



## **Moving Towards Register-based Censuses: Opportunities and Challenges** **Friday Seminar on Emerging Issues, 54<sup>th</sup> UN Statistical Commission**

*Geir Axelsen, Director General, Statistics Norway*

Thank you Chair and dear colleagues,

Norway went from traditional censuses towards register-based censuses over a sixty years' period. Reaching 2023 many countries have decades of failures and successes that can be shared and in addition IT technology and statistical methodology for register data have developed significantly since the 1960'ies.

Our main key takeaways are:

- The traditional census will be a good starting point as a data source when building a register.
- We recommend a step-by-step approach moving towards the register-based census.
- This development goes hand in hand with innovations and improvements in information technology, data management, statistical production, and products.

The road towards a register-based population and census system in Norway starts with the 1960 Census when we were able to establish a population archive that could be used to assign a personal 11-digit identification number for all residents. The natural next step was establishing a Central Population Register in 1964 partly based on data from the 1960 census. It was placed in a separate unit within Statistics Norway at the time and kept so until the beginning of the 1990ies.

The 1960 Census also brought a major investment in the IBM 1401 computer, one of several steps in moving from punch cards to modern technology. It constituted a major innovative and radical change at the time. With such investments the staff also developed relatively high competence in programming over time.

The 1970 census was still mainly conducted with paper questionnaires and enumerators in the field. But the traditional census data were used to check and update the Population Register. From then on, the register became the major source for all demographic statistics we produced. The 1970 Census was also used by Statistics Norway, in collaboration with the Ministry of Education, to establish the Register of the Level of Education - a statistical register which is still held by Statistics Norway.

The Population Register was primarily established for administrative purposes in 1964 and replaced several sector specific population registers such as the electorate and school enrolment, but Statistics Norway early saw the benefits of using the register for statistics production and influenced how it was set up by playing an active role in its establishment.

The 1980 census became the first census where administrative registers were widely used instead of traditional data collection. Several milestones were reached, and information on demography, education, and income, as well as geographical characteristics, were all collected from registers.

However, in 1990, the labour market data and data on households and housing were still collected by questionnaires for the census. A statistical register for labour market data was established and used as supplementary information when producing tables of labour market characteristics. But still, little or no information on households and housing were available from existing administrative data.

The 2001 census was the last census where traditional data collection was used, although we did introduce the option to respond via a web-form and twelve percent did so. At this point, the data on households and housing were the only remaining data collection area that was still based on a full coverage survey with a questionnaire for every household. The natural next step was then to develop a dwelling register. Statistics Norway had played an active role in arguing for such a register since the

1980-ies and in 1999 it was decided to allocate a unique address to all dwellings in the National Cadastre. The census forms were used to establish links between resident persons and their dwellings in the population register which opened for production of household statistics based on registers.

In the following period from 2001 to 2011, Statistics Norway, together with the national mapping authority and the tax administration, improved the quality in the Cadastre and the population register so that a complete register-based census could be conducted in 2011.

The first fully register-based census was conducted according to the plans in 2011. The 2011 Census project was organized as a project in consisting of a small group of 2-3 persons working full time with the Census, but in close cooperation with other units in-house. In addition to being the first census entirely based on administrative data sources, it was also the first census where an EU-regulation on statistics applied. The tradition of national dissemination and a unique census day in November was maintained, although almost every census topic was already published in the different subject matter statistics. A register-based statistics on housing was developed and published for the first time for the 2011 Census and has been produced annually.

The 2021 census was fully integrated with our regular annual statistics, with no specific census statistics disseminated and no public use of the term “census” nationally. The objective of the internal census project has been reduced to produce and deliver census statistics according to international recommendations to the EU and the UN. Hence no staff is working full time with the census in Statistics Norway, and it is organized purely as a data delivery project within division for population statistics.

The 60 years long history of development and innovation is obviously also a story about technology. We have taken advantage of software development and infrastructure investments, both within Statistics Norway and by partners in other governmental agencies. There is no need for other statistical organisations that aim for a more register-based census to start all over again and spend the next 60 years repeating our long history. In 2023 the available data processing power, the storage capacity and software for statistical production should provide a much better starting point.

There are some pre-conditions that have been key in the development towards a fully register-based census. This is firstly the legal base. Legislation provides a mandate for the use of administrative data sources for statistical purposes. National legislation must reflect the broadly held view that it makes good sense to take advantage of existing administrative data sources rather than re-collect data for statistical purposes. The statistics act gives Statistics Norway the right to access administrative data and to link them with other administrative registers for statistical purposes. This right was extended in the new statistics act of 2019 by including privately held data as well.

Secondly, I would point to the public approval and trust. The increasing number of administrative registers raise the awareness and need for protection of privacy for all citizens. Statistical use of administrative records normally involves linking data from several different registers, which may give the misleading impression that the statistical office can or want to study single citizens.

We must ensure that data are managed in a safe manner, and without the possibility for NSO staff to observe private information for separate citizens. Our mission, and legal basis, is to aggregate, study groups, geographic distribution, or development over time on relevant phenomena in the society. Confidentiality is not only a legal requirement, but one of our core values.

We also need proper information security systems and policies; we need to integrate privacy protection in all relevant parts of the production value chain. And we must communicate and explain our mandate and mission in the public debate. Our web pages must give transparent information about data sources and how we meet privacy concerns.

The benefits of using administrative registers are our ability to deliver more timely, more granular, and higher quality statistics – and with the potential to reduce operational costs. Summarized, we strengthen our innovation capacity.

Another major long-term benefit is that conducting a population census does not imply additional costs in census-years to our operations. We achieve a total coverage count at relatively low costs since the data already are collected for other purposes and we can access these data at any time. This gives in turn vast opportunities in the continuous production of statistics far beyond the scope of a census and with high frequency and granularity. We now do quarterly updates of the population statistics and are now working on possible monthly updates. On the population count only, we can produce this number every day in theory.

Another advantage is that registers can be linked together based on clearly defined keys (person numbers, business numbers, address codes etc.) and hence improved or new statistical products can be developed.

Furthermore, data from the individual administrative registers are usually consistent and of high quality, at least for administrative purposes. The fact that we on a regular basis publish statistics from these registers also contributes to improving the quality and serves as important feedback to the owners of the administrative registers. Still, we will not overcome all challenges. Registers, as well as traditional censuses, may be unable to be updated quickly and with accuracy for data related to migration.

Moving towards a register-based statistical system does not come without costs. Collaboration and coordination across governmental institutions and IT-investments are examples of investments needed, but by doing so stepwise there are good possibilities of moving forward.

Register-based statistics needs to be supplemented by surveys where needed. There are limited possibilities to include additional questions and variables when relying on administrative data only and some data needs to be collected directly from the respondents. In our view this holds especially for variables that can have a sensitive nature.

Another challenge of register-based statistics is that statisticians are to some extent restricted by the definitions and administrative practice of the authorities responsible for the registers. Furthermore, the statistical office is not close to the actual data collection and may not have the same knowledge about the precise data content, data processing and data quality among others. The best way to approach this problem is by cooperating extensively with the register owners.

The register-based statistical system is a data ecosystem, where the statistical office and governmental partners learn are mutually learning from each other. Innovations in one institution may impact another and one needs to be aware of the wider implication of possible changes in a register. IT competence and statistical competence can be exchanged, and the collaboration often goes hand in hand with dialogue about users' need for new statistics or new analyses.

Thank you.