Economic Commission for Europe
Conference of European Statisticians
Group of Experts on Population and Housing Censuses
Nineteenth Meeting
Geneva, 4–6 October 2017
Item 5 of the provisional agenda
Meeting users’ needs

Factors impacting the wishes of census data users

Note by Statistics Estonia¹

Summary

Estonia plans to conduct the next population census in 2021 as a register-based census. In order to better serve the interests of users, it was decided to conduct a survey among the potential users to identify their expectations and wishes for the next census round.

During the new independence period Estonia has organised a census data user survey many times. From these surveys it appeared that need and purpose of using census data changes over time but not much.

¹Prepared by Diana Beltadze
I. Introduction


2. The user surveys have had various goals. For instance, in 2005 it was to determine the general demand for statistical census information in society, while in 2011 the survey focused on census data user preferences in using different statistics products.

3. In the context of organising the next census, the results of user survey (2016) have been applied for better (faster) servicing of users and improvement of products (output map application).

4. The usual purpose of a census user surveys in Estonia is to identify user requirements for the topics mapped in the previous census, to acquire an overview of the need for new topics, and to prepare a census program.

5. Due to the selected method of the previous census, user needs had to be analysed in the context of the results of the pilot census in order to assess the accuracy and completeness of answers, while also identifying the questions that were difficult, time-consuming or uncomfortable to answer.

6. In the light of the new census, focus was users needed clarification on the census outputs than can be feasibly produced on the basis of registers.

7. This document is dedicated to analysing how census user groups can be divided and among other things, the author raises the question of how to explore the needs of census users at the international level.

II. Users of census information

8. In Estonia census programs have been compiled in close relationship with users of census information, and it is one of the crucial steps in the preparations for a census, which is initiated well before the census itself. It normally involves a group of experts, who thoroughly review international experience from other censuses and quality assessments of previous results. However, the most important part of the work has been determining the needs of census data users.

9. In addition to comparability and harmonisation in the European and global context, the country’s census program is compiled to ensure continuity and comparability in time. In Estonia, this is required for monitoring the economic and social developments across all local censuses, which have been organised since the 19th century.

10. The contents of the program are generally prescribed by the Official Statistics Act. This has always been communicated to users.

11. By users it has been taken into account as a fact, that census program is based on the following sources:

   (a) Eurostat’s mandatory output (since 2000), which includes the main questions of the population and housing census;

   (b) Traditional questions of Estonian censuses (since 1881), such as ethnicity and native language;

   (c) Special questions for required for Estonian government or important research projects conducted in Estonia.

\(^2\) [http://www.stat.ee/user-surveys](http://www.stat.ee/user-surveys)
12. The first output includes 38 census variables. The questions of the second and third output are internationally recommended, but not mandatory questions (sometimes, they can be mandatory alternatives). The additional characteristics added for the second output have been determined in cooperation by the analysts of SE and Estonian experts, taking into account the public policy needs and availability or lack of alternative sources of information. This includes for instance the characteristics of ethnicity, as well as availability and address of secondary residence.

13. The third output was specified based on the results of the 2006 census users survey of potential users, where invitations for participation were sent to research institutions and local government representatives. The responses received enabled to assess users’ areas of interest, but they are not directly applicable for judging the relevance of census questions, because some indicators used for local governance do not necessarily have a clear connection with census questions.

14. All modifications or omissions of topics have had to be negotiated with all these users (research and educational institutions, state and local government representatives) to identify alternative options (registers, surveys) for obtaining the indicators and estimates required by users.

15. All modifications have also required an assessment of the gains in terms of time, labour and money savings in relation to increased risks to the quality and usability of results.

III. Census data users

16. The surveyed users of census data have mainly included local governments, ministries, research institutions, universities, consultation firms, and media organisations. The representation of entrepreneurs among respondents has been limited due to their large number and a relatively difficult process, which is required to establish contact. The lower representation of entrepreneurs among respondents has been offset by increased cooperation with entrepreneurs’ umbrella organisations in the preparation of censuses.

17. However, it should be mentioned that users of census data in international organisations have never been surveyed in Estonia. Furthermore, there has been no opportunity to conduct an in-depth survey among Estonian researchers and analysts, whose work results (indicators, projections, assessments, analyses, information in public databases) are used to shape and implement Estonian policies, incl. strategic steps. This situation is caused by lack of information on the relevant actors, which makes it difficult to decide who should receive the questions. At the level of local governments, Statistics Estonia has been unable to obtain information on the indicators, which would be required for better organisation of local life. The responses of local governments have been limited to straightforward answers to the questions of Statistics Estonia.

18. The users of data from population and housing censuses (PHC) can be divided in three main categories: (1) international organisations; (2) Estonian researchers and analysts whose work results (indicators, projections, assessments, analyses, information in public databases) are used to shape and implement Estonian policies, incl. strategic steps; (3) local government representatives, general public, media, educational institutions, etc.

19. Consequently, the PHC output can be classified to match the user groups as follows:

- Mandatory output (‘hypercubes’) specified in the EU legislation. This is required for making comparative analyses and projections in the EU and the UN, and for determining Estonia's rank among European and world countries;
• Characteristics, which are used (in addition to mandatory EU characteristics) to calculate indicators for public policy development and are published in the public databases of Statistics Estonia;

• Characteristics required for local governance or in-depth studies of specific fields.

IV. How do the needs of census data users change over time?

20. Organisation of a census depends primarily on the needs of users. The first survey on the use of census data was conducted in the framework of the 2005 survey of the users of population statistics. The main goal was to identify the purposes for which census data from 2000 were used.

21. The responses indicated that the results of the 2000 census had been used for various purposes: drafting of a national development plan and integration programme, analysis of the geography of the school network, mapping of migration and urbanisation processes, development of the public transportation network, creation of a housing development plan, assessment of the social infrastructure of local centres, comparison of regional and ethnic differences in employment, study of the need for Estonian language education among non-Estonians, drafting of a regional development strategy, etc. Many local governments had used this data for creation of general and detailed plans and in development plans (K. Alamets, 2006).

22. It was also established that, even though the Estonian census took place in March 2000, the data from this census were rather actively used – 55% of active users of population data used the census data in 2005 (Survey of Population Statistics Users 2005).

A. Census users’ survey in 2006

23. A specific online survey of census data users was conducted from 25 August to 25 September 2006. The questionnaire was available on the website of Statistics Estonia, and it was deemed to be completed after a respondent had entered their own data as well as the data about their organisation. At that time, the census preparation team established sub-goals for the survey, which included:

• Inform users about the initiation of a census and the envisaged topics to be covered in the census;

• Map the needs related to the topics of the previous PHC;

• Determine the need for new topics;

• Identify experts for developing the questionnaire.

24. The results of the survey were used to a great effect:

• For the preparation of the census questionnaire in 2007–2008 and a review of topic groups with experts;

• For understanding the needs of users who use the products of the Department of Population Statistics, e.g., mapping the topics that could be covered by surveys.

25. The following table provides an overview of the respondents to the survey. According to these data participation rate was higher for ministries (incl. government area agencies, county governments).
Table 1

**Participation rate by users group**

<table>
<thead>
<tr>
<th>Participation, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Local governments</td>
</tr>
<tr>
<td>Ministries (including government area agencies, county governments)</td>
</tr>
<tr>
<td>Others (research institutions, universities, consultation firms)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>

26. Main results of the user survey of 2006:

(a) The most important census data for users included data on places of residence, migration and commuting, sources of subsistence, employment and work (90% of the respondents needed these categories of data in their work). The data were needed for the organisation of routine work (social benefits, budgeting) within the organisation, for creation of development plans and strategies, and for assessing the impact of policies;

(b) In the category of residence and migration, the important topics for users included the date of moving to the current residence, previous place of residence, work-related commuting, the year of moving to Estonia and the previous country of residence. As new topics in the next census, users wanted to see the reasons for commuting and change of residence;

(c) In the category of sources of subsistence, employment and work, users primarily needed data on sources of subsistence, employment status, employer’s area of activity, and distribution of professions in the population. They suggested adding questions on distance work and amount of wages;

(d) More than 80% of users needed data on dwellings and buildings. The data were mainly needed for drafting development plans, plans, analyses and assessments;

(e) In the field of dwellings and buildings, users were mainly interested in owners of dwellings and household’s tenure status. Type of heating system, time of construction of the building and number of rooms were important for more than two thirds of users. As new topics, users suggested availability of an internet connection, the year of renovation in addition to the year of construction, and possible loan commitments associated with dwelling;

(f) More than three quarters (77%) of users needed data on education and language proficiency. Over half (55%) of users need these data for analyses, projections and overviews, while almost every fourth (38%) uses them for drafting development plans and general plans. The most interesting categories for users in the field of education and language proficiency include the attained level of vocational or specialised education, study status, education acquired in general education school, and foreign language skills. The suggested new topics concerned participation in refresher training and retraining;

(g) Health data (health status and restrictions) were important for two thirds of the respondents. These data are used for evaluating and planning social services, but also for drafting general plans and action plans. In this field, users were interested in long-term health problems and self-reported health assessment. A suggested new topic for the census questionnaire concerned the availability of public sports facilities;

(h) The population present at the moment of census and the type of institution, mentioned under other topics, were slightly less important for users; these data are used mainly for drafting development plans and area plans. Users had limited interest (30% and 18%, respectively) in more specific topic (year of birth of the first child, and religion).
B. Census users’ survey (REGREL) in 2016

27. Estonia plans to organise the next population census (2021) as a register-based census. As serving the interests of users is a significant purpose of any data collection, it was decided to conduct a survey among potential users of REGREL to ascertain the expectations and needs of potential users of the data.

28. As usually researchers, local government representatives, legislators, journalists and analysts were invited to participate in the survey.

29. In the 2016 survey, respondents were given a list of those census characteristics from 2011, which Statistics Estonia (SE) intends to measure in the coming census as well, and were asked about their use of these characteristics.

30. Goal of the survey was to identify user requirements for the topics of the previous census and to acquire an overview of the need for new topics. In short, to learn which census data are needed the most by users.

31. Sample size: 700 persons, response rate ca 40%. The survey was conducted from 19 April to 30 April 2016. The largest groups among the respondents included analysts or researchers (29%), associate managers (22%) and specialists (21%); 12% of the respondents were senior managers. 43% of the respondents represented ministries or other public authorities, 24% represented education and research institutions, and 21% represented county or local governments. The majority of the respondents (61%) were women (Ene-Margit Tiit 2016).

32. Compared to the survey contacted in 2006, the response rate improved in ministries and other public authorities, as well as in education and research institutions. The response rate fell by half compared to the previous survey among local government agencies; see Table 2.

Table 2
Number of respondents in census user surveys in 2016 and 2006 by the category of organisation

<table>
<thead>
<tr>
<th>Category of organisation in the survey</th>
<th>Number of respondents 2016</th>
<th>Number of respondents 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ministry, agency or another public authority</td>
<td>99</td>
<td>25</td>
</tr>
<tr>
<td>County or local government</td>
<td>47</td>
<td>97</td>
</tr>
<tr>
<td>Education, research or culture institution</td>
<td>56</td>
<td>19</td>
</tr>
<tr>
<td>Media company</td>
<td>10</td>
<td>–</td>
</tr>
</tbody>
</table>

33. The introductory section of the survey sheet explained the principles of a register-based census and notified the respondents that, in a register-based census, questions cannot be posed directly to the persons enumerated. The virtual census questionnaire can only be supplemented with questions covered by an administrative register from where the needed information (of sufficient quality) can be extracted and added to the census data. In doing this, Statistics Estonia gave respondents directions in a way, which was not perhaps strictly proper in the context of the survey. However, one of the objectives of this user survey, like in the previous one, was to inform users about the general policy adopted in the organisation of the census.

34. The substantive questions of the user survey were presented in several blocks. It started with a list of all planned topics of the REGREL questionnaire (EUROSTAT’s mandatory output characteristics) and each respondent was asked to mark the information that they had used in their work. The respondents then had to state whether they believed that more questions in this area would be needed in the census and what would be the appropriate interval of censuses in the future.
35. The first block of questions covered personal data in the census. One question inquired whether there are respondents who have not used any of the personal characteristics. The share of such respondents was 5.2%.

36. This was followed by two questions about relevance of the characteristics where a negative answer was also possible; see Table 3.

Table 3
Relevance of data on occupation and workplace

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes, %</th>
<th>No, %</th>
<th>No answer, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevance of data on workplace location</td>
<td>63.6</td>
<td>35.9</td>
<td>0.4</td>
</tr>
<tr>
<td>Relevance of data on occupation</td>
<td>54.1</td>
<td>45.5</td>
<td>0.4</td>
</tr>
</tbody>
</table>

37. This information was used in the meetings with different target groups in 2016 and 2017 to explain the need for register-based collection of data on occupations and workplaces.

38. The second block of questions pertained to households and families. The respondents were asked whether some of them had not used any of the listed characteristics. The share of such respondents was 32.9%.

39. The third block of questions covered dwelling characteristics and respondents who had not used any of the dwelling characteristics. The share of such respondents was 36.8%.

40. This was followed by a block of questions about applications for the information obtained. The survey sheet listed a number of activities and the respondents had to state whether they have used data for those activities.

41. In order to identify the purposes of use of the census data, the survey included a list of potential purposes (see Table 4) and the respondents could mark more than one purpose.

42. On average, each purpose was marked by slightly more than one fifth (21%) of the respondents and each respondent selected three purposes. Comparison of Estonian regions was clearly the most popular purpose. 6% of the respondents could not find a suitable purpose in the list provided.

Table 4
Purposes for which census data are used

<table>
<thead>
<tr>
<th>Description of purpose</th>
<th>Percentage of yes-answers, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Making projections, impact analyses</td>
<td>37.2</td>
</tr>
<tr>
<td>2. Legal compliance</td>
<td>14.7</td>
</tr>
<tr>
<td>3. Describing the initial situation</td>
<td>31.2</td>
</tr>
<tr>
<td>4. Investment analyses</td>
<td>8.7</td>
</tr>
<tr>
<td>5. Allocation of state resources</td>
<td>6.9</td>
</tr>
<tr>
<td>6. Allocation of local government resources</td>
<td>11.3</td>
</tr>
<tr>
<td>7. Policy monitoring</td>
<td>22.9</td>
</tr>
<tr>
<td>8. Comparison of Estonian regions</td>
<td>69.3</td>
</tr>
<tr>
<td>9. Comparison of countries</td>
<td>30.7</td>
</tr>
<tr>
<td>10. Responding to international queries</td>
<td>8.7</td>
</tr>
<tr>
<td>11. Research and studies</td>
<td>28.1</td>
</tr>
<tr>
<td>12. Professional development</td>
<td>15.2</td>
</tr>
</tbody>
</table>
43. The results indicated that applicability of data among users has diversified in the past 10 years.

44. The next set of questions covered the use of census outputs produced by Statistics Estonia (see Table 5).

**Table 5**

**Use of the products of Statistics Estonia**

<table>
<thead>
<tr>
<th>Use of SE products</th>
<th>Yes, %</th>
<th>No, %</th>
<th>No answer, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Database</td>
<td>85.3</td>
<td>14.3</td>
<td>0.4</td>
</tr>
<tr>
<td>Map application</td>
<td>42.9</td>
<td>55.4</td>
<td>1.7</td>
</tr>
<tr>
<td>Publications</td>
<td>64.1</td>
<td>34.2</td>
<td>1.7</td>
</tr>
<tr>
<td>Mobile app</td>
<td>11.3</td>
<td>86.6</td>
<td>2.2</td>
</tr>
</tbody>
</table>

45. It revealed that 93% of the respondents believed that the range of offered products is sufficient. 46. The next block of questions pertained to applications of the information obtained from the data. The survey sheet listed a number of activities and the respondents had to state whether they have used census data for those activities.

46. We also took a risk and asked about interest in new topics. The responses indicated that 18% of all respondents believed that the census could cover more subject areas than the mandatory international output indicators specified in the survey sheet. It resulted in a fairly long and diverse list of suggestions.

47. In order to identify the purposes of use of the census data, the survey included a list of potential purposes and the respondents could mark more than one purpose.

48. On average, each purpose was marked by slightly more than one fifth (21%) of the respondents and each respondent selected three purposes. Comparison of Estonian regions was clearly the most popular purpose. 6% of the respondents could not find a suitable purpose in the list provided. Suggestions for additional topics were made by 18% of respondents. (E.-M. Tiit 2016).

49. Notably, the respondents were generally very supportive of the organisers of the census and commended the high quality of census data, while also having very strong faith in the organisers’ ability to obtain all the different types of data they wish.

50. The opinions of different groups of respondents were characterised by three general suggestions/observations (*Ibid*):

   - Add characteristics that can be extracted from administrative databases (e.g., income, health data, migration information, marital status and children, education, etc.);
   - Find more opportunities for using the data of private structures to obtain up-to-date data on current processes, which could be used in decision-making;
   - Data on certain fields cannot be obtained in a register-based census.

51. According to the author (Ene-Margit Tiit) the first suggestion represents the group of respondents that has understood and accepted the ideology of register-based statistics, including register-based censuses. The second position is typical for innovative users, who
can envisage novel ways of collecting census data, even though they are rather hypothetical at this time. The third group of users represents the supporters of the traditional census who believe that a fully register-based census is currently not possible in Estonia and it should be supplemented with interviews, i.e., by returning to a combined census.

52. There were also suggestions on how to (further) simplify data downloads and analyses for the users, on having the opportunity to access source data for compiling customised cross-tables, and on linking the databases of SE with other registers.

53. As the result we can say that the respondents would like to receive additional information on wages, the statistics of the social welfare and health care field, waste handling, work- and study-related migration and general spatial mobility. They would like to see more basic characteristics in map applications. They were also interested in any information concerning big data.

54. Generally, users did not note the efforts made to compile census outputs, such as making comprehensive census data, covering all Estonia, available with comparisons to international data and the results of the previous census of 2000. The products created for regional users found greater recognition (most tables of the PHC 2011 database provide information at county level and a relatively large number of tables also cover lower settlement hierarchy levels), but the usage level of the respective tables was not very high.

55. The list of suggested new topics for the census program was rather extensive and much of it was similar to the user survey of 2006.

56. Similar topics:
   • Employer’s address field (i.e., the address of actual workplace, not employer’s registered address);
   • Means of transport used for going to work;
   • Number of ambulatory visits and hospital days from the health information system;
   • Previous/interim countries of residence;
   • Estonian citizens abroad;
   • Distance between residence and workplace, educational institutions, public services, grocery store, etc.;
   • Dwelling renovation status: when and what was renovated, was Kredex support used?
   • Native language and/or self-reported ethnicity;
   • Existence/absence of any special needs;
   • Ethnic origin, self-determination, home language;
   • Length of stay in the secondary residence. Addresses of main and secondary residences and workplaces.

57. New suggestions:
   • Actual place of residence and Population Register data;
   • Additional professional skills: driver’s licence, languages, computer skills;
   • Details about internet use. For in-depth analysis of the virtual gap; what smart devices are used (smartphones, tablets, etc.);
   • How many dwellings are owned by companies? How many dwellings owned by companies have been rented out? How high are the rent expenses for families who rent their dwelling?
- Ratios between family size, composition and number of rooms in dwelling;
- Relations between dwelling, workplace location and commercial and service establishments: What distances are covered? Where and how often people go / earn or spend money?
- The most important need is to receive direct information from persons about their place of usual residence;
- Religious preferences of respondents (identifies / does not identify with a particular religion, etc.);
- In addition to water supply and sewerage facilities, the issue of waste management in households should be studied, incl. options for the disposal of sorted waste (especially paper and biological waste) on the property;
- Involvement in sports and sporting opportunities for the entire population;
- Acquired profession, name of the school from which the respondent graduated;
- Existence of social insurance cover;
- Workplace organisation: ranges of employee numbers and form of employment (freelance, rental work, fixed-term employment contract, part-time work, etc.).

58. How have data needs changed over 10 years? It can be observed that there is a growing demand for census data, which could help to evaluate life quality, living environment and living conditions in the society.

59. Concerning the increased use of census data, census users surveys provided an overview of the diversification of the purposes of data use. It was not mentioned that most respondents in the 2016 user survey had a negative opinion about reducing the interval between censuses, with 38% of the respondents believing that censuses should be conducted with an interval of 5 years and only 5% of the respondents supporting an even shorter interval (Ibid).

V. Conclusion

60. In conclusion, it should be noted that users were very appreciative of the census outputs as a source of information; many of them said that it is the only reliable source of information on the issue of interest. This has been facilitated by active communication during censuses and explanatory work in presenting the data after census.

61. Secondly, the suggestions of information sources (such as private structures) and methodologies (combining registers with interviews) revealed certain naivety and optimism of users, as the achievement of expected results would mean unreasonably high cost in terms of resources and the effectiveness would remain questionable even then.

62. Thirdly, users would like the census to provide more data on the quality of life and assessment of the costs of the living environment.

63. Fourthly, there is a better understanding of feasibility of the use of census data and the areas of use of census data are more diverse than they were in 2006.

64. A challenge is to provide international data comparisons and create user-friendly outputs based on innovative solutions.
65. It is possible to categorize the use of census data by consumer groups. At the same time, it can be observed that there is a growing demand for census data which could help to evaluate life quality and living conditions in the society through new perspectives.

66. The relative importance of the identified factors varied by users groups according to the level of involvement with the census programme.

67. Factors that have influence on users wishes are largely determined by the:
   (a) Census method;
   (b) Specific purpose of study of users groups;
   (c) Absence of another kind of data sources for desired information;
   (d) Free and easily available detailed census database for users;
   (e) Level of cooperation or relationship with census data users; and
   (f) Purpose of data use.

References


Data set for 2006 census users survey.
Data set for 2016 census users survey.