where they live for most of the year, including people without legal residence or those registered in non-conventional dwellings. Furthermore, since the implementation more than 25 years ago of the *Padrón* as a continuous record, with INE as the coordinating body of more than 8,000 municipal registers, the *Padrón* has matured greatly as a population register, establishing, among other things, procedures for expiration and periodic verification of the residence of foreigners that make it an updated, reliable and accurate population register.

11. In making the decision for a register-based census, it was taken into account that there has been a progressive decline in the response rates of surveys (Keeter, 2018; Stedman et al., 2019), while register data do not suffer from non-response, bias or response errors, and while the quality of administrative data is increasing every day (Keeter, 2018; Stedman et al., 2019). One could represent the quality of censuses with a figure similar to this one:

Figure 2
Quality of censuses



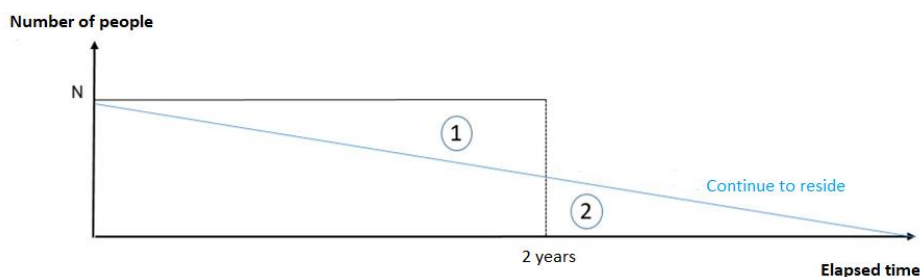
III. The census population figure

12. The method for counting the resident population used in the 2021 population census is completely novel for Spain; it is known in the literature as the ‘signs of life’ or ‘signs of presence’ method.

13. This method consists essentially of starting from a population register, in our case the *Padrón*, and evaluating the plausibility of the residence of each person based on their presence in other registers, so that for each individual it is decided whether he or she is counted as a resident or not. This is more precise than the administrative accounting of the population register, *Padrón*, on its own.

14. Nowadays most of us leave ‘traces’ in some register, but for the census the set of files that we can use is limited, since there are still barriers that prevent us from accessing certain records such as social networks, use of credit cards or medical consultations. Specifically, in the method of signs of presence applied in the 2021 census, we used employment and educational records and applied the method exclusively to the foreign population, with the aim of improving the census counts that result from the expiration processes or ex officio verification of residency.

Figure 3
Foreign population count scheme according to the date of expiration of their registration



15. If we analyze a cohort of foreigners entering Spain in a given year, and we assume that the registration of all of them expires after two years, we would have the situation illustrated schematically in Figure 3. According to this scheme, the N persons in the cohort would be counted as long as their enrolment had not expired, and all of them would automatically cease to be counted after the expiration of their enrolment. But the cohort of immigrants in fact leaves the country in a manner more similar to what is shown by the descending blue line.

16. Thus, there is a group of foreign population who continue to live in Spain but whose registration has expired ((2) in Figure 3). If we cross check this population with the labour or

educational files, we find foreigners who were initially ‘expired’ (excluded from the count) but who, because they show sufficient signs of presence, are ‘rescued’ and counted in the census population figure. If we were to rely exclusively on *Padrón* counts, without using the information from these records, we would be making a significant error in the estimation of the relationship with labour activity, for example.

17. Conversely, there are people who no longer reside in Spain but whose registration has not yet expired ((1) in Figure 3), so they are still counted as inhabitants. Non-presence in these records does not seem sufficient to stop counting a registered person. The decision made for the 2021 census was to include in the count only the foreigners who did show a sign of life the previous year. This way, if a foreign person appeared in the pre-census file on January 1, 2020 as employed or unemployed or a pensioner or a student or in any other situation denoting his or her presence, but on January 1, 2021 they no longer appear in any of these registers, we consider that person to have left the country, even if his or her registration has not yet expired.

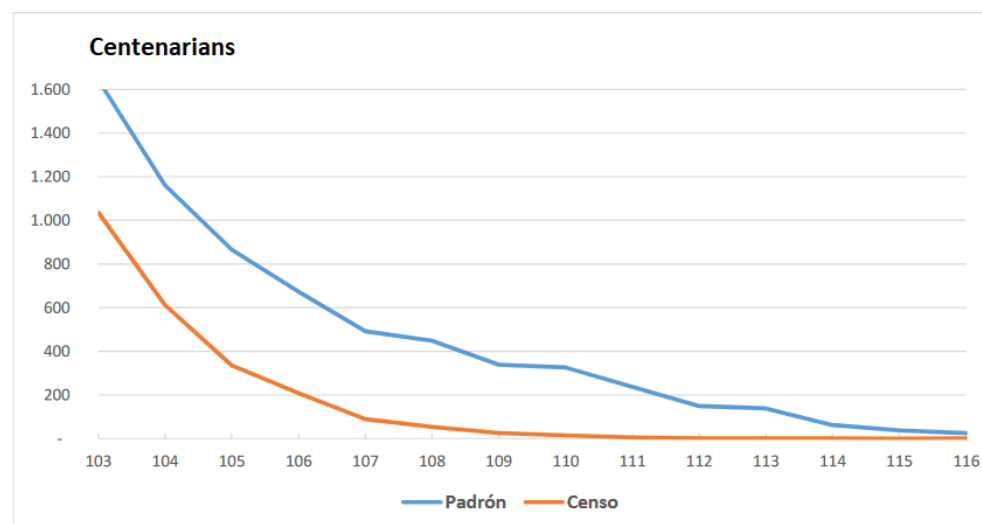
18. The application of signs of life in the foreign population solves more a problem of quality of variables than quantity of population, as shown in table 1.

Table 1
Application of signs of life in the 2021 census

	<i>Padrón</i>	<i>Census</i>	<i>Number</i>
(1) in Figure 3	Yes	No	152.599
(2) in Figure 3	No	Yes	158.044
Centenarians	Yes	No	5.228

19. Census population counts also differ from the *Padrón* for specific groups such as centenarians, or children aged between 0 and 1, for whom processes are carried out to correct the imperfections associated with the administrative management of the population register.

Figure 4
Centenarians



IV. Construction of census variables

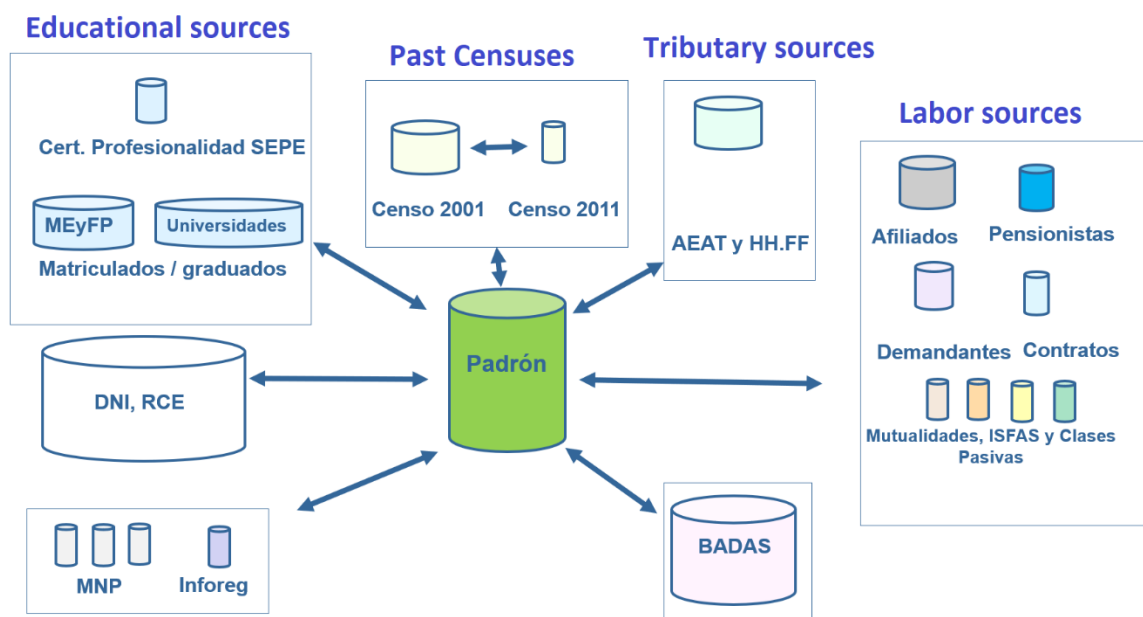
20. Once the census population has been obtained, the associated census variables are constructed for each person, so that the final product is similar to the data files that would be obtained if the person, dwelling, and household questionnaires of the entire population had been collected in a traditional fashion.

21. The identifying data of each person that appears in the *Padrón* will be used to link with the same person in the various different administrative sources and to capture and store all their information together.

22. Some variables such as basic demographics (sex, nationality, country of birth) or those related to migration (years of arrival and previous places of residence) are obtained by directly exploiting the data and variations contained in the *Padrón*. But for the rest of the variables (education, employment, marital status...) in general there is no single source containing all the information, making it necessary to apply specific algorithms that take into account all the information of the person in the different registers to obtain each one of the variables.

Figure 5

Padrón as skeleton and other sources of information



23. Regarding household information, the *Padrón* provides us with the address, so we have a first version of the census households as the group of people who share the same *Padrón* address. This information is contrasted with other sources, such as the cadastral reference (see section V) that each person declares as their usual residence, so that sometimes the census households are formed by joining or separating the initial registered households from the *Padrón*. To establish the relationships between the members of the household and determine the composition of the households, the information accumulated over the years on links between parents and children and couples contained in the files of births, marriages, taxes and residence permits for foreigners is used.

A. Level of education

24. The collection of information on educational attainment and education in progress has been one of the most labour-intensive elements of the 2021 census and has been achieved mainly thanks to the leadership of the Ministry of Education and Vocational Training (MEyFP) over the past few years. It has been necessary to strengthen collaboration with the education ministries of the Autonomous Regions to improve the quality of the administrative information.

25. In order to meet the objectives of the 2021 census and to maintain comparability with the results of previous censuses, information is provided on educational attainment and education in progress. In particular, for the level of studies, a slightly more detailed breakdown is offered than that required by Eurostat, with twelve categories based on the

classification of programmes, degrees and certifications in levels of training attained (Spanish CNED-A):

- (a) Illiterate;
- (b) Incomplete elementary/primary education;
- (c) Elementary/Primary education;
- (d) Lower secondary education and similar;
- (e) Second stage of secondary education with general orientation;
- (f) Second stage of secondary education with vocational orientation;
- (g) Post-secondary non-tertiary education;
- (h) Vocational training, fine arts and design and sports education at the higher level and equivalent; university degrees that require a bachelor's degree, with a duration of 2 years or more;
- (i) University degrees of 240 European Credit Transfer and Accumulation System (ECTS) credits, university diplomas, university-specific degrees of expert or specialist, and similar degrees;
- (j) University graduates with more than 240 ECTS credits, bachelor's degrees;
- (k) Master's degrees, specialities in health sciences by the residency system and similar;
- (l) University doctorate.

26. The Spanish CNED-A includes both current and non-current formal education programmes. It should be noted that we do not have full coverage of all categories of the CNED-A, since, for example, information on university-specific degrees is not available. In addition, we have partial coverage of post-secondary education and lower levels of professional certificates. Most of these more problematic categories fall under other categories of the classification.

27. The estimation of the level of education is carried out by incorporating the administrative information on the highest level of education attained, which is taken as the highest level of education found in the records. The sources available are the following:

- (a) *Padrón*: because of the formation of the Electoral Roll, the school or academic diploma is a registration variable that must be recorded by the municipalities in the *Padrón*. This information is collected in the population register as a continuous record and filtered with the information received every six months from the MEyFP;
- (b) MEyFP qualifications: information is available on the files of non-university degrees, university degrees, and ratifications of foreign degrees to university and non-university degrees, which the Ministry of Education continuously sends to INE;
- (c) 2001 census: for the nearly 34.5 million people aged 16 and over who responded to the census questionnaire in 2001 with the reference date of 1 November 2001, we have information on the level of education attained up to that time;
- (d) 2011 census: for the nearly 3.5 million people aged 16 and over who responded to the census questionnaire in 2011 with the reference date of 1 November 2011, we have information on the level of education attained up to that time;
- (e) Job seekers from the State Public Employment Service (SEPE): information on the level of studies completed is available for persons registered in the public employment services;
- (f) Certificates of professionalism (SEPE): there is a file of certificates of professionalism that accredit competences recognized by the Educational Administration, having the effect of validation of the corresponding professional modules;
- (g) University graduates: historical information is available on university graduates from 2010–2011 to the year prior to the census;

(h) Non-university graduates: historical information is available on graduates from the 2014–2015 academic year to the year prior to the census;

(i) Central Registry of Foreigners: information is available on foreigners with study authorization in Spain;

(j) SEPE contracts in force: information on the educational level of the worker is available since 2001 on the last contract of each person as of 1 January of the reference year;

(k) Enrolled in the educational system in non-university education of the MEyFP: information is available from the 2014–2015 academic year to the year prior to the census;

(l) Enrolled in university studies of the Council of Universities: information is available from the 2010–2011 academic year to the year prior to the census.

28. No administrative register exists in Spain that provides the educational level of the whole population with the required disaggregation. Moreover, the coverage of each data source is different. In general, the *Padrón* or MEyFP qualifications have nearly complete coverage, but are less detailed than university and non-university graduates files; and the *Padrón* or prior censuses may not have the most up-to-date information on each individual.

29. We could simplify by saying that the use of many sources is needed in order to update the information on level of education obtained from the *Padrón*. For example, if available, enrolment studies can be used to impute a higher level of education than that found in the *Padrón*, due to the entry requirements of many university courses, vocational training programmes, etc.

30. It is worth noting that less than 2 per cent of people have their educational level imputed. Moreover, the vast majority of those imputed are found in one or more than one source of educational information but at an insufficient level of disaggregation.

B. Economic activity

31. In order to meet the objectives of the 2021 census and to maintain some of the breakdown from previous censuses, the following seven categories of economic activity were proposed, i.e., those of the corresponding European Regulation plus one, which would be given for any breakdown (municipal or lower) whenever possible:

(a) Active population:

(i) Employed;

(ii) Unemployed;

(b) Inactive population:

(i) Retired, early retired;

(ii) Students;

(iii) Receiving a permanent disability pension;

(iv) Other inactive situations;

(v) Not reaching the national minimum age for working activity.

32. There is no single source that provides information on the entire population. The estimation is approached through the union of several different administrative records.

(a) Economic sources:

(i) Workers contributing to Social Security (General Treasury of Social Security) who were registered during the week prior to 1 January of each year;

(ii) SEPE job seekers;

(iii) Civil servants who do not contribute to the Social Security and belong to one of these three mutual insurance companies: MUFACE, MUGEJU and ISFAS;

- (iv) Persons in the Social Security system who received a pension, and the reason for the pension (retirement, disability, orphan's, widow's, widower's);
 - (v) Persons who receive a pension but are not members of the Social Security System ('passive classes');
 - (b) Educational information on students enrolled in the educational system, both for university and non-university education;
 - (c) Tax information (Tax Agency and similar forms of the Regional Tax Authorities).
33. For each person 15 years of age and older, information is sought from economic and educational sources, first of all.
34. When taking into account so many sources, it is normal that different situations coexist and lead to conflict for each person. For example, a person who has registered as a job seeker and has, at the same time, enrolled in a certain course of study.
35. Most commonly, each person appears in only one source or, if he or she appears in several, the information is the same. But in about three million cases there is a discrepancy, and a single situation must be assigned to each person. To do this, a series of priority rules are established based on the recommendations of the United Nations and the European Census Regulation. For example: If the person is found to be employed in any source, he or she is considered to be employed.
36. The percentage of unassigned people stands at 9.26 per cent of the population of 15 years of age and over for the 2021 census. The fact that a person is listed as unassigned should not be taken as unusual. In fact it is logical that many of the economically inactive do not appear in these records. The curve by sex and age of the unassigned is, moreover, very similar to the curve by sex and age of the economically inactive, which would indicate that a large part of these unassigned are economically inactive. For this reason, all of them are imputed to the category of 'other situations of inactivity'.

V. Dwellings

37. The 2021 housing census posed a major challenge compared to previous editions: it is the first time that it is based entirely on administrative records.
38. Just as the *Padrón*, the municipal register of inhabitants, constitutes the 'backbone' or skeleton of the population census, it was also necessary to find an administrative source to serve as a skeleton for the housing census. After many tests, the cadastre was chosen as the main source.
39. Using the cadastre as a source has many advantages. To begin with, the registers have an identifier, the 'Cadastral Reference', which is not always a unique identifier for the dwellings, but it is an excellent starting point for obtaining this unique identifier. However, the main advantage of the cadastre is that it contains geographic coordinates. This allows us to georeference the dwellings as well as all the cadastral information associated with them. Furthermore, by linking each inhabitant with their dwelling, the information associated with people can be georeferenced too, which constitutes a transversal element in the dissemination of census data.
40. In addition, the cadastre is a rich source of information on its own. The surface area of each dwelling, the year of construction, the type of dwelling, etc. can be obtained from the cadastral files.
41. Regarding the type of use, each dwelling will be classified as principal if it is the habitual residence of its components and as non-principal otherwise.
42. For non-principal dwellings, the classification between secondary (vacation home) and vacant (empty house) is, for the first time in Spain, established according to their electricity consumption in the year prior to the census reference date. In previous censuses

the distinction between secondary and vacant dwellings was determined by what the census agent perceived when visiting the building.

43. The source of information on electricity consumption is the ‘annual declaration of electricity consumption’ (tax form 159). It provides, for each electricity meter, its consumption throughout the year. The quality of the file does not allow for linking each meter with its corresponding dwelling. For this reason, aggregated information is provided, at the municipal level, on the number of dwellings according to consumption brackets. Ultimately, we provide information on the number of dwellings with no consumption or with consumption below certain determined thresholds, which can be assimilated to the concepts of empty dwellings or secondary dwellings.

VI. ECEPOV – The Survey of Essential Characteristics of the Population and Dwellings

44. It should be noted that there are a number of topics that traditionally appear in the census series but are not available in administrative sources. For this reason, the Survey of Essential Characteristics of the Population and Dwellings has been launched, whose questionnaire has been prepared with the collaboration of various organizations (CSIC-Higher Council for Scientific Research, Centre for Demographic Studies of Barcelona, some universities, Statistical Institutes of the Autonomous Communities, Ministry of Transport, Mobility and Urban Agenda,...),

45. The ECEPOV offers complementary information to the census on issues such as knowledge and use of languages, commuting, second homes, housing equipment, characteristics and facilities of buildings, etc.

46. This survey has a considerable sample size (172,444 dwellings in which 424,493 people reside), allowing us to offer estimates for municipalities with more than 50,000 inhabitants and provincial capitals.

VII. Conclusions

47. The 2021 census was the first fully administrative register-based census in Spain, which makes us the largest country in the European Union to have one. A strong legal base and a file serving as the skeleton of the population are two of the main criteria that Spain meets in order to be able to build a register-based census.

48. The signs of life procedure is applied to the *Padrón*, the municipal register of inhabitants, for the final accounting of the population.

49. Dozens of files from several institutions are needed to accumulate the information required to build each variable, as we exemplified with the level of education and the economic activity variables in this paper. In order to do that, linkage between the *Padrón* and each of those files is required.

50. In conclusion, it has been a great challenge to build a register-based Census in Spain, but it comes with great advantages as well: from now on, annual population censuses will be made, and each three or four years a housing census will be carried out. This, in addition to being a clear improvement in quality if compared to a traditional census, is the preferable course of action given the fact that there is a tendency to have more and more administrative records, and of increasing quality, in Spain.