



**Economic and Social
Council**

Distr.
GENERAL

ECE/CES/2006/24
29 March 2006

ENGLISH
Original: FRENCH

ECONOMIC COMMISSION FOR EUROPE

STATISTICAL COMMISSION

CONFERENCE OF EUROPEAN STATISTICIANS

Fifty-fourth plenary session
Paris, 13-15 June 2006
Item 6 of the provisional agenda

SEMINAR ON POPULATION AND HOUSING CENSUSES

SESSION 2

The population census in France: from general census to “rolling census”

Submitted by INSEE, France

I. SUMMARY

1. In 2004, France introduced a new census method which combines two basic principles - a five-year rolling cycle of data collection plus surveys of municipalities of 10,000 inhabitants or more. This new method will yield detailed results every year at all relevant geographical levels, from the country as a whole down to the municipality or even the neighbourhood level (part II below).

2. The census has been redesigned in response to the ever-growing need for fresh demographical data, particularly at fine geographical scales. It also helps spread the financial and human costs of the census and improve organization and supervision (part III).

3. The census as it is now conducted in France meets the five fundamental criteria set by the United Nations (part IV).

4. The success of such a method depends on several factors, some technical (the availability of sub-municipality-level geographical information systems and the administrative data needed to keep them up-to-date), some methodological (sampling techniques in particular), and some political (extensive consultation with the municipalities, as partners in the census, and with census users in general) (part V).

5. Midway through its first five-year cycle, results from the French census are positive, but a great deal of work remains to be done (part VI).

II. METHOD

6. The first basic principle of this method is that data collection is spread over a five-year cycle, generating information each year relating to the median year of that cycle. Thus, in a given year y , the information output is based on data gathered in years y to $y-4$ and represents year $y-2$. The second basic principle is the use of surveys in the biggest municipalities, those with a population sufficiently large for a survey to yield robust information. This survey component has been introduced as a means of relieving the burden that the census puts on INSEE, municipalities and census respondents; it also makes for improved quality control of the collection process and in particular fewer omissions - always a bigger problem in large towns.

7. Given the great number and extreme diversity of France's municipalities - nearly 37,000, half of them with populations of less than 400 - a special sampling mechanism has been developed.

8. A census is taken of all municipalities with populations of less than 10,000, one fifth of them being covered each year.

9. The 35,750 municipalities with populations of under 10,000 - which account for half the population of France - have been divided into five groups. Every year, each of the municipalities in one such group are thoroughly surveyed. The municipalities in the first group are surveyed again five years later, and so on.

10. The five groups are balanced on the basis of some 10 criteria, some demographic (population, sex and broad age groups), others relating to housing stock (number of dwellings, number of principal residences). Balance is ensured at the national level and also for each of France's 26 regions.

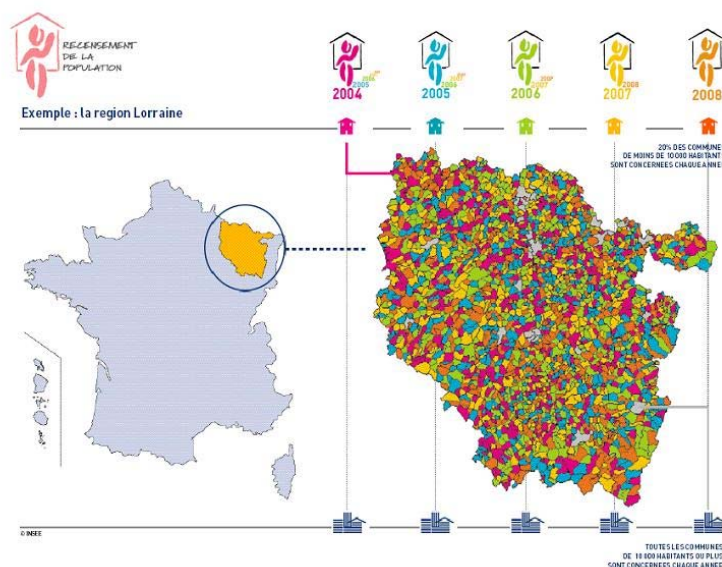


Figure 1. The five groups in the Lorraine region

11. The 900 municipalities with populations of 10,000 or more also conduct census surveys every year, but only of a portion of their populations. Each year's survey covers 8% of dwellings, selected from across the municipality. Thus by the end of the first five years 40% of the population will have been surveyed: enough to guarantee robust information on the municipality and its neighbourhoods.

12. The sampling frame in each large municipality is the relevant located buildings register (*répertoire d'immeubles localisés - RIL*), which is a list of all buildings, whether residential, administrative, industrial or commercial premises, identified and located using a geographical information system (GIS).

13. The RIL was originally compiled on the basis of the 1999 general census, and has been kept up to date since then using administrative data such as building permits and local tax records, and postal records such as Post Office address lists. It is checked by the municipalities every year before final validation by INSEE.

14. The census sample for each municipality is based on the subset of residential premises using the following sampling method:

(a) To avoid cluster effects, the largest addresses in the municipality are exhaustively surveyed; they are divided into five yearly groups;

(b) New addresses are also exhaustively surveyed, since there is no information on which to base a sample; these, too, are divided into five groups;

(c) Other addresses are divided into five balanced groups following demographic or housing-stock criteria, each group being uniformly distributed across the municipality (the same street will thus contain addresses in different groups); the sample of addresses to be surveyed

each year is drawn from the current yearly group, so that all the addresses to be surveyed (large addresses + new addresses + other addresses) account for around 8% of total dwellings in the municipality.

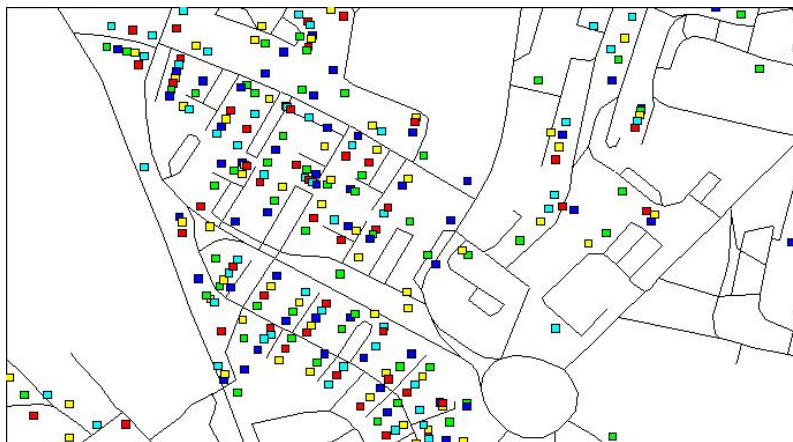


Figure 2. The five address groups

15. In this way, a comprehensive census is taken each year of one fifth of the municipalities of fewer than 10,000 inhabitants and 8% of the population of the municipalities with 10,000 inhabitants or more, i.e., some 4.5 million dwellings and 9 million individuals. After five years, all the small municipalities and 40% of the population in all the large municipalities will have been surveyed comprehensively. In all, 70% of the population is covered in the course of the five-year cycle.

16. The data collection protocol remains based on delivery and collection of a census form by an enumerator. Sampling is a major innovation, but data collection still follows the method used in general censuses, i.e., enumerator delivery and collection. Each household receives two kinds of form, one on the household plus individual questionnaires. The questionnaires are not very different from those used in earlier censuses. The household questionnaire requests a list of occupants and asks some 15 questions about the features and standard of the dwelling and the number of motor vehicles. The individual questionnaire contains 25 questions on age, sex, place of birth, nationality, place of residence five years previously, educational qualifications, occupation and place of work or education.

17. For principal residences which it has not been possible to survey (occupants cannot be contacted, are away for a long period or refuse to reply), the enumerator completes a non-surveyed dwelling form giving the presumed number of occupants, which helps in adjusting the population figures.

A. Method of estimation

18. The method used in municipalities of 10,000 inhabitants and over relies on a rolling average based on the samples over five years. From the aggregate of the five samples from

years $y-4$ to y , an average population per dwelling is calculated, which is representative of the situation at mid-period (year $y-2$). This ratio is then multiplied by the number of dwellings at the start of $y-2$, as given in the RIL, to obtain the population of the municipality.

19. For municipalities of less than 10,000 inhabitants it is necessary to calculate, at the end of year y , the population at the start of year $y-2$, since it is important to provide information with the same reference date as for municipalities of 10,000 or more. The municipalities of less than 10,000 inhabitants are surveyed over the entire course of the cycle, i.e., one fifth in $y-4$, one fifth in $y-3$, and so on. For those municipalities surveyed in $y-2$ the census results are saved, while for those surveyed in $y-1$ and y , the $y-2$ population is obtained by interpolation between the census survey and the last result published. For the municipalities surveyed in $y-4$ and $y-3$, the calculation involves extrapolation between the results of the census survey and $y-2$: this is based on local housing-tax data, which indicate any change in the number of dwellings per municipality, and is adjusted to take account of the differential between the change in the number of dwellings and the change in the number of inhabitants. The differential measured between the last two censuses is applied to the change measured by the housing-tax data to give the change in population.

III. REASONS FOR REDESIGN

20. There are two reasons for the census redesign: (a) the need for more up-to-date information, more regularly; and (b) the desire to spread the financial and human burden that used to be associated with a general census.

21. The need for fresh data was being voiced ever more explicitly by INSEE's clients. With general censuses getting ever further apart - nine years between the last two - it was becoming increasingly difficult to keep track of changes in France's population, whether shifts in society and the family, trends in peri-urban settlement, or urban developments. As major users of the census whose interest in such data had been sharpened by a gradual devolution of new powers under decentralization, local communities were particularly insistent: municipalities or groups of municipalities, for example, have acquired new responsibilities for urban transport, as have the departments in the areas of social welfare, help for the elderly, and investment in and operation of middle schools, and the regions in the areas of vocational training, town planning and investment in and operation of high schools.

22. The discharge and monitoring of these mandates required a regular flow of geographically specific, up-to-date information. The new census meets this need by providing detailed information every year at all geographical levels, from national down to municipal, and even the neighbourhood level.

23. Besides being widely spaced, the general censuses were also very cumbersome and their financial and human costs were not easy to absorb. It was for budgetary reasons that the census originally planned for 1997, for example, had to be put off until 1999. By spreading the load, the new method makes it possible to stagger the cost and operate with a steady workforce and budget, while the smaller scale makes for better organization and more manageable operations.

IV. INCORPORATION OF THE ESSENTIAL FEATURES OF CENSUSES AS DEFINED BY THE UNITED NATIONS

A. Individual enumeration

24. Information gathered separately on each individual or household surveyed. Starting in 2008, a detail file (one record per individual or household surveyed) will be produced annually by compiling the files from five years of surveys, thereby allowing cross-tabulation as in the general censuses. Each of the 9 million individuals and 4.5 million dwellings surveyed every year, adjusted by the sample weighting, figures in this file. Thus the detail file takes account of every individual.

B. Universality

25. All municipalities with a population of less than 10,000 are surveyed in the course of a five-year cycle. In these municipalities the entire population is enumerated, thereby guaranteeing universality as in the general censuses. For the municipalities of 10,000 and over, the entire territory must by law be covered in the course of a five-year cycle. The sampling frame comprises the total stock of residential buildings in the municipality and thus fully meets this requirement and complies with the principle of universality. It is updated annually on the basis of the Post Office's address list and building permit or tax record files. The update is a cooperative effort by INSEE and the municipalities. Using cumulative data from five years of collection does indeed enable the entire population and territory of each municipality of 10,000 and more to be described.

26. The cumulative sampling rate over five years is extremely high (40% in the large municipalities), and this guarantees accurate results. By contrast, in the general censuses some variables were applied only to a sample of individuals or dwellings. Similarly, in some countries the "long form" is sent to only a fraction of the population, in many cases around 15 or 25%.

27. The use of sampling, and the annualization of the data-collection process help INSEE, the coordinators responsible for the local censuses in each municipality, and the enumerators to concentrate their efforts, which makes for better-quality data collection. Enumerators in municipalities of 10,000 or more have lists of addresses to be counted. This is a very important point, since it means that they know exactly where they need to go. This method is more efficient than the sweep technique used in general censuses and, combined with the use of the non-surveyed dwelling form, helps to reduce and remedy the problem of non-response.

28. Lastly, groups living in collective households, people living in mobile homes and the homeless are comprehensively surveyed. In this way every individual living on French territory is counted.

C. Provision of data relating to small areas or small populations

29. Meeting the criteria of universality and individual enumeration means that this requirement is automatically fulfilled. The rate of survey coverage (100% in municipalities of less than 10,000, 40% in those of 10,000 and over), and the fact that the information is organized

in files of individual weighted data, allow highly detailed information to be extracted at fine geographical scales and for small populations. The cumulative file comprises some 45 million individuals and 22 million dwellings, which makes for very detailed tabulations.

30. The census will yield detailed information every year on each municipality and, for the larger municipalities, on blocks of roughly 2,000 inhabitants, taken together for statistical purposes, known as IRIS (*Ilôts Regroupés pour l'Information Statistique*).

31. Clearly, for smaller geographical areas, information updated every year is preferable to very old information, even if the random nature of a survey makes it somewhat imprecise. A rolling census will be far more useful in pursuing an urban renewal policy, for example, than a 10-yearly general census.

D. Simultaneity

32. Information gathered in different years is adjusted to apply to a single date at the mid-point of the five-year cycle. For the bigger municipalities, the average of the five samples is taken, and then “nudged” to match the number of dwellings given in the RIL for the median year. For smaller municipalities, the method used is one of interpolation or extrapolation between the census survey and the population reference date. Extrapolations are consolidated with the aid of data from the housing-tax registers, which give an indication of changes in the housing stock from year to year. With a five-year cycle, interpolation/extrapolation over more than two years is never necessary, so the figures are quite solid.

33. Census surveys take place on the same date each year (starting on the third Thursday in January) in all the municipalities concerned.

E. Regularity

34. The fifth essential feature is regularity: censuses should be conducted regularly, every n (e.g., 5 or 10) years. The French census will produce yearly findings, and thus meet this criterion. In addition, the availability of annual data means that users will no longer be tied to the dates imposed by 5- or 10-yearly censuses.

V. ENSURING SUCCESS

35. A far-reaching change of approach such as this depends on a number of technical and methodological factors and on collaboration with data users and census partners.

36. Technically, a high-quality sampling base is important: administrative files must be available for updating the base and instructions on their use must be drawn up. Here the partnership with municipalities is of great importance, for it means the contents of these files can be properly checked at the local level, which is very valuable because their knowledge of conditions on the ground is so much better. In the absence of administrative files, updating the sampling base is a slow and onerous business, requiring officials to be sent into the field to identify dwellings, as happens in France's overseas departments (DOM).

37. Another key factor in the success of the operation is the availability of administrative sources allowing the sampling base to be updated and survey results to be “nudged” if necessary.

38. Enough methodological expertise in surveying and modelling to be able to optimize samples, extrapolate survey results and produce synchronized statistics regardless of data collection date is also essential. Familiarity with the balanced-sample technique and its operation is vital for sample optimization and hence for budgeting and the quality of the estimates.

39. There must be an adequate, constant budget: a rolling census does not reduce census costs, it simply helps to spread them over time and secure a better “yield” since information is produced every year. The annual cost of the census in France each year, for example, is around one seventh of the cost of a general census. Once under way, the operation cannot be interrupted, for the cost of recovering one lost year would be very much higher than the budget for a normal year.

40. Thus the shift from a traditional to a rolling census relies on a number of important preconditions and thorough preparation (creating sampling frames, sampling, negotiating the budget, etc.), and therefore requires a considerable collaborative effort.

41. A key stage in preparing for the change is consultation with census users: moving from a “snapshot” system of 10-yearly data to a system of annualized data, a statistical probability representative of a period, is not something that can be done overnight. In France, there was also very close consultation with the municipalities, as long-standing participants in the census process and partners in the new census. Consultation must also involve political decision-makers at the national and local levels who rely on the census figures. If they are to accept the new system, they must have full confidence in the statistics service.

42. On the organization of data collection and dissemination, collective efforts are still under way. A national census evaluation committee has been set up at the National Council on Statistical Information (CNIS). Chaired by a Senator, and bringing together INSEE, the municipalities and census users (administrative bodies, associations of elected representatives, researchers, etc.), it monitors the processes of collection and supervision, suggests modifications to the various protocols and will be expected to give an opinion on amendments to the regulations governing census operations. As to dissemination, a working group of INSEE and the municipalities provides a forum for discussion on products for distribution and services at the sub-municipality level and beyond.

VI. PROMISING RESULTS MIDWAY THROUGH THE FIRST CYCLE

43. After three data-collection exercises it is possible to give a first progress report on the operation of the new census, and this is generally positive. There is widespread acceptance of the new method among both municipalities and the population at large. All municipalities took part in the censuses in 2005 and 2006 and the two that refused to do so in 2004 took part in the following years. As to the general public, the reaction one might have feared to a survey-based method - less interest than in the general censuses - did not materialize, thanks to highly effective publicity, nationally and locally. Lastly, INSEE has been able to provide initial population estimates to all municipalities surveyed of less than 10,000 inhabitants and around 100 of the

larger towns since late 2004. These results were sent out via the Internet two days before the start of the 2005 exercise. The procedure was extended in late 2005 to the municipalities of less than 10,000 surveyed in 2006 and to another 100 or so large towns, the results being published on 17 January 2006. Releasing the initial findings in this way has done much to convince users of the value of the new approach.

44. On the organizational level, the positive effects of annualized data collection are beginning to make themselves felt through better organization, computer applications, and greater professionalism among all involved, both at INSEE and in the municipalities.

45. Much remains to be done before the end of the first cycle, to build on the experience of these initial exercises and move up to cruising speed. Two areas will be considered for improvement at the next stage: the possible adoption of new data-collection methods and survey questions to meet users' new expectations; and the possible use of new administrative sources of data on matters such as employment.
