

## **The Population and Housing Census in a Register Based Statistical System**

Anita Lange\*

Statistics Denmark, Copenhagen, Denmark [anl@dst.dk](mailto:anl@dst.dk)

### **Abstract**

The population and housing census has for decades been the back bone in the social statistics in most countries, and so it has been in Denmark. However, a traditional census is an extremely costly operation and, especially back in time, a long lasting process, which makes data less useful. On this background the direct use of administrative registers in statistics was introduced in Denmark in the 70's. The paper describes the way to the world's first totally register based population and housing census in 1981, and how the use of administrative registers changes the production of statistics and also the demand for statistics. In the register based systems most statistics can be produced annually or more frequently, for example can population statistics in principle be produced daily. The first register based census in 1981 turned out to be also the last census published in Denmark. Denmark still undertakes censuses based on registers, but only to fulfill international commitments. So the conclusion is that in a register based statistical system, like the Danish, the census has lost its importance and hardly exists.

Key words: administrative registers, frequency, production time

### **Introduction**

The first population census in Denmark dates back to 1769 and since 1840 censuses have been conducted regularly, first every 10th year and from 1901 every 5th year. Housing information for cities has been collected since 1911 and in 1955 the housing census became a part of the population census, covering all country. The census was the most important, and for a long time, the only source for social statistics. The census delivered bench mark data on demography, family and household composition and labour market conditions for the total population, and was the main source for statistics on small areas. The content of the census has changed over time, but there has been some fundamental information that has been the principle elements in all censuses. The number of people, their sex, age and marital status, their place of living, the family or household composition, the source of living/income - often given by occupation and industry- and education, these are the most important variables in a census. However, conducting a traditional census is an extremely demanding process as regards both organization and costs. Already in the 60's there was an ongoing discussion among the Nordic statisticians about how to use administrative data as source for statistics to lower cost and increase the efficiency in the production. In 1968 the Central Population Register (CPR) was established in Denmark with a unique number (PIN) identifying every person in the population. This was the beginning of the extended use of administrative data in social statistics in Denmark.

### **From a traditional census in 1970 to the start of a register based census in 1976**

Following the 5-year rhythm there should be a population and housing census in 1970, but the introduction of CPR had resulted in a widespread expectation that censuses from now on would be unnecessary. From CPR it was possible to extract information on the population's address, sex, age, place of birth, citizenship and marital status, but this only covered the most elementary data from a census. Information like housing, employment and education was missing. Moreover there was still a lack of experience in transforming data from administrative registers into statistics. On that background it was decided to carry through a traditional census in 1970, but the intention was to continue the work with register based statistics towards a register based census. This last traditional census in Denmark in 1970 underlined some of the problems in census taking. The time and resource consuming work in collecting the data, register and check them, code them, compile tables and publish them was evident. Even if the work with distributing and collecting forms was done in the municipalities, more than 200 people were hired in Statistics Denmark to handle the coding and data checking. There was a comprehensive publishing program connected to the census. The first publication with data from the 1970 census was published in 1972 and the last one in 1977. In a changing society seven years old data are a problem.

A necessary condition for substituting the traditional census with a register based one was the establishment of new relevant sources for variables like housing, employment and education. The trends in society supported the development in that direction. Already in 1970 tax at source was introduced and a lot of different tax registers were built and in the 70's a number of other administrative registers were introduced in Denmark. Of vital importance was that all registers on persons used the PIN as a key, which made it possible to match information from the different sources. An administratively oriented business register (CBR) was set up in 1975 with an identification code (CBR-code) for all business units on enterprise level (legal level).

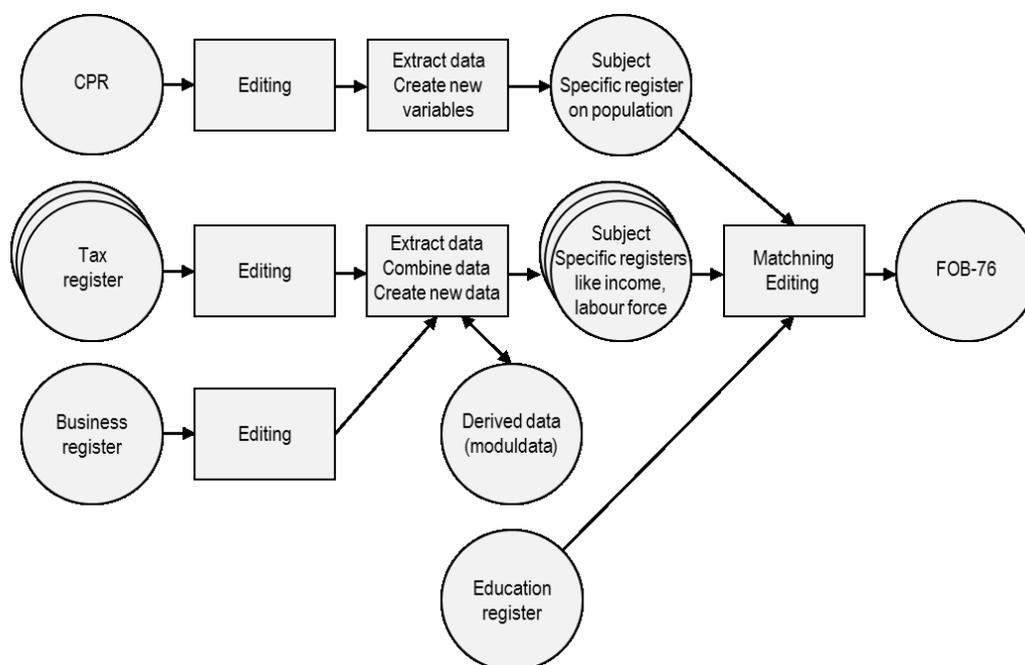
In Statistics Denmark an intense use of different administrative data started. At an early stage it was decided not to organize data in one all-comprising data base but in a system of sub registers with the common identifier (PIN) in all of them. As a rule the sub registers content was taken from one administrative register.

Already in 1970 the first annual population statistics were compiled based on data from the Central Population Register with distribution by sex, age and marital status. As the address in the register allows aggregating data at any regional level the statistics could be published for local areas. Also statistics on families and households (dwelling households) were developed in the early 70's by using information in CPR. A register on the population's educational attainments was established based on data from the 1970 census supplemented by the current education statistics. From the tax registers not only information on income could be compiled, but by matching data on enterprises from the Business Register (CBR) with a register on employers' payment of wages and salaries to their employees, a kind of activity distribution for the economic active population could be established. Information on people's occupation could be found in the tax-registers, but the quality was not the best as the information was depending on the single persons updating. So, in 1974 it was possible to cover a considerable part of traditional census information by using administrative data as

primary sources. An exception was housing data, where administrative data did not exist. However, it was decided not to carry through a traditional census in 1975/76, but replace it with a population census, based on the existing register sources. By combining information from those registers it would be possible to produce a reduced population census but without any housing information.

Combining data from several administrative sources was at this point still a relatively new exercise and conflicting information in different registers for the single person was one of the challenges. For example a person could receive pension in one register, but also have an income as employee in another. In those cases it was necessary to have a set of rules for how to handle the situation. Developing and testing different set of rules, together with improving methods for matching and checking large amount of data were important issues in preparation for the 1976 population census.

Figure 1. The 1976 Population census



### A totally register based population and housing census 1981

The 1976 census was, as it is mentioned, not complete. Not only was housing information missing but information on education was without details and the information on economic activity was taken from the enterprise level and not from the local workplace of the person. However, this census proved that the thought of a census completely based on data from administrative registers was a real possibility. And when there in 1976 was a political decision to establish a Building and Dwelling Register (BDR) for administrative and statistical use, it was clear that a traditional population and housing census was not any more a real alternative. The next population and housing census was scheduled to 1981 so there was a time limit for the continued work to bring the register based census up to level where it could be a real substitute for a traditional census. For the 1981 census there was an EU agreement on the variables to be included and the output to be delivered, so it was a closely defined project that should be carried through.

To solve the problem with the missing link between the person and the local work place, that in 1976 caused the inaccurate distribution of the population by economic activity, the so called workplace project was established. In this project an annual enquiry to all private enterprises with more than 10 employees was introduced to map all local workplaces and get the information into the administrative systems. For the public sector the link between workplaces and employees could be derived from pay systems in central and local government agencies. The workplace project was also a condition for future commuting statistics and structural business statistics.

To improve the information on people's occupation in the tax register, data from a number of different registers was collected to supplement this information.

The continued development of the education statistics made it possible to get detailed information on attained and ongoing education.

The Building end Dwelling Register was established with basis data from the general public assessment of real property for 1977, and the updating of the register was the responsibility of the local government authorities. It was decided that the single dwelling should be identified by a 17 digit long address code (digits and letters). This code was the same as the address code used in CPR for persons, and in that way it was possible to match persons with dwellings. In the beginning the matching between CPR and BDR was not without problems. As both registers are updated locally many different people have hands on it and with a 17 digits long code there are possibilities for mistakes. However, those initial difficulties were not long lasting.

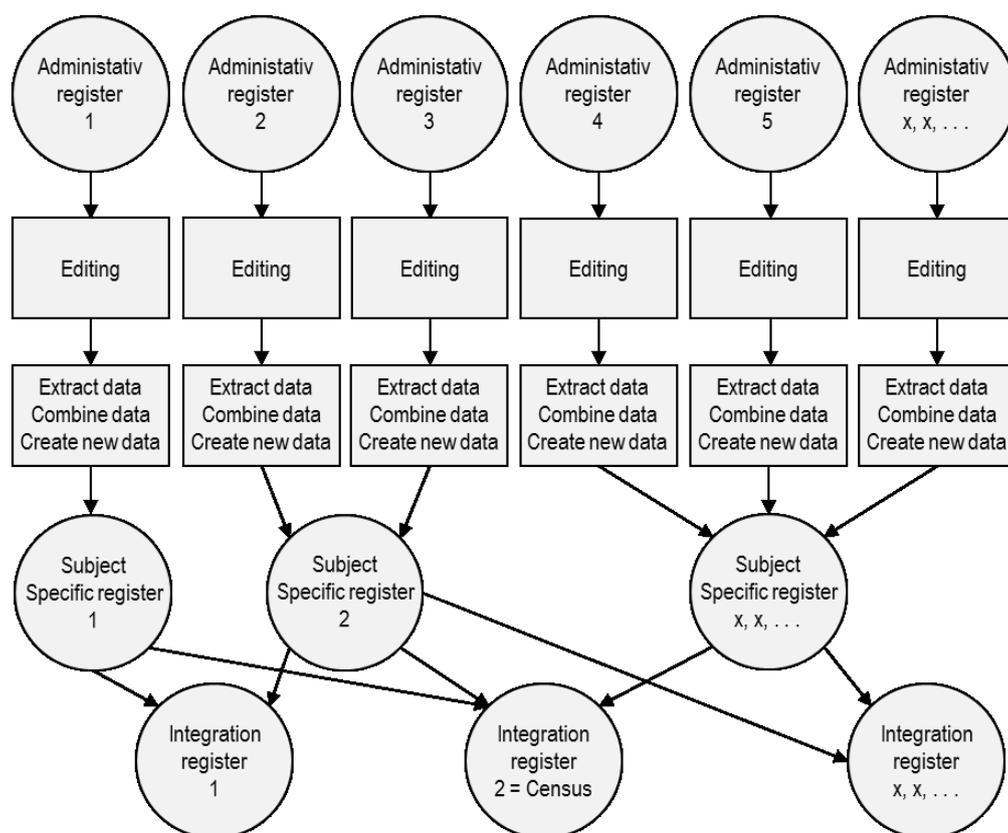
In 1981 it was possible- by using the three identifiers, PIN, CBR- code (including workplace code) and address-code - to combine information on persons, dwellings and employment from different administrative sources. In other words the basis for the first register based population and housing census was a fact. As different administrative registers not are updated and ready for use at the same time the last primary data for the census reached Statistics Denmark at the end of 1981 and the beginning of 1982. For some of these data processing had not been tried out earlier and not surprisingly this led to some delay in the final compilation of census data. So the timeliness in this first register census was not much better than in the traditional census. When data from the traditional census in 1970 are compared with data from the register based census there is one unit - the housekeeping household - that is missing in the register based system. In this system it is only possible to define dwelling households. In a country like Denmark, where only less than one per cent of the households includes more than one family, the dwelling household definition is satisfying.

### **The Population and Housing Census in a register based system**

There is no doubt that the goal to develop a register based population and housing census was of crucial importance for the progress in utilization of administrative sources in the whole area of social statistics. The work with the different administrative sources had shown that in most of them a lot more relevant information was available than was directly needed for the population and housing census. So at the same time as the basis for a census was established a number of subject specific statistical register were developed. These registers, like for example the population

register, the register of income, the dwelling register, the education register and the register of labour force employment, they all included information that was a necessary part of the census, but moreover a lot of additional information. Some of the subject specific registers were based on a single administrative register, others on a combination of two or more registers. Different administrative registers are not updated with the same frequency, for example are most tax registers only updated once a year while CPR is updated daily. Consequently it became possible, depending on updating frequency, to compile detailed statistics with relatively short intervals in a number of the most important statistical areas. For example has population statistics been compiled and published quarterly since the 80's. This development of the subject specific statistics made the demand for census data disappear almost completely. In the Danish statistical system based on administrative data it is possible to compile a population and housing census every year, but the census cannot compete with the subject specific statistics, not in timelines, nor in particularity. That is way the world's first totally register based population and housing census in 1981 was both the first and the last ever published in Denmark. Denmark still compile censuses based on administrative registers, but only to fulfill international commitments and the data are not published by Statistics Denmark. The census is now only one of several so called integration registers in a statistical system where administrative data, transformed into statistical data, are used and reused in a number of statistical products.

Figure 2. A simplified picture of a register based statistical system



In the European population and housing census from 2011 are the Danish results now ready to be delivered to Eurostat. This Census has been produced by two persons working on the project since 2011. The cost of the two persons can of cause not in any

way be seen as the cost of a register based census, as a prerequisite for the process is a number of well-functioning administrative registers and a fully developed statistical system built on these registers.

### **References**

Jensen, P. (1983) "Towards a register based statistical system- some Danish experience", Statistical Journal, Volume1, Number 3,341-365