



Economic and Social Council

Distr.: General
27 April 2012

Original: English

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Population and Housing Censuses

Fourteenth Meeting

Geneva, 24-25 May 2012

Item 5 of the provisional agenda

Internet data collection

Internet data collection in 2011. Population and housing census of Lithuania

Note by Statistics Lithuania ¹

Summary

The paper reviews the efforts and work undertaken by Statistics Lithuania in using the Internet for data collection in the 2011 Population and Housing Census of Lithuania.

Despite the fact that Statistics Lithuania had been using the data from the main administrative registers for the 2011 Census needs, a decision on the total enumeration (by the Internet and by face-to face interview through paper questionnaires) of the population was made due to the lack of information available in the registers. The e-Census was carried out using virtual IT resources.

The 2011 Census showed that the population is ready for e-government services and is willing to use them – more than a million (about 32 per cent) of the residents of Lithuania took part in the e-Census.

I. Introduction

1. Preparation for the Population and Housing Census of the Republic of Lithuania 2011 (hereinafter referred to as “Census”) started in 2007. After some pilot studies, a

¹ Prepared by Dalia Ambrozaitienė, Vanda Vaitekūnienė, Violeta Šalnienė and Rasa Balandienė.

decision was made that despite the possibility to use all the main State registers (the Population Register, the Real Estate Register, the State Social Insurance Fund's database and others) for the Census needs, the total enumeration of the population is needed due to high undeclared emigration of Lithuanian population, lack of data on educational attainment, occupation, etc.

2. In 2010, about 60 per cent of the country's population used the Internet; therefore, Statistics Lithuania decided to invite the residents of Lithuania to use this modern method and register themselves electronically.

3. Statistics Lithuania decided to use the outsourcing of some phases of work and services for the Census: the creation of a computerised system, the scanning and verification service; for the e-Census – to use virtual IT resources, i.e. opportunities provided by cloud computing technologies. All public procurement procedures were followed.

4. The Census was conducted in March–May 2011. It was the first time that inhabitants were afforded an opportunity to answer the Census questions on the Internet. The e-Census lasted for two weeks – from 1 to 16 March 2011. A special Statistics Lithuania e-Census portal was available for the population. On 17 March–4 April, the preparation for the population enumeration was carried out. On 5 April–9 May 2011, the second stage of the Census took place: those who had not participated in the e-Census at all were visited by enumerators and asked to answer the questions of the Census questionnaire. In June 2011, the processing of the data collected started. The Census data were scanned, coded, logical control was exercised; the data were compared with those available in State registers and other administrative sources. Detailed Population and Housing Census results will be published by Statistics Lithuania by June 2013.

5. Based on the provisional data of the e-Census and the population enumeration, at the Census moment (1 March 2011), there were 3 million 53.8 thousand permanent residents in Lithuania, which makes up about 94 per cent of the population of Lithuania estimated based on the data of the Population Register. In total, 1 million 415 thousand dwellings were enumerated. The residents of approximately one per cent of dwellings did not participate in the Census (refused to participate or did not let enumerators in at all)

II. General Information about the Computerized System of the Population and Housing Census of Lithuania 2011

6. A modern Census data collection and processing technology – a computerized system – was implemented at Statistics Lithuania

7. The computerized system should ensure: maximum speed; quality and accuracy; data confidentiality and protection; minimum of errors due to the human factor; economy of financial and human resources; comparability of the Census results and possibility to transmit them to Eurostat

8. It was planned to computerize the following Census processes:

- E-Census – to afford an opportunity for the population to submit data about themselves, their households and dwellings electronically;
- Preparation of the Census area maps and lists of dwellings for the population enumeration;
- Monitoring of the Census processes;
- Scanning and verification of questionnaires;

- Data processing: coding, correction, identification, elimination of double entries, imputation of missing data, anonymisation of records;
 - Data analysis and dissemination.
9. Two portals were created (distinguished by function) (see the Annex):
- a) The external Census portal (e-Census portal), affording individuals an opportunity to transmit the Census data electronically;
 - b) The internal Census portal, where the rest of the functions were implemented:
 - Monitoring of the Census processes;
 - Scanning and verification subsystem – for entering data from the Census questionnaires;
 - Image archiving subsystem – for storing generated e-Census and scanned Census questionnaires images;
 - Census data processing (data coding, logical control and data editing, identification of persons and buildings, elimination of double entries, transfer of data of temporarily present persons to the address of the permanent place of residence, data imputation, transfer of arranged data to the database of arranged primary data);
 - Data analysis subsystem – for Census data analysis, preparation of reports, creation of tables and maps (using GIS technologies).

III. Functioning of the E-Census

10. The Population and Housing Census was started on 1 March. The President of the Republic of Lithuania became the first participant of the e-Census and invited all the residents of Lithuania to participate in it. Everyone could fill in the e-questionnaire on the Internet, having connected to the Census system of Statistics Lithuania.

11. For the authentication of an individual, the Interoperability Platform of Public Administration Information Systems (IPPAIS) of the Information Society Development Committee under the Ministry of Transport and Communications of the Republic of Lithuania was used. The IPPAIS supports and improves as well as expands the use of its infrastructure in transferring public and administrative services to the electronic environment (e-Government gateway) and providing them. This system allowed user authentication through various possibilities (electronic signature, Internet banking systems, etc.).

12. Additionally, a special system allowing the population to connect to the e-Census portal using a personal document (passport or ID card) number was created.

13. After the authentication in their e-Census questionnaires, the respondents could see certain pre-filled personal data and data on the dwelling, taken from administrative data sources (Table 1). They were also able to correct the pre-filled data, if necessary.

Table 1. **Pre-filling of the e-questionnaires with data from administrative sources**

Administrative registers	Variables
Population Register	Sex, age, country/place of birth, legal marital status, citizenship
State Social Insurance Fund Board database	Current activity status, status of employment, workplace
Real Property Register	Type of living quarters, period of construction, useful floor space, number of rooms, type of heating, facilities (electricity, hot water, bath/shower)
Addresses Register	Address

14. After connecting to the e-Census portal, an individual could fill in the Census questionnaire in online mode:

- In the browser, the data fields of the e-Census questionnaire were displayed with the help of a wizard. The wizard was programmed in such a manner that data entry would be performed step-by-step, i.e., first, a few values had to be entered into the fields, and then the rest of the fields would be displayed for data entry depending on the previously entered values.
- In case of text information, it was allowed to select values from a classification / list of values. If the selection of values was not obligatory, one could enter information manually;
- After the e-Census questionnaire had been successfully filled in, the system changed the status of the Census questionnaire to “Submitted”. In the course of the e-Census, individuals were afforded an opportunity to correct their entries and confirm the newer version of the Census questionnaire. The user could see the errors detected on the portal and correct the information. If the questionnaire had no errors, the user could save it in his/her personal computer.
- One could download instructions on how to fill in the Census questionnaire. The elements of the portal had contextual help: if one had been holding the cursor on an unclear element for some time, a pop-out window (hint) with a brief explanation was activated.

15. The recent Population and Housing Census is one of the examples of cloud computing in the public sector of Lithuania. No IT infrastructure was purchased for the e-Census project: the Census was carried out using virtual IT resources, i.e. opportunities provided by cloud computing.

Table 2. **Preconditions for cloud computing**

Precondition	Situation in Lithuania
Fast Internet connection	Internet connection speed is among the highest in the world.
Information and data security	Security can be ensured by data centers with global recognition and auditing.
Reliable service provider	A few dozens of reliable IT service providers operate in Lithuania, including reliable data centers, software developers, installation companies, etc.
Privacy	IT service providers guarantee strict confidentiality requirements.

16. The e-Census required powerful servers; however, Statistics Lithuania made an innovative decision to hire rather than to buy the servers. The virtual infrastructure service allowed Statistics Lithuania to save a considerable amount of money – no investments were made in additional servers, which would have been necessary for the 2011 Census project, and Statistics Lithuania could use the necessary resources in a flexible manner.

17. Additionally, the IT service provider took care of the system integration, server maintenance and guaranteed a high level of data availability and security. The e-Census software and database were installed in a virtual server infrastructure. Cloud computing technologies allowed flexible adjustment to the number of persons taking part in the e-Census. Had Statistics Lithuania used the usual solution, it would have had to purchase a few servers the resources of which would be fully used only during peak hours.

IV. E-Census Results

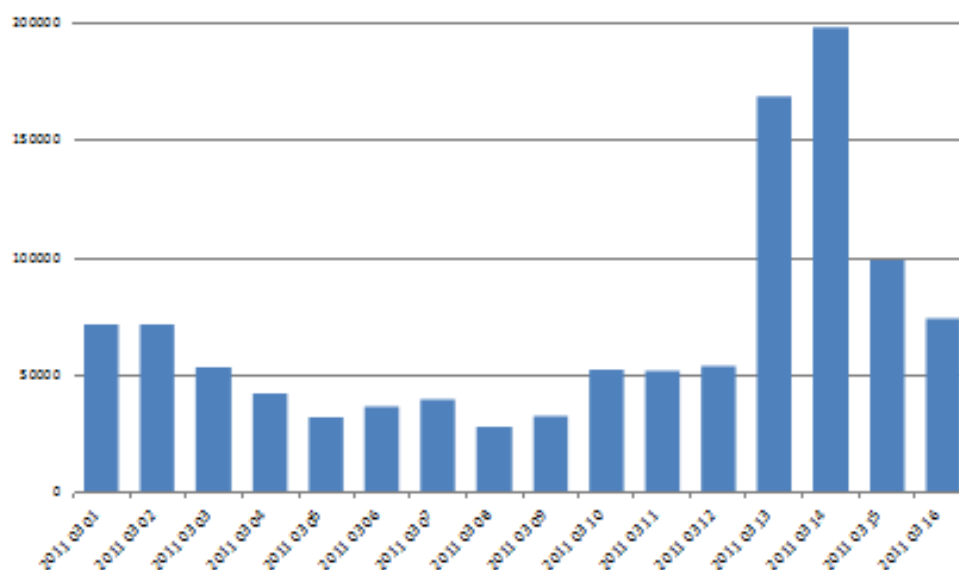
18. During the e-Census, 1 million 39 thousand persons enumerated themselves or were enumerated electronically.

19. Statistics Lithuania had a possibility for the everyday monitoring of the Census processes. Connections to the system via:

- Personal documents (passport and ID card) – 76 per cent;
- Internet banking system – 24 per cent.

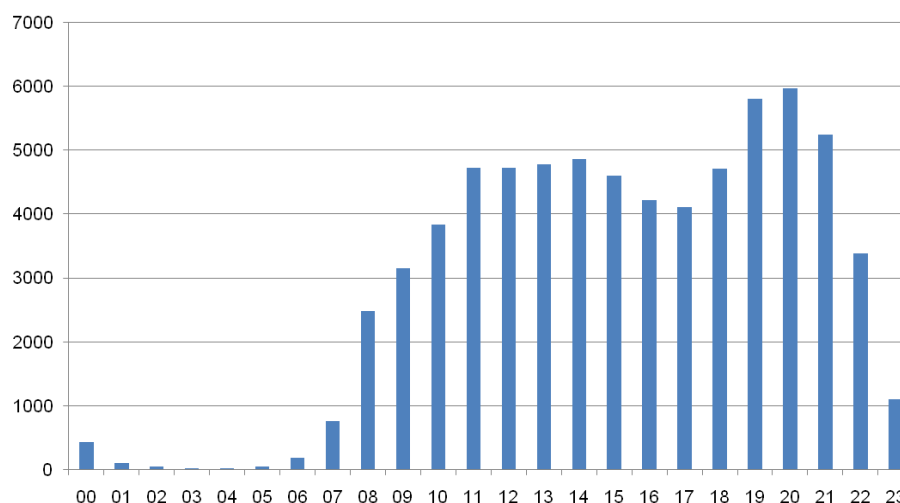
20. Residents were most active in the first and in the last days of the e-Census – as many as 17 thousand persons, on average, would be connecting to the system per hour

Figure 1. **Number of connections to the e-Census system**



21. Due to the extremely active participation of residents in the e-Census and at the request of the population, the time of the e-Census was prolonged: those wishing to participate in the e-Census could do it until 24:00 on 16 March 16.

Figure 2. Average number of connections to the e-Census system per day by hour



22. The Census specialists of Statistics Lithuania worked as consultants and provided consultations by phone and e-mail. There were possibilities for the population to come to the premises of Statistics Lithuania or other regional statistical offices and to fill in the e-Census questionnaire. On average, one specialist of Statistics Lithuanian gave advice to 172 respondents per day. Consultations were provided on weekends and on weekdays until 9 p.m.

V. Lessons Learned

23. Lessons:

- To have enough time for the preparation and testing of the e-Census questionnaire design and user-friendly functioning;
- To dedicate enough time to the testing of the questionnaire; an e-questionnaire should differ from the paper one (navigation between questions, broader explanatory information);
- To foresee time, resource, technical capacity reserves;
- The population should be consulted by an e-questionnaire consultant centre;
- More attention should be paid to the teaching of the population, explaining how to use the e-Census system.

VI. Conclusions

24. Conclusions:

- The 2011 Census showed that the population is ready for e-government services and is willing to use them – more than a million residents of Lithuania took part in the e-Census.
- The innovative e-Census was a successful project that saved time and money for Statistics Lithuania and the State budget.

Annex

2011 Lithuanian Population and Housing Census processes

