

Integrating administrative data in the statistical production process of population Statistics in Lithuania



Milda Šličkutė-Šeštokienė
Population statistics division
Statistics Lithuania

Changing profile of census

- 2001 – Traditional approach
- 2011 – Combined approach – administrative sources were extensively used:
 - Residents were able to fill-in electronic questionnaire (34%)
 - Enumerators visited the rest of population at home
 - Some data were derived from administrative sources – for prefilling questionnaires and for editing purposes
- 2021 – Completely register-based, no field-work:
 - Pilot register-based censuses:
 - 2019 – testing the system
 - 2020 – testing the algorithms of variables and data
- After 2025 census indicators will be produced annually

Coherence of Census with Population statistics

- Usual resident – the same definition is applied
- Population statistics based on Population Register since 2005
- Different estimation methods produce different results:
 - 15 000 difference in 2021
 - 180 000 difference in 2011
- After each census revision of population statistics is done taking into account census results
- Since 2023 the methodologies are synchronized (No difference between Census statistics and Population statistics)

History of administrative sources in SL

- Administrative data - a new field in SL but very quickly evolving:
 - In 1995, Statistical business register was established and the first sample survey was accomplished (before that, a complete enumeration of enterprises was always used)
 - In 2004, first administrative (social insurance) data were used for the production of official statistics at the estimation stage
 - negotiations for getting data took 2 years and also government were involved trying to get the data
- 42% of results published by Statistics Lithuania are based on administrative sources. In Social Statistics - 52,3%
- Wider usage of administrative sources is one of the top priorities of Statistics Lithuania.

Administrative sources in SL

- 182 Administrative data sets at the micro level
- 155 Administrative data sets at the aggregated level
- 42% of results published by Statistics Lithuania are based on administrative sources. In Social Statistics - 52,3%
- This numbers constantly changing

Administrative sources for Census 2021

- 34 administrative sources analyzed
 - Inventory of indicators
 - the quality of data was analyzed
- 19 administrative sources were chosen for census 2021
 - Population Register and Real Property Register are most important

Legal basis

- EU population and housing censuses regulations - KEY legislation which would be absolutely enough
- The number of national measures introduced - in order to facilitate procedures
 - Resolution of the Government of Lithuania
 - List of administrative data sources – with obligation for data holders to provide data to NSI
 - Law on reduction of administrative burden
 - Ensures the monitoring of administrative burdens and reduction of unjustified administrative burdens
 - Prohibits to collect data which is already collected by other institution

Quality issues

- Quality of administrative data sources is a state priority – a lot of institutions using the same data and quality has to be guaranteed in the original source
- Basic checks - Number of records, counts by different categories, dates and missing values
- Input data quality - Quality guidelines for administrative data prepared and approved in 2021, still needs improvements
- Output data quality - Quality indicators for statistics produces on the basis of administrative data sources, under development

Statistical register of population

- **Statistical register of population created**, but improvements is under development
- Better coherence with population statistics, history of vital events, more precise migration flows and other related statistics
- All samples are drawn from the same sampling frame (which is equal to census population)
- A variable for different survey comes from the same source using the same algorithm for estimation

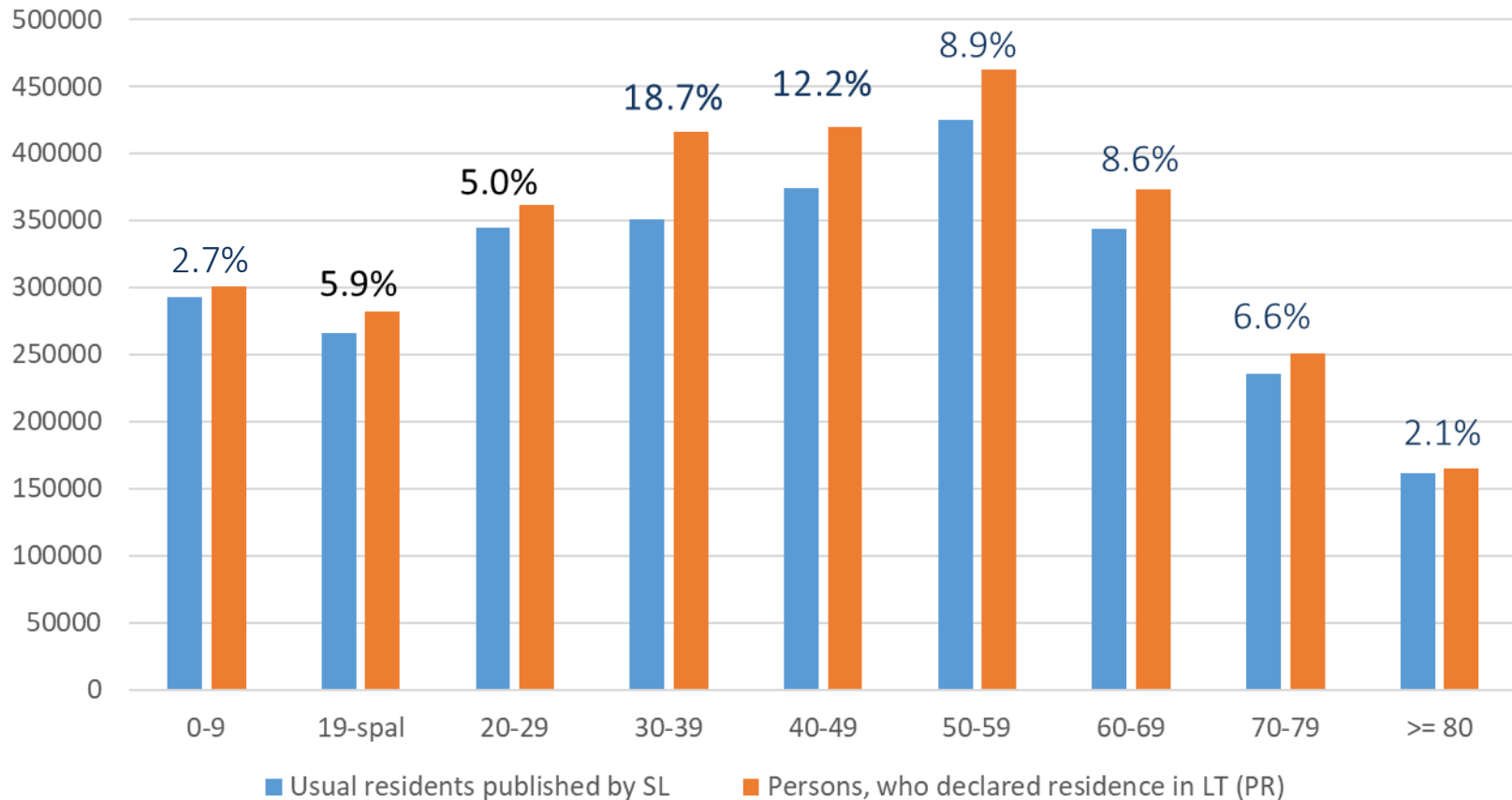
Technical implementation

- **2021** – database created using **Oracle Data Integrator (ODI)**: outsourced
- **After 2025** – **Palantir**: data warehouse will be developed by staff of Statistics Lithuania

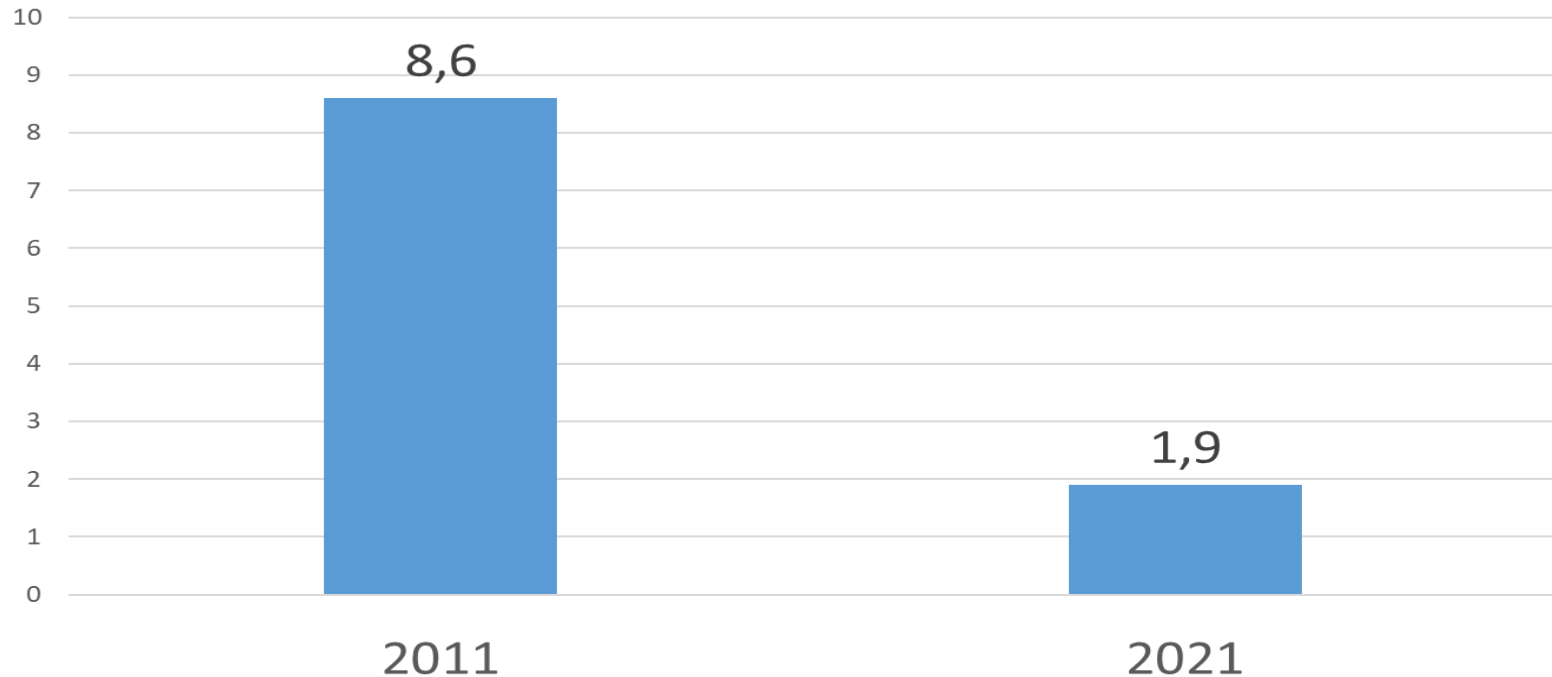
Biggest challengers

- Identification of **usual residents** (difficult to identify „signs of life“ ant „rate of stay“)
- **Household and family formation** (>20% of persons declared only the locality of residence but not the full address)
- Maintenance of the **integrated data warehouse**
- **Not all variables available** in administrative data sources:
 - Ethnicity, Mother tongue, Other languages, Religion – were collected by a separate survey

Number of usual residents compared to persons who declared place of residence in LT (total difference – 8.5%)



Census Budget 2011 and 2021, in mln. euros



Census date: 2021-01-01

Publication date: 2021-12-21 (provisional)

More information

<https://osp.stat.gov.lt/gyventoju-ir-bustu-surasymai1>



Official Statistics Portal



All indicators

News

Services

Classifications

E-library



Database news



News releases for this week



Home > All indicators > Censuses > Population and Housing Census

Census 2021



Census 2011



Census 2001

