

Population resident in group establishments

(Group Survey in the
2011 Population and
Housing Census)

Methodology

Subdirectorato General for
Sociodemographic Statistics

July 2013

Index

1 Introduction

2 Objectives

3 Research scope

4 Basic concepts and definitions

5 Target characteristics of the study

6 Questionnaire design

7 Sample design: type of sampling, size of the sample, selection of the sample, and sampling errors

8 Information collection: collection method, incidences and their processing, response rates

8.1 Information collection. Collection method

8.2. Incidences and their processing

9 Processing the information

10 Dissemination of the results

11 Population resident and population registered in group establishments, and comparability with the 2011 Census

Annex I: Questionnaire

Annex II: List of persons

1 Introduction

The Group Survey is part of the set of operations constituting the 2011 Population and Housing Censuses. The general methodological framework in which the project for Spain is developed is established by the recommendations from the Conference of European Statisticians for the 2010 census circuit, and more specifically by European Parliament and Council Regulation 763/2008, and the three regulations from the Commission that develop the previous¹(regarding definitions of variables and classifications, regarding data hypercubes and regarding the quality of the operation).

In addition to implementing the compulsory nature of carrying out the Census during the year 2011, this norm ensures the comparability of the results on a European Union level with regard to the methodology, definitions, associated statistical data and metadata and quality of the operation.

Naturally, beyond these international rules and recommendations, the project also needs specific information from Spain that is not already included in the aforementioned regulations. We have begun with an analysis of the variables studied on previous occasions and of the new proposals arising as a result of the social evolution over these years, to help in their determination.

A significant part of the analyses of the population included within the censuses is aimed at the population resident in group establishments.

The census targets those persons who live in both family dwellings and group establishments (hotels, retirement homes, barracks, convents, etc.) The means of facing the data collection for this group of persons must necessarily be different from that for the rest of the population (those residing in family dwellings), and the data is studied through a sample of dwellings. Thus, due to the specific characteristics of the group establishments, they are studied through an independent field operation. In order to avoid counting this population twice, as much as possible, the group establishments are excluded from the research corresponding to the sample of dwellings that is also part of the set of operations constituting the 2011 Population and Housing Censuses.

In the same way that the data on buildings and properties is obtained from the Building Census circuit, and the information corresponding to family dwellings is obtained from the sample of dwellings in the Population and Housing Census, the information corresponding to the groups is obtained from the Group Survey.

Group establishments include both group dwellings themselves (convents, barracks, retirement homes, long-stay hospitals, prisons, etc.) and hotels, guest

¹A) Regulation from the Commission regarding technical specifications of the variables and breakdowns thereof, passed on 30 November 2009 (published 15 December 2009).

B) Regulation from the Commission regarding statistical data and metadata that must be transmitted, passed on 16 June 2010 (published 17 June 2010).

C) Regulation from the Commission regarding the quality of the data and the format of its transmission, passed on 8 December 2010 (published 9 December 2010).

houses and similar establishments, so long as at least one person resides there regularly. If at the census reference date, 1 November 2011, there were no persons residing there regularly, it was considered to be, not a group establishment, but rather a locale. This specification was introduced due to international recommendations, and it thus appears in the census regulations.

2 Objectives

This survey enables determining the number of group establishments on the census reference date, and their type, as well as the number and basic demographic characteristics of the persons who reside therein. These variables and their breakdown level, less in the case of the population in family dwellings, are included in the regulations developed by the European Union, with the exception of the type of establishment that is included based solely on the needs of users in Spain.

3 Research scope

The target population of study of the Group Survey is the set of group establishments in Spain with at least one person resident on the census reference date 1 November 2011, as well as the persons resident therein.

On a geographical level, the research extends to the set of municipalities within the Spanish territory, including Ceuta and Melilla.

The reference date is 1 November 2011, carrying out the information collection during the period from January through March 2012.

4 Basic concepts and definitions

Dwelling: Structurally separate and independent venue that, given how it was constructed, reconstructed, transformed or adapted, is conceived to be inhabited by persons or, even though this is not the case, is used as a person's regular residence. As an exception, dwellings will not include venues that, despite being initially conceived for human habitation, at the time of the census are dedicated totally to other purposes (for example, those used exclusively as commercial premises).

Family dwelling: A dwelling designed to be inhabited by one person or more, not necessarily linked by family ties, and who do not constitute a group.

Group dwelling (or group establishment): a dwelling designed to be inhabited by a group, that is, by a group of persons subject to an authority or common regime, not based on family ties or on cohabitation, such as retirement homes, barracks, convents, institutions for persons with disabilities, prisons, etc. , so long as at least one person resides in the establishment regularly. If at the

census reference date, 1 November 2011, there were no persons residing there regularly, it was considered to be, not a group establishment, but rather a premises. The group dwelling may only partially occupy a building, or more frequently, the entirety of the building.

Premises: Venues that are structurally separated and independent (in the same sense used to define dwellings) that are not used exclusively as family dwellings and in which the economic activities of a company or institution may be carried out. The premises should be located in a building, occupying it totally or partially.

Regular resident: That person who has resided or plans to reside for a year in a group establishment is considered to be a resident, but questions are also asked of persons regarding this circumstance.

This considers the following **types of group establishment:** hotels or the like, worker residences, psychiatric hospitals or long-stay hospitals, retirement homes, institutions for persons with disabilities, other social care institutions for children and young persons, etc., religious institutions (monasteries, abbeys, etc.), military establishments (barracks, etc.), penitentiary institutions (prisons, reformatories, etc.) and other types of group. In particular, in the first type of group, unlike the others, it is common for there not to be persons residing there regularly, but rather simply spending a time there, so according to the census definitions, in this cases they are not considered groups, but rather premises.

Unlike the previous censuses, the following types are not considered to be group establishments in this census: general hospitals and short-stay speciality hospitals, where there are no persons residing regularly, since stays at these hospitals last less than 12 months, hostels for persons in a situation of social exclusion, on considering that those persons who go there have a residence or are homeless, nor university residences, student residences, boarding schools, military academies and schools or seminaries, as in all these cases, residences must be registered in the family dwelling.

5 Target characteristics of the study

The geographical identification variables that include the place where the group establishment is located are the municipality, province and Autonomous Community.

The variables collected relating to persons who regularly reside in group dwellings are the following:

- Sex
- Age
- Legal marital status

- Registration situation: registered in the group or not
- Educational level
- Relationship with economic activity (employed, unemployed or economically inactive)
- Professional status (businessperson, self-employed worker, worker working for others, other)
- Place of birth (Spain, another EU country, outside of the EU, other)
- Country of nationality (Spain, another EU country, outside of the EU, stateless)
- Place of regular residence one year ago (same municipality, another municipality in the same province, another province, another country)
- Size of the municipality of residence
- Number of living children born to women aged 16 years old and over

The following is collected regarding group establishments:

- Persons who reside therein: size of the group
- Type of group establishment

All these variables are included in the regulations developed by the European Union, with the exception of the last of each group (number of children for women and type of group) included, considering solely needs of users in Spain, as well as the registration situation of residents in groups included for the purpose of enabling the calculation of the census population figures, as is set out in the last section, **Resident population and population registered in group establishments, and comparability with the 2001 census.**

The names of the variables are the same as in the project for the 2011 Demographic Censuses, and they are broken down conceptually and territorially the same way, with the exception of some that require a lower breakdown level and are mentioned in parentheses in the above list.

The educational level is collected, according the following classification, which enables its conversion to the CINE, to one digit:

1. Cannot read or write
2. Can read and write but went to school for less than 5 years
3. Went to school for 5 years or more, but did not reach the last year of OSE, GBE or the first stage of secondary education
4. Reached the last year of OSE, GBE or the first stage of secondary education or has the School Certificate or Certificate of Primary Studies
5. Second-stage secondary education (LOE, LOGSE), Secondary Education, Post-Secondary Graduate, Post-secondary Non-higher Education, Pre-University Course
6. Intermediate Vocational Training, VTI, Industrial Technician or equivalent , Intermediate Music and Dance, Certificates from the Official Language Schools
7. Advanced Vocational Training, VTII, Industrial Master or equivalent
8. University Diploma, Technical Architecture, Technical Engineering or the equivalent

9. University Degree, Architecture, Engineering, Doctorate or the equivalent

The size of the municipality of residence refers to the size, as the number of inhabitants, of the municipality in which the group establishment is located. The following classification in eight categories is considered: municipalities with fewer than 2,001 inhabitants, 2,001 to 5,000, 5,001 to 10,000, 10,001 to 20,000, 20,001 to 50,000, 50,001 to 100,000, 100,001 to 500,000, more than 500,000 inhabitants.

The groups are collected according to the types of establishment set out below, though according to the information to be disseminated, different aggregations are used of the types of group indicated later.

1. Hotel, hostel or guesthouse
2. Residence for workers
3. Long-stay or psychiatric hospital
4. Asylum or retirement home
5. Institution for persons with disabilities
6. Hostels for those persons marginalized from society (excluded)
7. Another social care institution for children, young persons, etc.
8. Religious institution (monastery, abbey, etc.)
9. Military establishment (barracks, etc.)
10. Penitentiary institution (prison, reformatory, etc.)
11. Another type of group

The first grouping (7 types) is used for those tables on an Autonomous Community level, and for national results:

Health institutions G1: 3

Asylum or retirement home G2: 4

Institutions for persons with disabilities or for social care for children, young persons, etc. G3: 5, 7

Religious institutions G4: 8

Military institutions G5: 9

Penitentiary institutions G6: 10

Another type G7: 1, 2, 11

A greater grouping (5 types) is used for the tables on provincial and municipal levels:

Health institutions A1: G1

Asylum or retirement home A2: G2

Institutions for persons with disabilities or for social care for children, young persons, etc. A3: G3

Religious or military institutions A4: G4, G5

Another type A5: G6, G7

6 Questionnaire design

The questionnaire targeting each group is comprised of two parts. The first part includes the identification data of the establishment, and its objective is to ascertain the characteristics thereof: type of group and number of persons who reside there regularly, broken down by sex. Moreover, it includes instructions that enable completing it.

The second part includes in individual questionnaire, the information on the variables referring to each person selected in the centre sample. If the centre does not have more than 20 residents, it needs not select a sample of persons, but instead collects information on all the residents

The selection process of the sample of persons, when the establishment has more than 20 residents, requires the availability of an orderly listing of all those persons resident in the centre, and the use of a table for the selection of persons that appears on a sheet apart from the questionnaire, but which is submitted together with it.

This complementary sheet has a selection table on one of its sides, with the instructions to use it, and on the other side, it has a space for writing the given names and surnames of the "List of persons" comprising the sample of residents in the establishment, identified through a number assigned to the person, as established on said sheet. This is a sheet that complements the questionnaire, but it is not necessary to submit it with the latter once completed, as its use consists of allowing the individual questionnaires for each person to be identified through this number, thus not being necessary to include the given name and surnames.

The individual questionnaire for each person includes the following: the number of the person identified, the sex, the date of birth, the place of birth, the nationality, the marital status, if the person is registered in the establishment, the educational level attained, the relationship with economic activity, the professional status and the place of residence one year ago. Moreover, for women, if they have had children, and how many of them were born alive.

The individual questionnaires include the questions necessary for providing Eurostat with the information it requires, and only those questions, with the exception of the number of live children born to women, which Eurostat does not requires, but that has been considered of interest. They are identified only with the person number, not including the given name or the surnames.

The formulation of the questions is the same as that of the individual questionnaires that are included in the questionnaire that is submitted to the family dwellings in the sample of dwellings in the Population and Housing Census, even if there are no filters by age (women aged 16 years old and over for live-born children, persons aged 16 years old and over for the employment situation the last week, persons aged 16 years old and over for the educational level with the question of whether they attend a school centre in the case of begin under 16 years old).

In general, the information is provided by a person, with enough responsibility and knowledge of those resident in the group, in charge of collecting all the questionnaires and transmitting them to the INE, by any of the channels provided. Nonetheless, the instructions to the group state that, if any of the persons selected wish to submit the data from their individual questionnaire personally, they may use a photocopy of the corresponding individual questionnaire or download it from the INE website, thus preserving the privacy of their answers. For these cases, the centre questionnaire has a space to show, for each person selected, whether or not they are sending their questionnaire individually, including a contact telephone number if necessary.

7 Sample design: type of sampling, sample size, sample selection, estimators and sampling errors

The estimation of the population in groups has been carried out using the Group Survey, which is an operation that is independent from the rest of the Population Census. This operation has been comprehensive in the first-stage units, the group establishments (with the exception indicated below), and by sampling in the second-stage units, persons residing in group establishments.

The framework of the Group Survey is the directory of groups obtained from different sources, using information available in the INE, both from the previous census, and from different surveys (for example, the Tourist Accommodation Survey or the DIDSS-08 Disability Survey) and from directories provided by external institutions, such as the Ministry of Defence.

The sampling design of the operation depends on the type of group. For the groups from the first group (hotels, guesthouses or the like), there was a module in the surveys on hotel occupancy for the month of October 2011, carried out by the INE, in order to determine whether there were persons (employees or clients) residing there regularly, in such a way that only those establishments were studied in the Group Survey.

A comprehensive study is done of the prior directory of group establishments, for all the establishment types, except hotels, guesthouses or the like, which are studied via sample. A threshold to study is established within each group, based on the size (number of persons residing) of the establishment, in those cases in which the number of persons in the group exceeds that threshold, a sample of persons is selected:

1. if the centre does not have more than 20 residents, they are all studied,
2. if there are 21 to 509 residents, a sample of 20 residents is selected,
3. if there are 510 to 800 residents, a sample of 30 residents is selected,
4. if there are 801 to 1,000 residents, a sample of 40 residents is selected
5. in the groups with more than 1,000 residents, a sample of 50 residents is selected.

For the selection of the sample of persons, the Kish method has been used, which assigns an equal probability to all the persons in the group.

An elevation factor is used for the centres, in order to process non-response, and an elevation factor is used for the persons. Hotel-type establishments and the like have an elevation factor that is different from the rest of the groups, on considering not only the adjustment for non-response, but also the probability of selection in the tourist occupancy samples.

For the purposes of the estimation of the characteristics, each one of the types has been considered as a stratum.

The estimator for the total persons in group establishments with a given characteristic X is expressed as follows:

$$\hat{X} = \sum_t \sum_{i \in t} f_i \cdot f_i^p \cdot X_i$$

Where subindex t represents the types of group, subindex i represents the group, f_i the factor of group i, f_i^p and X_i are the factor of the person and the total persons who have the characteristic studied in group i, respectively.

Factor f_i in the case of type-1 groups, is the factor assigned in the sample of the survey it is in. In the rest of the types, on being comprehensive research, it should have been $f_i = 1$, but due to the incidences presented during the fieldwork, there is a correction factor obtained as follows:

1. The groups, according to the final situation in the survey, are grouped as E surveyed, FA outside of the scope, DU duplicated, A not contacted and N refusal to participate.

Both FA and DU are framework errors, and are not a part of the research.

2. The correction factor for non-response due to groups that refuse to participate or that are not contacted is calculated **within each type** as follows:

$$f_i = \frac{E + N + \hat{A}}{E}$$

3. Prior to the calculation of the above factor, the number of groups considered to be A must be estimated, due to the fact that among them, there might be groups that are FA or DU.

The total groups A among those of a certain type are calculated in the following way:

$$\hat{A} = A \cdot \frac{E + N}{E + N + DU + FA}$$

Where A, E, N, DU and FA is the total groups classified as such within said type.

4. The groups may have a type in the framework that is different from the type obtained after studying them. The factors are made according to the type in which they appear in the framework, and the characteristics according to the type observed in the research.
5. The **person factor** within the group, given that the persons are selected with an equal probability within the group, is obtained as the quotient between the total persons in the centre and the total persons in the sample of that centre.

The estimated total persons in group establishments takes the following form:

$$\hat{P} = \sum_t \sum_{i \in t} f_i \cdot p_i$$

In which p_i the total persons in group i .

The estimation of the characteristics of the population resident in groups is affected by a sampling error.

The procedure used has comprehensively studied all of the group establishments (except those hotels in which the sample size is very large), and therefore, the calculation of the sampling errors may be approached by using the expressions of simple random sampling.

The following table shows the relative sampling errors (Variation coefficients) that are obtained for different totals.

TOTAL	CV(%)
5	77.46
10	54.77
25	34.64
50	24.49
75	20.00
100	17.32
250	10.95
500	7.75
750	6.32
1,000	5.48
2,000	3.87
5,000	2.45
10,000	1.73
20,000	1.22
50,000	0.77

Sampling error enables obtaining the confidence bracket, within which the real value of the estimated characteristic is found with a given probability.

The variation coefficient is expressed as follows:

$$C(\hat{X}) = \frac{\sqrt{\hat{V}(\hat{X})}}{\hat{X}}$$

Sampling theory determines that, in the

interval between

$$\left(\hat{X} - 1,96\sqrt{\hat{V}(\hat{X})} , \hat{X} + 1,96\sqrt{\hat{V}(\hat{X})} \right)$$

there is 95% confidence in finding the real value of parameter X.

8 Information collection, collection method, incidences and their processing, response rates

8.1. INFORMATION COLLECTION, COLLECTION METHOD

The information collection of the Group Survey was completed in two subsequent waves, according to the availability dates of the directories of the different types of centre. The first took place between October and December 2011, and considered 9,512 centres, excluding hotels and the like, religious, military and penitentiary institutions, whose directory was made available afterward. In the second wave carried out between January and March 2012, aside from continuing with the previous centres that had not yet responded, the remaining 3,835 centres of the aforementioned types were included.

The collection was carried out for all of Spain, from two Provincial Delegations of the INE, sending the questionnaire to be filled out by ordinary post, and offering several channels for mailing the information: the Internet, post, telephone, fax or email. The information received with errors was filtered using telephone contacts with the respondents, with a collection management application as support.

A first documentation mailing was posted to those groups selected to be part of the sample, including the following documentation:

- Survey letter of introduction.
- Print questionnaire, including the URL and keys for access to completing the questionnaire online, if the respondent wishes to use this method. Regarding the response paid so that the information may be submitted via this channel (post) if it is preferred.
- Sheet with the list of persons selected in the sample, containing instructions for the selection thereof.
- 900 telephone number, free of charge, for queries and the resolution of doubts of respondents, and the fax number, also 900, so that the completed questionnaire may be submitted free of charge, if they prefer to do this following this procedure.

This mailing was completed with two additional mailings, or more as necessary, claiming the completion of the questionnaire, or collecting the incidences relating to non-response (closures, outside of the scope, etc.).

The information received by post or fax was recorded with an application with controls over the completeness and consistency of the data, which led to telephone calls to complete the data in those cases in which the information received did not have sufficient quality to be processed.

Intense telephone support was carried out during the collection period. The interviewers made telephone calls for the following actions: location of ideal respondents in each group, verification of receipt of the documentation,

complementary information regarding the survey objectives, resolution of doubts and queries for the correct completion of the data, claims of questionnaires, resolution of inconsistencies, request for absent data following the recording and automatic validation of the questionnaire content.

8.2 INCIDENCES AND THEIR PROCESSING. RESPONSE RATES

The initial directory had a total of 13,347 group establishments, 2,679 of which belong to the first type, hotels and the like, from the research carried out in the Tourist Accommodation Survey. This directory was filtered from duplicates and those outside of the scope (addresses that do not meet the definition of group establishment) during the collection.

Questionnaires have been obtained for a total of 8,359 group establishments (62.7%), and for the remainder, different incidences have been obtained, which may be grouped as follows: duplicates (787 centres, 5.9%), establishments outside of the scope (2,715 centres, 20.3%), establishments that qualify as absent or not contacted (1,468 centres, 11%) and refusals (18 centres, 0.1%).

The questionnaires submitted to the establishments consist of one part to be completed by the establishment itself and the individual questionnaires for the selected persons, offering them the option of filling out and posting their own questionnaire. Individual non-response was 3.6%.

The response rates obtained for each response method have been the following: 63.5% by post, 30.5% by the Internet, 2.5% by telephone, 2.1% by fax, 2.3% by email or other means.

9 Processing the information

The information collected by the different channels habilitated for this purpose is integrated and controlled by a centralised collection and management application from each one of the two collection units established. Likewise, this application enables the recording and subsequent telephone filtering of questionnaires received by channels other than the CAWI, including a system for error detection and classification thereof, according to type, and for validation of questionnaires. Once the period established for the information collection has concluded, the files resulting from the control and management process of each unit are unified and newly filtered in a centralised way, in order to carry out the imputations necessary in case of lacking data.

10 Dissemination of the results

Dissemination of the results has been carried out in electronic publication format through the INE website, making it possible to obtain detailed results consisting of tables relating to population figures in group establishments, as well as different statistical tables regarding the characteristics of those persons resident in groups, or regarding the number and types of groups on a national level, by Autonomous Community or Autonomous City, by province and by municipality.

11 Resident population and population registered in group establishments, and comparability with the 2001 Census

In the 2011 Census, an important innovation was introduced at the time of counting the population resident in group establishments, which it is necessary to consider for comparing it with the 2001 Census, and also for analysing the figures of the population resident in family dwellings.

That person who has resided or plans to reside for a year in a group establishment is considered to be a resident, but questions are also asked of persons regarding this circumstance. Thus, it is determined that, though 444,101 persons reside, only 54% of them (241,187) are registered in the group establishment, with prisons being noteworthy, as barely 10% of the population resident there is registered.

Given that the population census is built using a file treatment that is, in turn, derived from the Municipal Register, those persons resident in family dwellings count where they are registered. Thus, the 202,914 persons (the difference of the two above figures) who, residing in group establishments, are not registered therein, are being counted as residents in the family dwellings where they appear registered, in order to avoid counting them twice.

This fact also differentiates the 2011 Census from the prior Census, in that the method for obtaining data, the figure for residents in group establishments was very close to the figure for the population registered therein, which hindered making comparisons, the more the figure for registered persons differed from the figured for residents. INEbase publishes detailed figures for residents and registered persons, by type of establishment, in order to enable analysis.

Annex I. Questionnaire

Annex II. List of persons