HOUSEHOLDS AND DWELLINGS FOR REGISTER-BASED CENSUS: A GRAPH-BASED APPROACH

H. Visk[1], V. Levenko[2], K. Lehto[3], E. Maasing[4], E.-M. Tiit[5]

[1] Statistics Estonia
e-mail: helle.visk@stat.ee

[2] Statistics Estonia
e-mail: vassili.levenko@stat.ee

[3] Statistics Estonia
e-mail: kristi.lehto@stat.ee

[4] Statistics Estonia, Tallinn University of Technology
e-mail: ethel.maasing@stat.ee

[5] Statistics Estonia, University of Tartu
e-mail: etiit@ut.ee

Abstract

The 2021 census was the first in Estonia to produce all EU-mandatory census characteristics solely relying on the administrative data. One of the greatest challenges was caused by the low, just 80% accuracy of the place of residence data in the Population Register (Äär, 2017; Sõstra *et al.*, 2019) and its impact on households.

Register-based censuses define household as a set of people living in the same dwelling. When using the place of residence from Population Register to determine households, the resulting household and family statistics suffers from heavy bias towards more lone parent families and less married and consensual union couples' families (Statistics Estonia, 2017).

To obtain better statistics on households and families, we have developed a graph-based method which uses input from multiple administrative registers. We consider the people and addresses as vertices of a graph. A connection between two persons (such as marriage, parenthood, care leave) or a person and a place (such as registered address, real estate ownership) form the edges of the graph. A household is viewed as a subgraph containing household members and their dwelling. Then, determining households and their dwellings equates to finding densely connected subgraphs, in other words, to community detection.

To find connections between people or people and places we used data from 17 administrative registers. Each edge in the graph was assigned a weight describing the probability of people living in the same household or a person living on an address. The probability models were fitted on existing household data from Estonian Social Survey and Labour Force Survey.

The new framework was used to compute place of residence in the census. In the presentation, we will give an overview of the evolution of the register-based households methodology in Statistics Estonia with the focus on graph-based approach.

Keywords: register-based census, households, graphs, place of residence, population register

References

Äär, H. (2017) Coincidence of actual place of residence with Population Register records. *Quarterly Bulletin of Statistics Estonia*, **1**, 80–83.

Sõstra, K. *et al.* (2019) *Leibkondade kontrolluuringu (LEKU) tulemuste analüüs*. Statistikaamet. Available at: https://www.stat.ee/sites/default/files/2020-11/Leibkondade%20kontrolluuringu%20tulemused%202018.pdf (Accessed: 9 May 2022).

Statistics Estonia (2017) *Estonia's first register-based pilot census*. Available at: https://www.stat.ee/sites/default/files/2020-07/Results%20of%20the%20first%20pilot%20census%202016%20%281%29.pdf (Accessed: 3 June 2022).