Marja Tammilehto-Luode
marja.tammilehto-luode@stat.fi

Register-based statistics and geographic information

Register-based statistics are compiled from data that usually cover domains exhaustively, for example, all inhabitants, houses or businesses. The total coverage of the data makes it possible to aggregate statistical units depending on the accuracy of their location. All buildings in Finland have map co-ordinates, which make it possible to locate residents and businesses into them. So at a first sight there appear to be no special challenges in Finland to the compilation of statistics on small areas from a register-based statistical system. In theory, data secrecy is the only restriction to the production of small area statistics.

However, there are also other challenges to the production of good quality small area statistics from a register-based statistical system. In my presentation I will discuss these challenges from two perspectives: the register-based statistical system itself and the use of geographic information in statistics production.

A register-based statistical system has many pros and cons. It has been said that the results from a register-based census are at least as reliable as the results from a conventional census made by interviewers or questionnaires. However, there is certain controversy about how the quality of a register-based system should be described and documented. There is no existing theory for assessing the accuracy of statistics based on administrative registers (Platek and Särndal 2001). As a matter of fact, there is lack of methodological knowledge about the quality of register information, e.g. nature and meaning of errors and missing information. In my presentation I will compare the quality factors of a sample survey and a statistical register and discuss some special characteristics of statistical registers and register-based studies.

When geographic information is used in statistics production, the location of statistical units and /or statistical regions does not only help in the classification of data. Accurate locational identifiers of statistical units, such as building co-ordinates in Finland, make it possible to delimit statistical areas flexibly but also to calculate distances between objects and formulate comparable density indicators between different areas as examples. However, mappable statistics also present new challenges. Geographical information possesses quality factors that cannot be measured or managed by statistics. Instead of uncertain information one could talk about imprecise information (Niskanen 1998). In my presentation I will consider why geographic information is an essential part of a register-based statistical system and discuss major challenges to the use of geographic information in statistics production.