

The Finnish Social Statistics System and its Potential

**Life after the Census:
Using Administrative Data to Analyse Society
Wednesday 9 May 2012, Belfast
Kaija Ruotsalainen, Statistics Finland**



Contents

1. Background
2. The basic units of register-based statistical system
3. Codes for combining the registers
4. Sources of Population and Housing Census in Finland
5. Data warehouse of social statistics system
6. New possibilities

Register-based Data Collection

- Long experience, since the 1970 population census
- first in social and demographic statistics, nowadays more and more also in business statistics
- about 96% of input data comes from administrative sources (as measured in number of statistical units times number of variables)
- direct data collection also very important
- these two methods are complementary to each other

Data collection

- Indirect data collection (proportion about 97%)
 - taxation registers
 - population register
 - Social Insurance Institution's material
 - central government accounts, etc.
- Direct data collection (proportion about 3%)
 - web questionnaires
 - paper questionnaires
 - interviews

Some factors facilitating the increased use of administrative registers in Finland

- Use of uniform identification numbers
- Administrations own interest in building nation-wide databases
- Legal basis
- Growing need for statistics
- Acceptance of the population
 - cost efficient
 - reduces response burden

Statistical basic registers



Organisation number

Statistics Finland



*Personal identification
number (PIN)*

*Population register
centre*

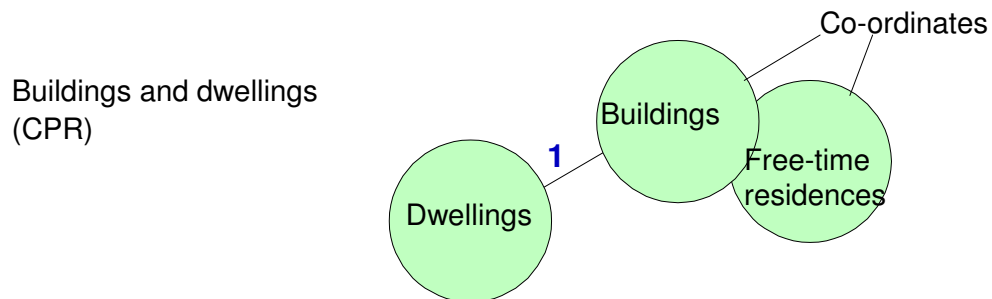


*Building/dwelling
number*

*Population register
centre*

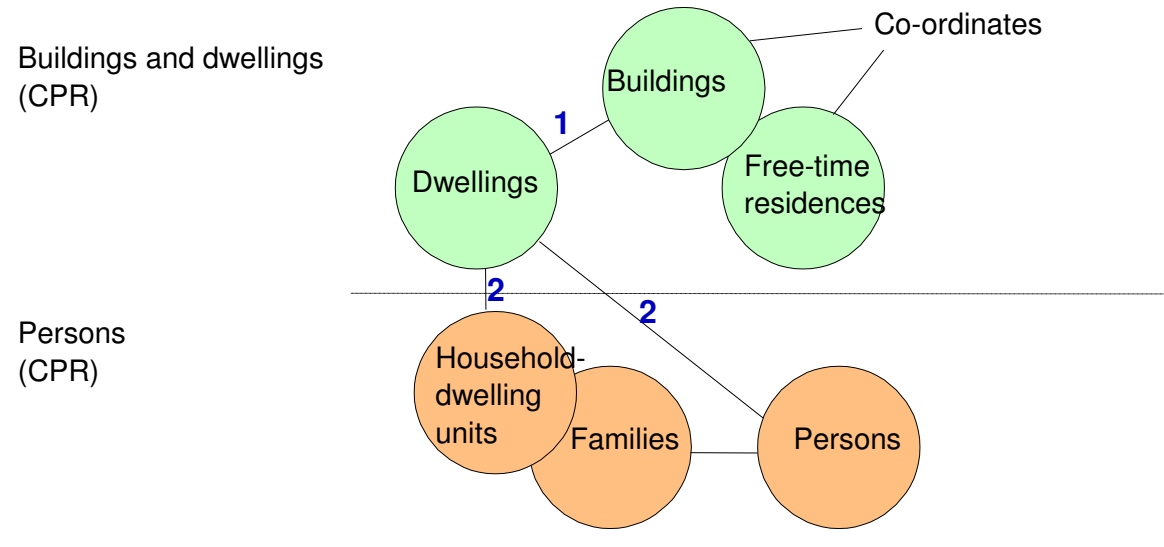
The basic units of register-based statistical system

- 1 Building code



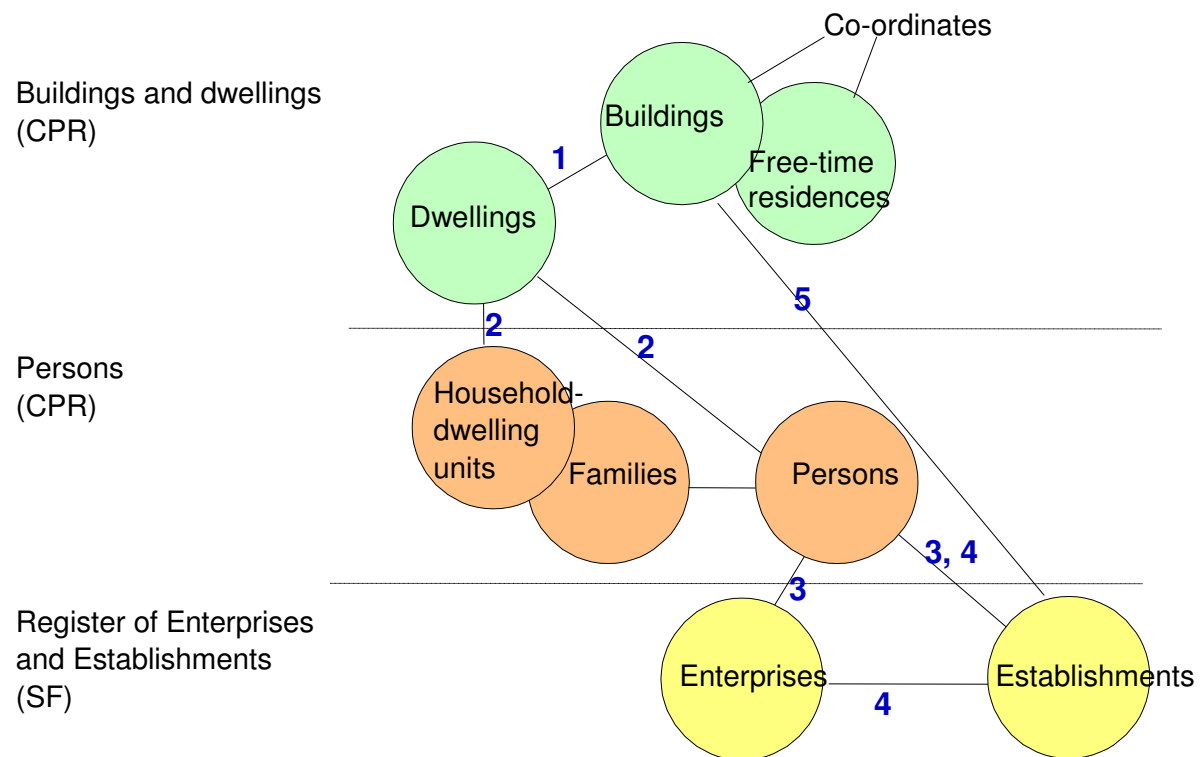
The basic units of register-based statistical system

- 1 Building code
- 2 Domicile code



The basic units of register-based statistical system

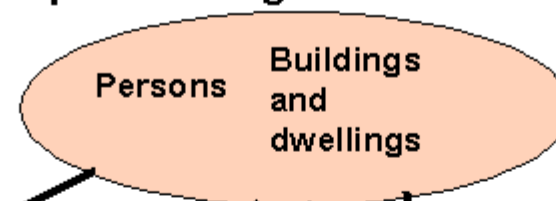
- 1 Building code
- 2 Domicile code
- 3 Enterprise number
- 4 Establishment number
- 5 Address



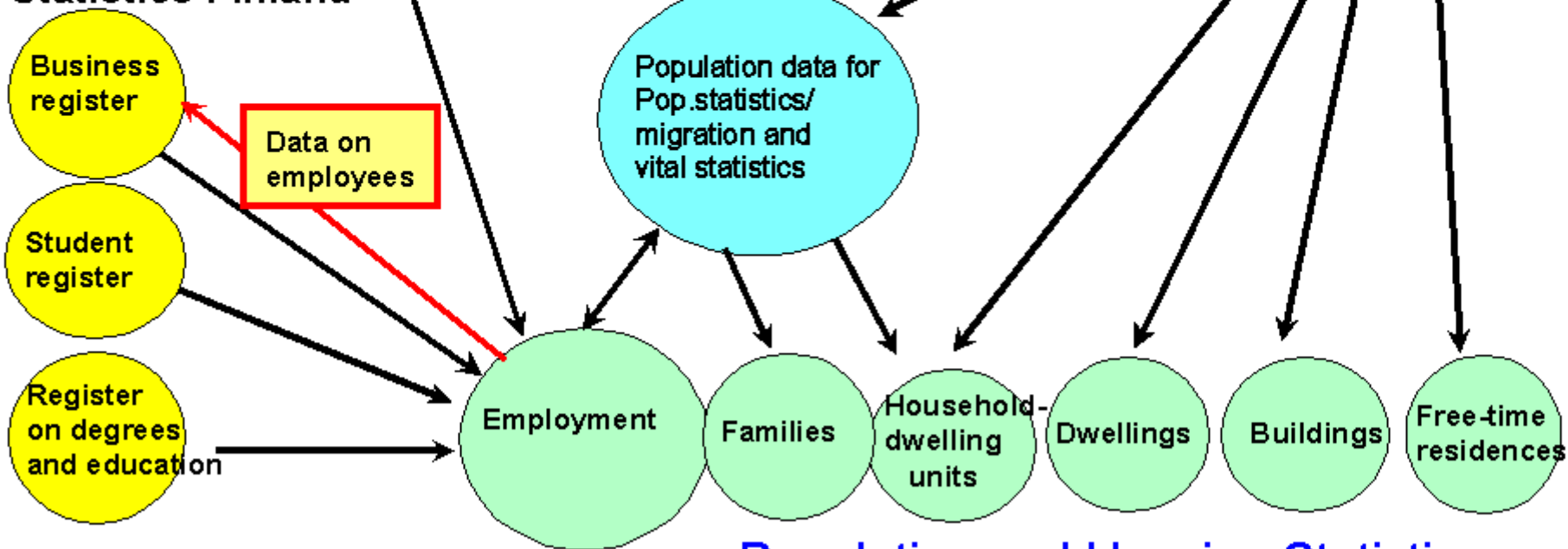
Administrative registers



Population Information System of the Population Register Centre



Statistics Finland



Population and Housing Statistics

GIS and register-based census

- Traditionally, the most important regional unit in statistics has been the administrative area. The area code system for administrative areas has been the cornerstone of regional statistics.
- However, administration is dynamic and keeps changing, which means that the boundaries of administrative areas are also constantly changing => may occur difficulties to keep up these changes
- The Finnish register-based statistical system is therefore point-based rather than area-based.

GIS and register-based census

- This building-based code system with its coordinates has provided a solid foundation for reliable and flexible statistical areas.
- Despite major changes in administrative areas, it is still possible to produce time series for different regions.
- The adoption of map coordinates for buildings has also paved the way to more flexible determination of statistical areas.

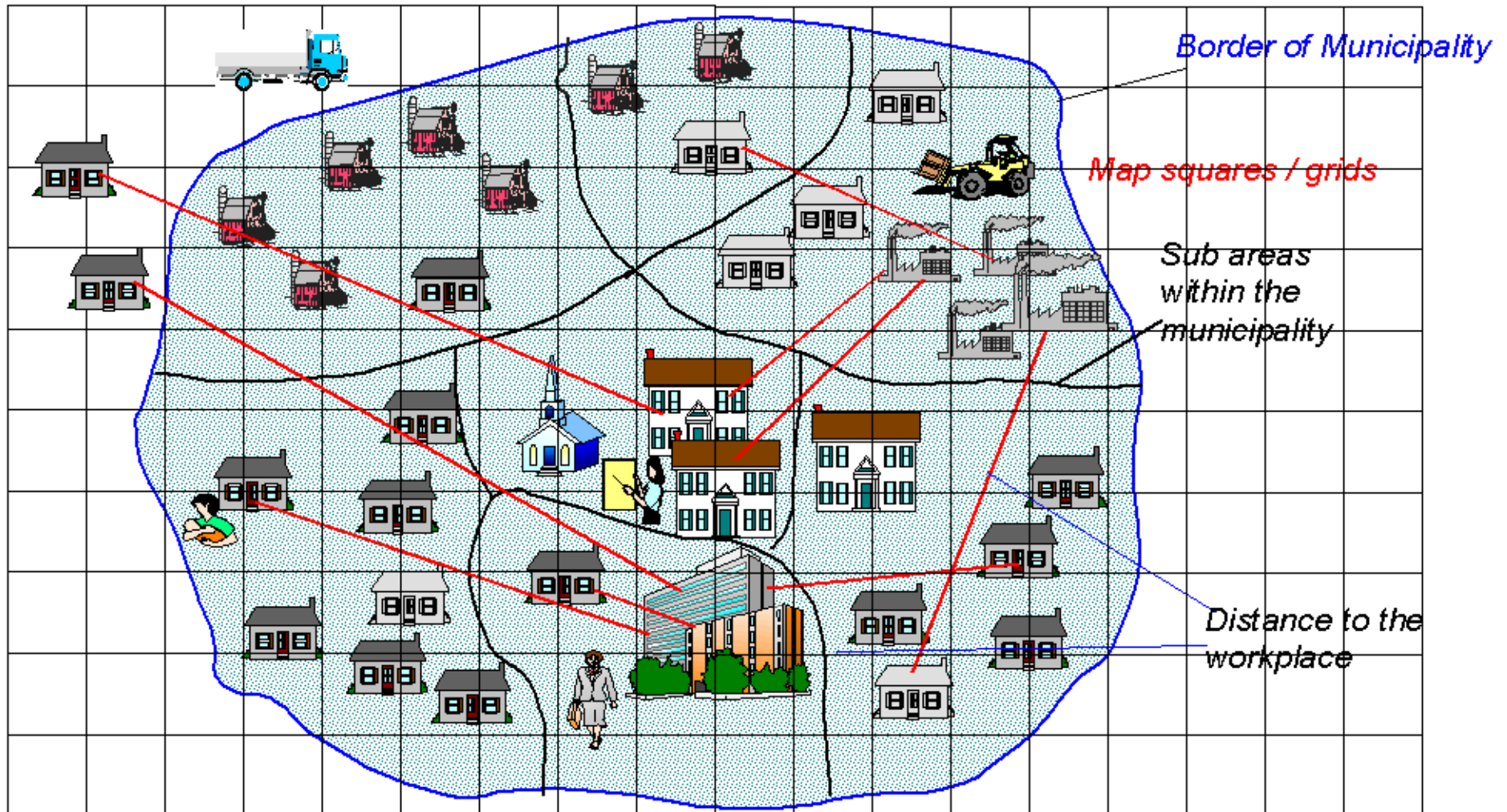
GIS and register-based census

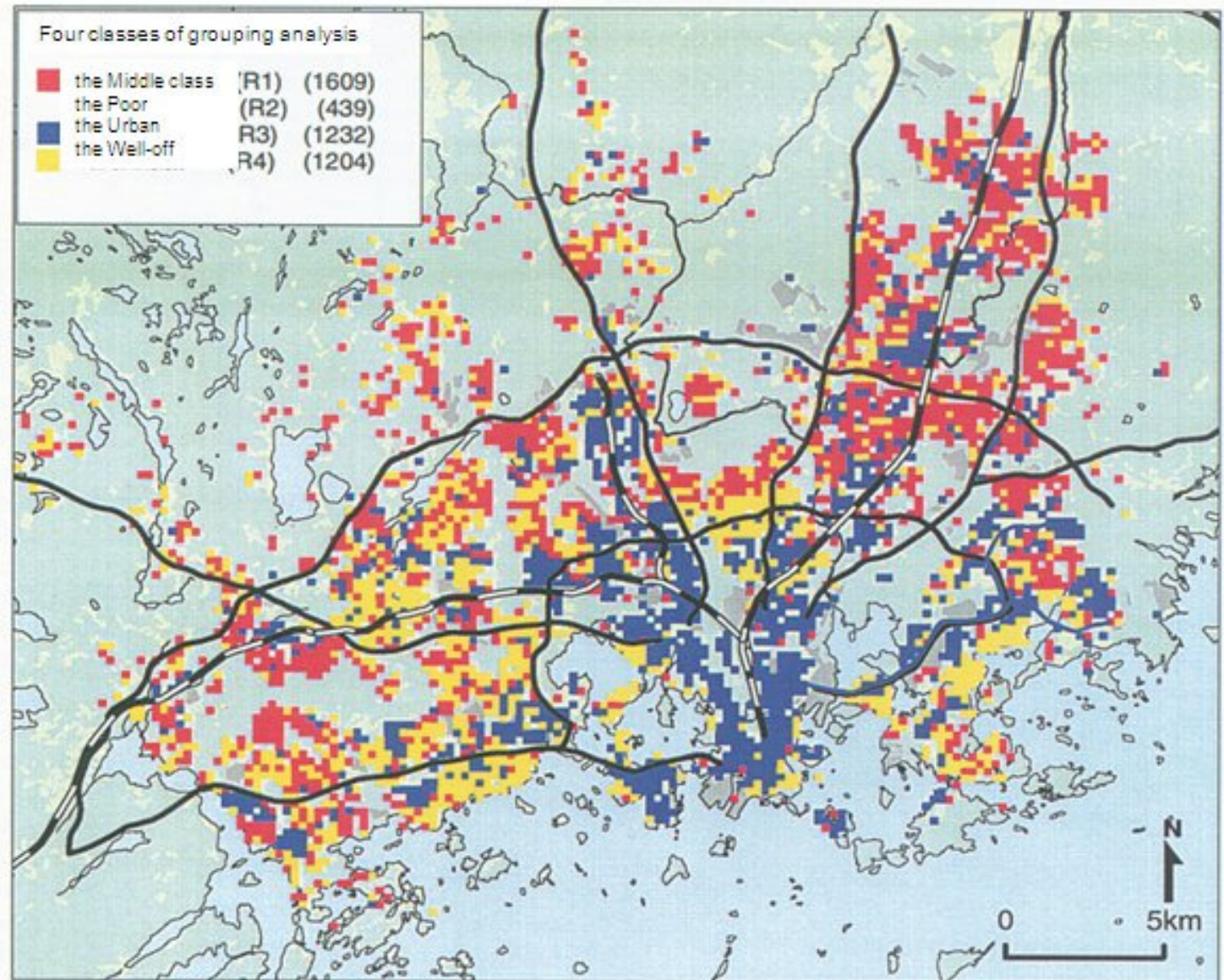
- The most commonly used non-administrative areas in statistics production are as follows:
 - a statistical classification between urban, semi-urban and rural areas
 - localities (urban settlements)
 - municipal sub-areas
 - post code areas
 - 1 km x 1 km grid squares.

The links between units

- person -> dwelling -> building -> **map co-ordinate**
- employed person -> enterprise (employer) -> establishment -> building -> **map co-ordinate**

Annual Statistics system for Small Areas in Finland





Data: Statistics Finland

Data warehouse

- Database in MS SQL Server environment
- Data from new production system (2005 -)
- Data from previous years (1970- 2004) transferred to database
- Database is used in dissemination processes
 - Tabulations (SuperSTAR databases, SAS, SQL)
 - Microdata (Samples for research purposes)
 - GIS analysis (ArcInfo) in GIS database (situated in same server)

Data warehouse - Contents

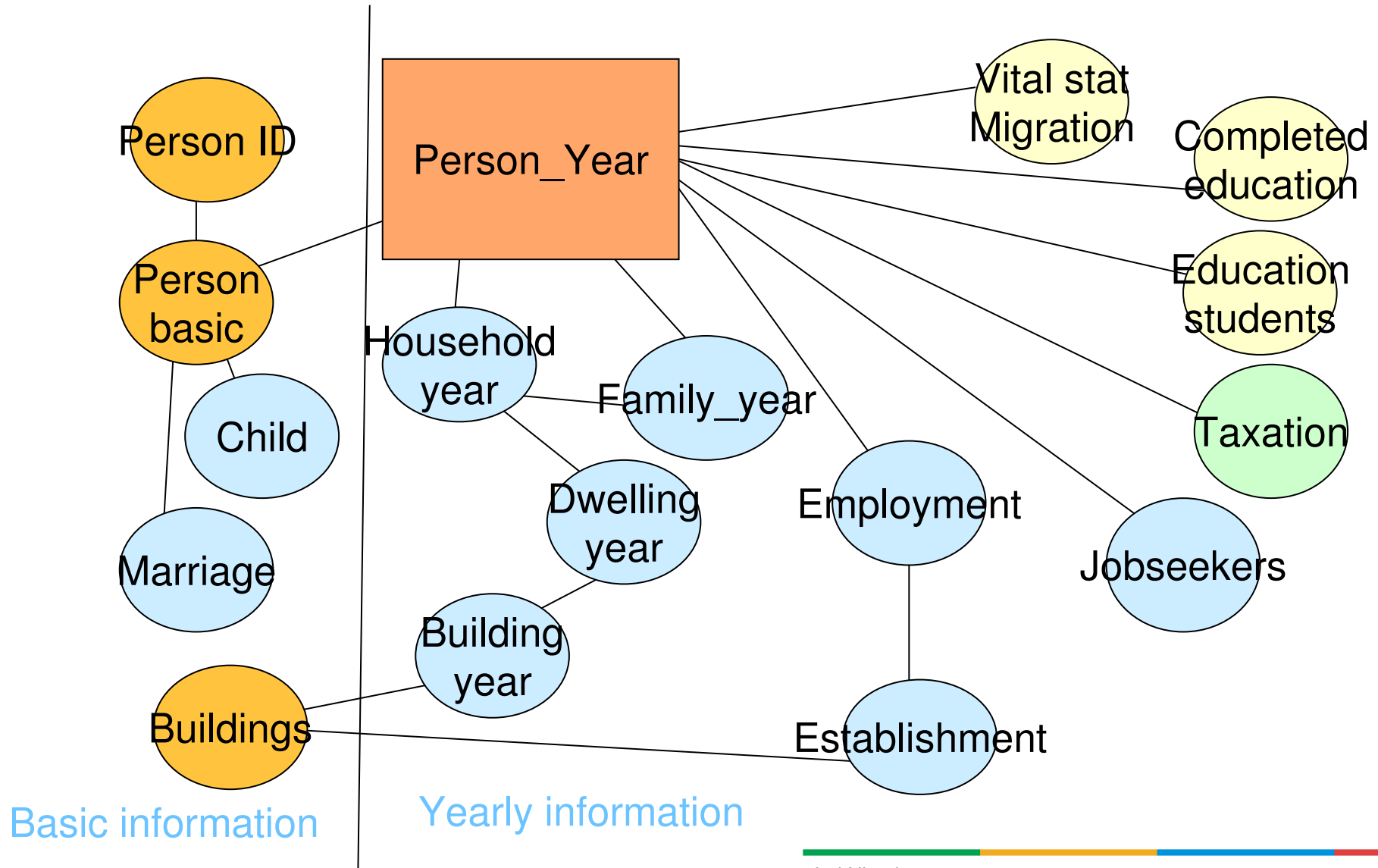
- Register based statistics in population statistics unit
- SQL Server database
- about 40 tables, almost 1000 columns/(variables)
- Population censuses
 - Population structure, Families, Dwellings and housing conditions, Buildings and free-time residences, Buildings and free-time residences, Employment
- Areal information (Subregions of buildings, localities, grid data, postal code area) - GIS integration
- Demography, Migration
 - Births, Changes in marital status, Deaths, Migration
- Education
 - Register of Completed Education and Degrees
 - Students and qualifications of educational institutions
 - Providers of education and educational institutions

Data warehouse - Contents

- [Taxable incomes](#) and [Total statistics on income distribution](#)
- Links (views) to [Classification database](#)
- Linked to Business statistics (business id code, GIS integration) [Finnish enterprises](#)



Data warehouse - Structure



Data warehouse - Conclusions

- In Statistics Finland we have now annual 'census' data 23 years period (1987-2010) and data from 'old' censuses (1970 -1985)
- Solution to flexible use of this data was building up a census data warehouse
- Data warehouse essential part of production processes used for all dissemination (tabulations, statistical database and microdata)
- New statistics (some examples later)
- Data mining
- Valuable source of data for other statistics in Statistics Finland

New possibilities for research and statistics production (1)

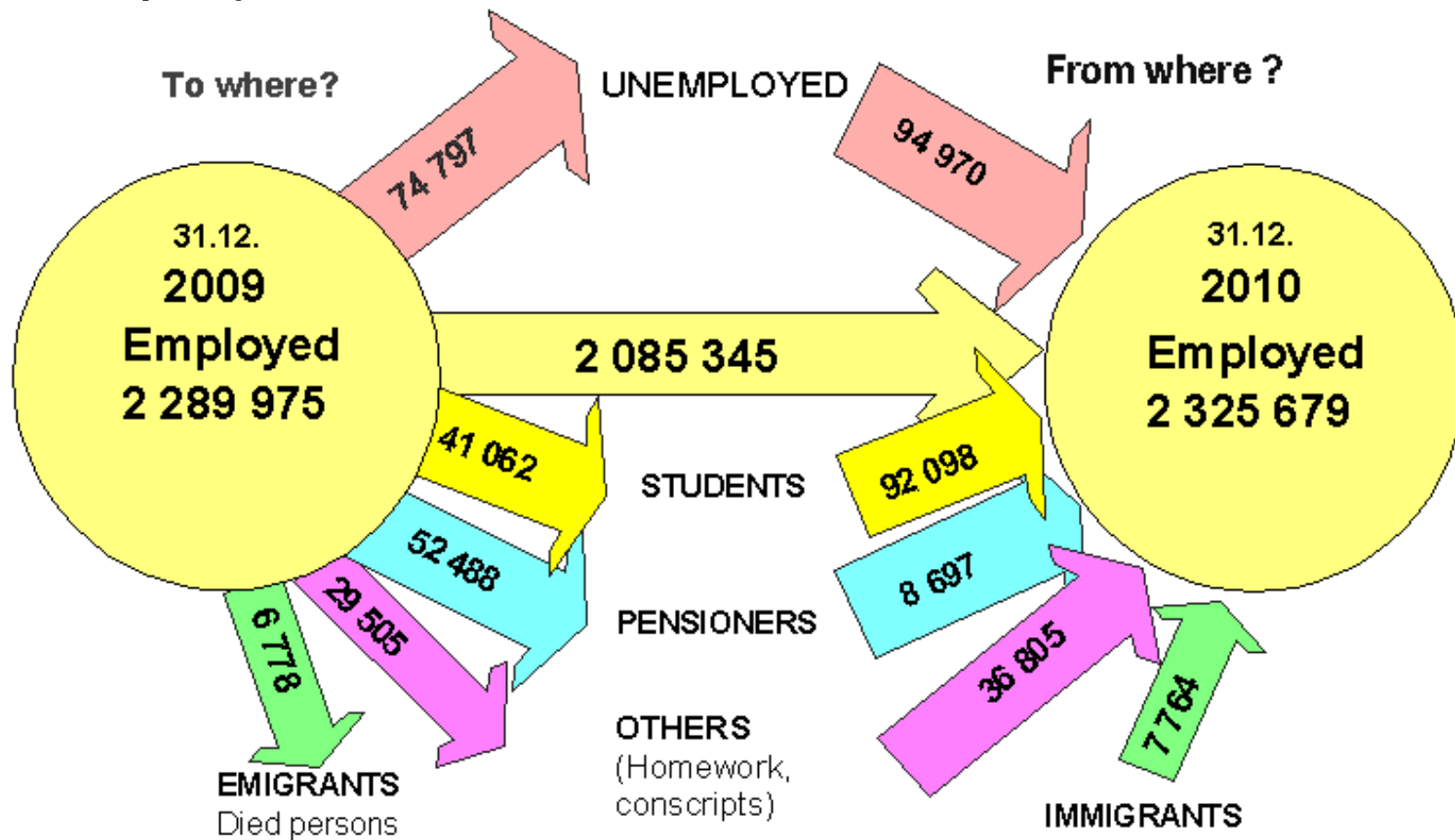
- Statistics usually provide cross-sectional information on a variable at a given point in time, such as population number or the number of people in gainful employment;
- on this basis we can see to what extent these figures have changed.
- The register system offers the added advantage of allowing us to identify the individuals behind these changes: who has got a job, who has completed a degree.
- Changes can be monitored by linking unit data from consecutive years

New possibilities for research and statistics production (2)

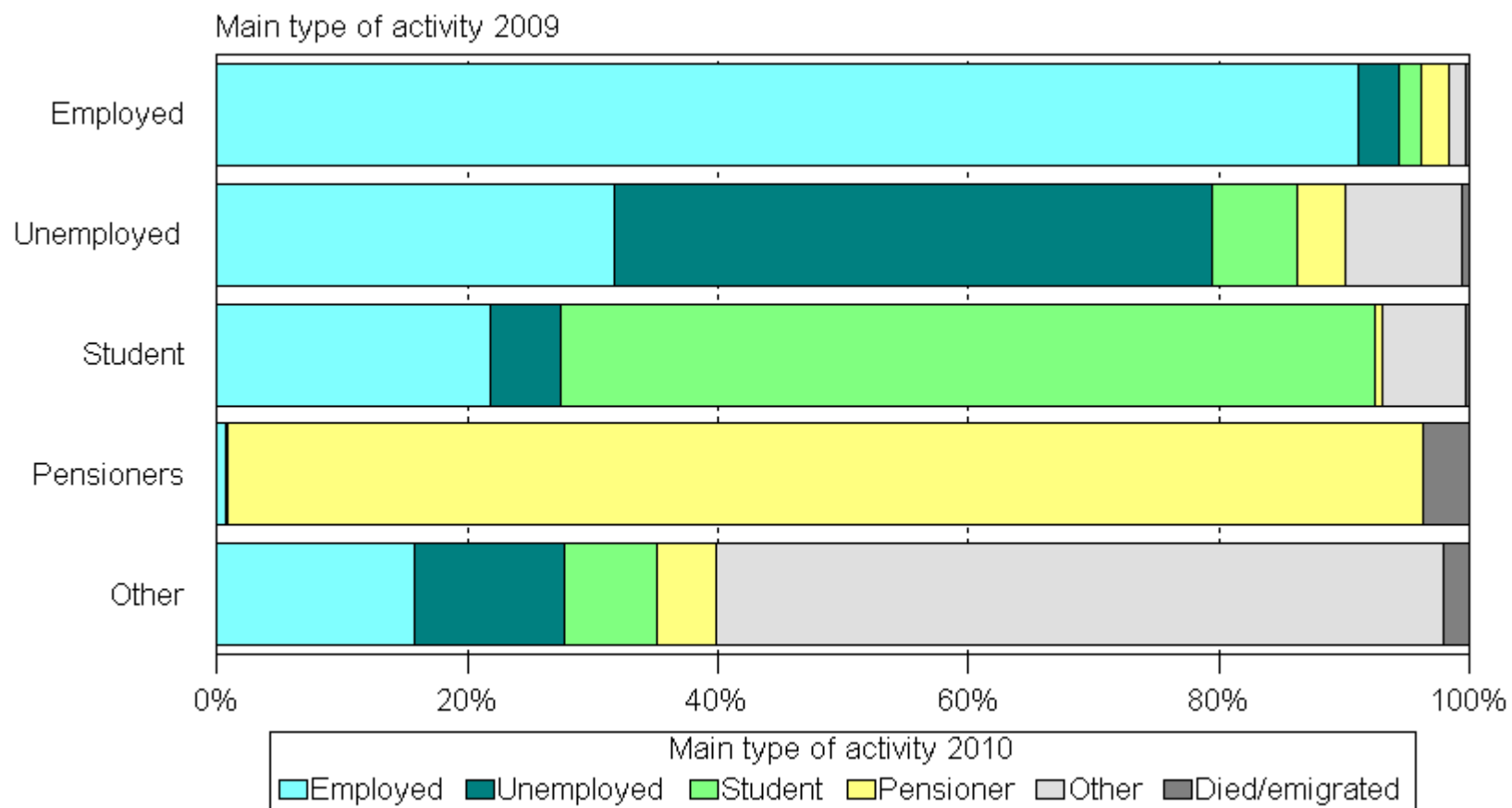
For instance:

- Flow statistics
 - employment flows
 - student flows
 - flows between branches of industry etc.
- Placement statistics
- Longitudinal researches

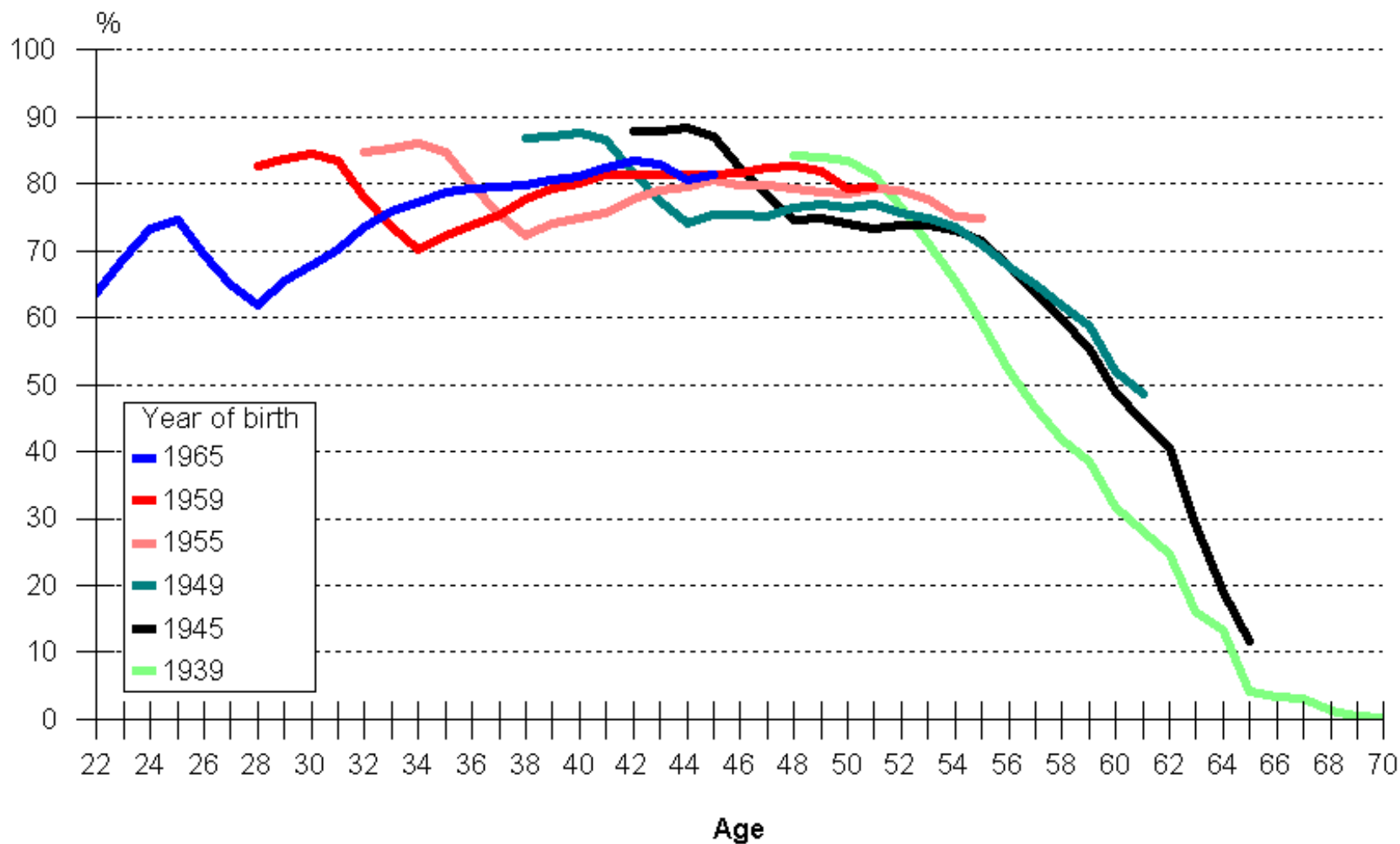
Flows between different activity groups: Employed 2009-2010



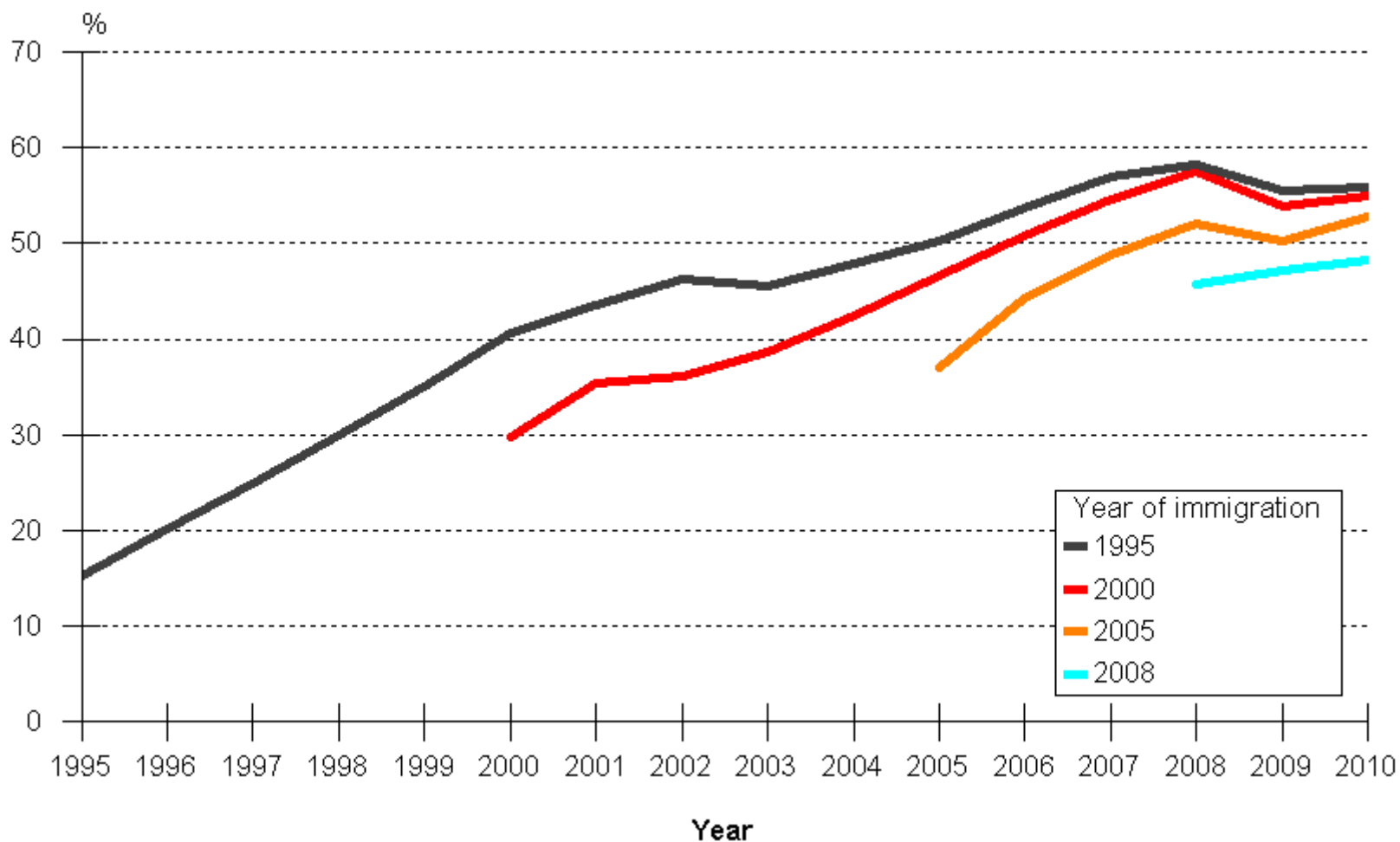
Changes in the population of age 15 years or more in 2010



Employment rate in different birth cohorts

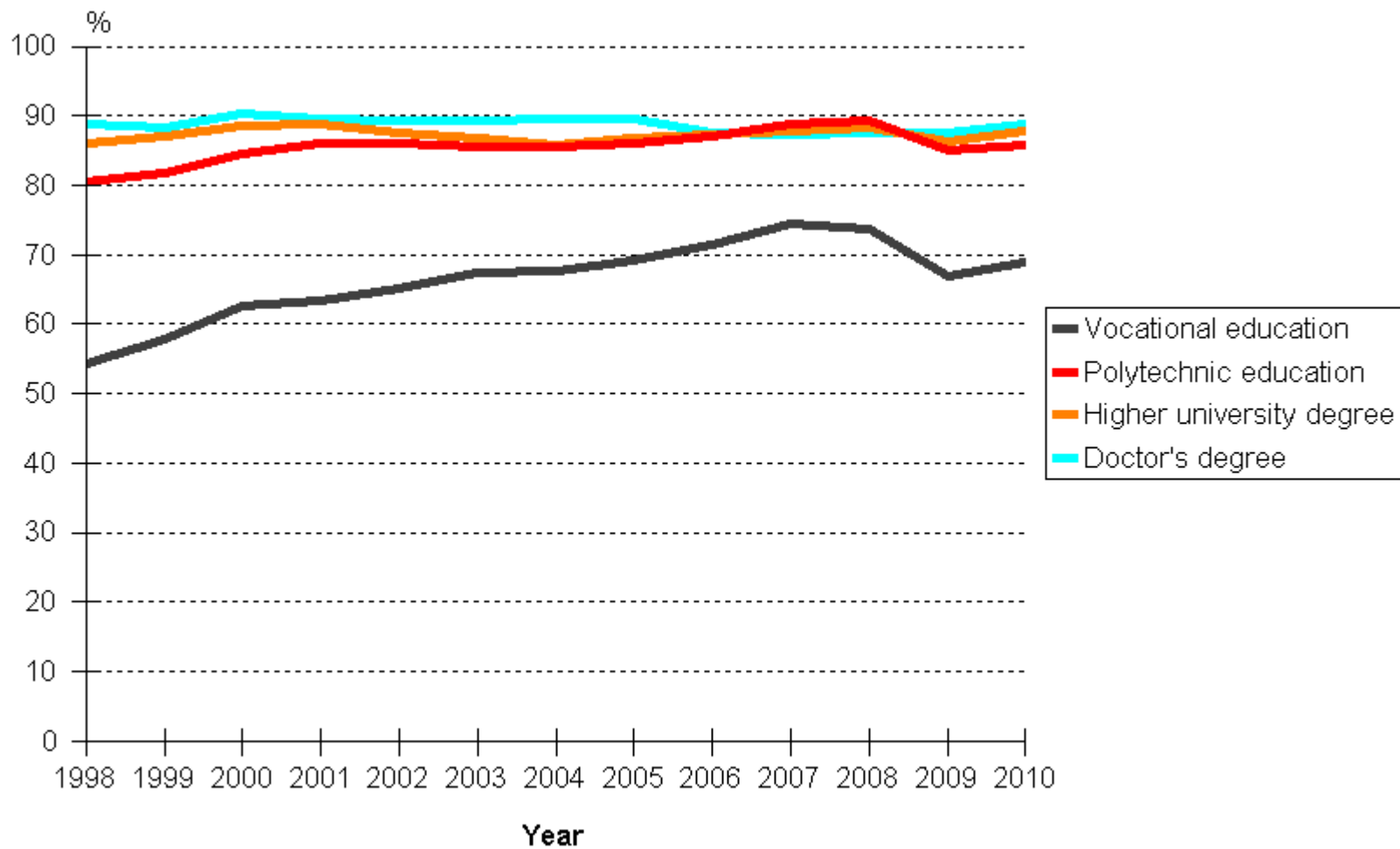


Employment rate of foreign immigrants in different immigration cohorts



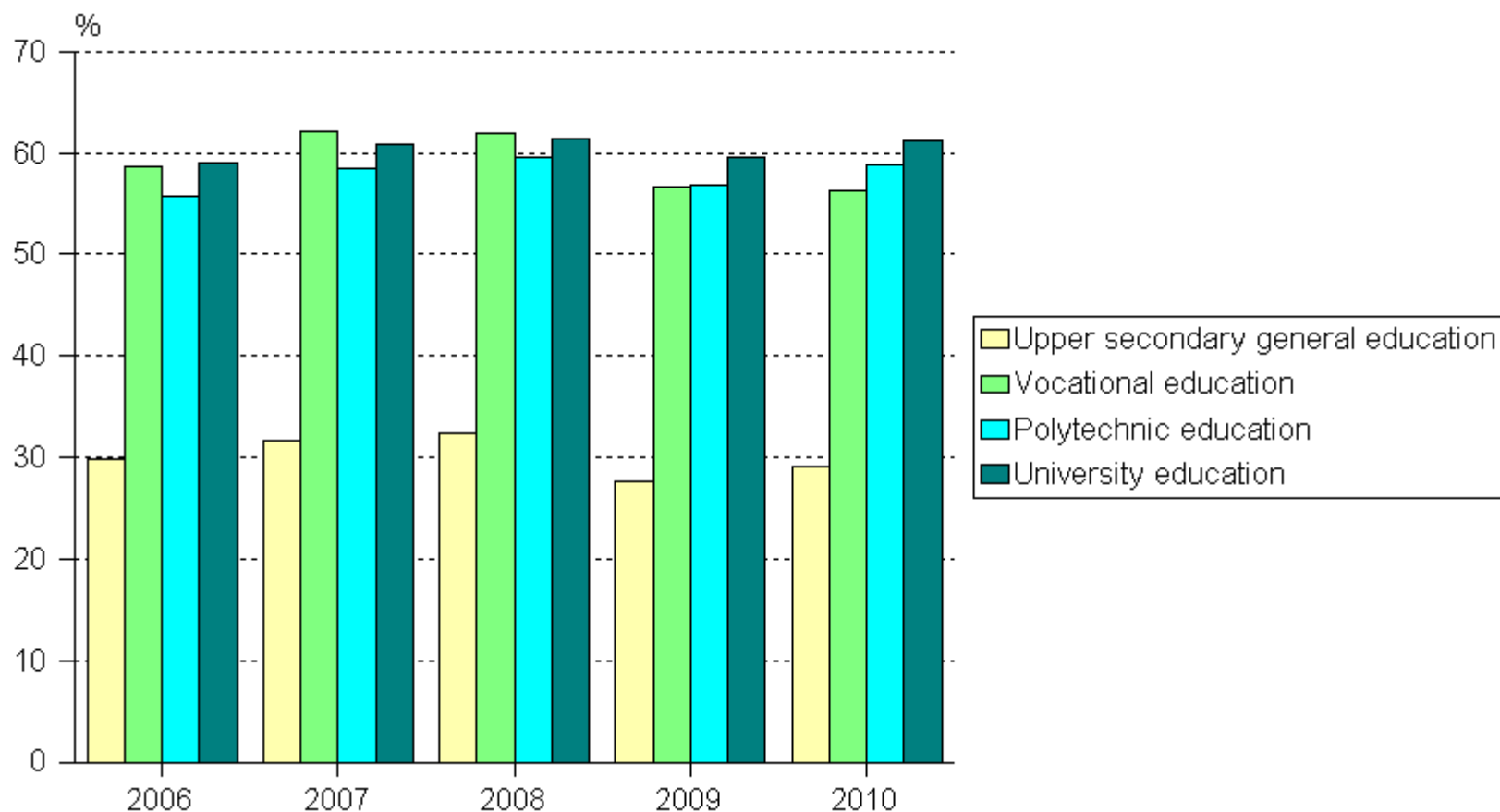


Employment of graduates one year after graduation 1998–2010, %





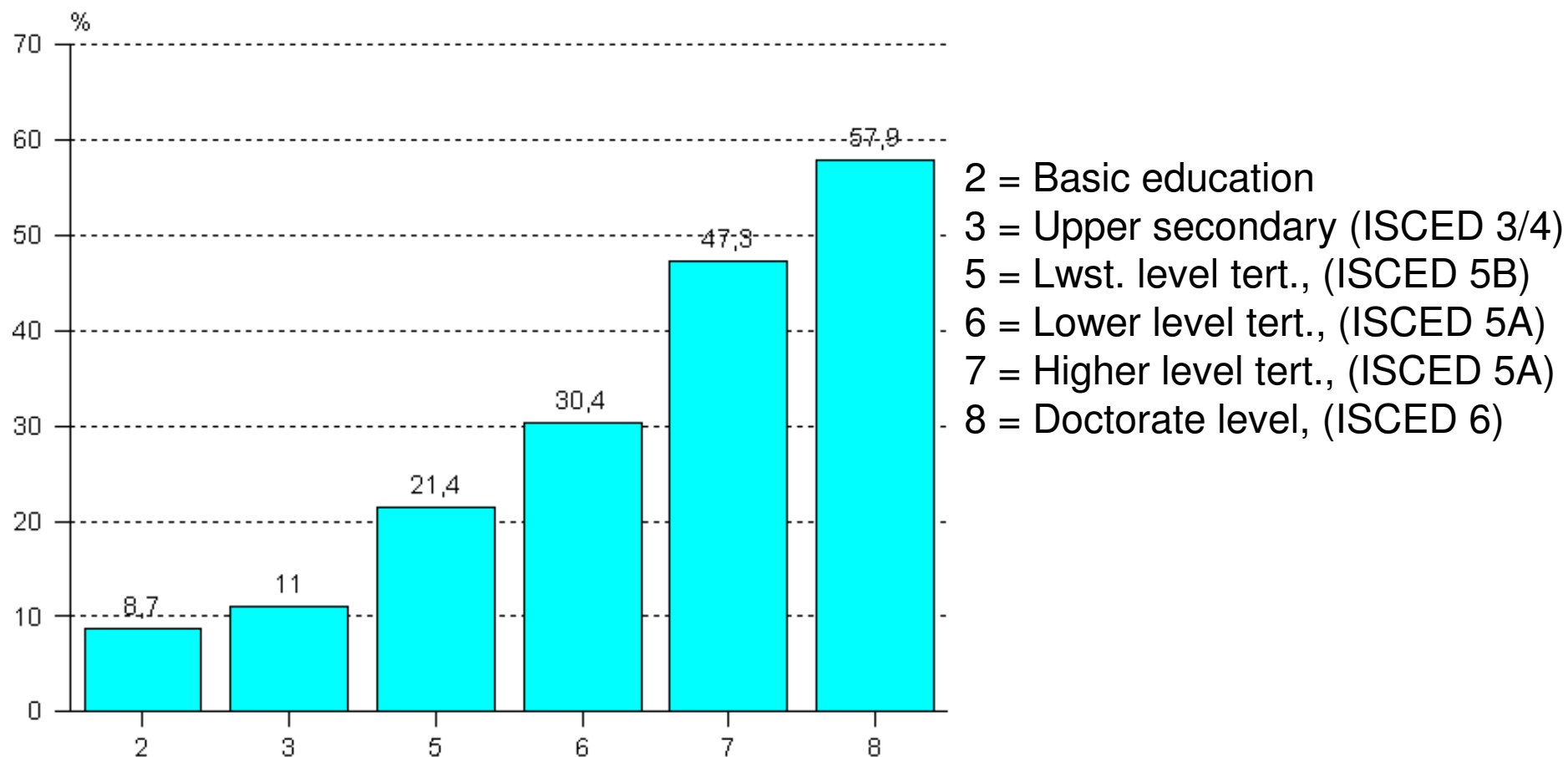
Shares of employed students aged at least 18 of all students in 2006-2010



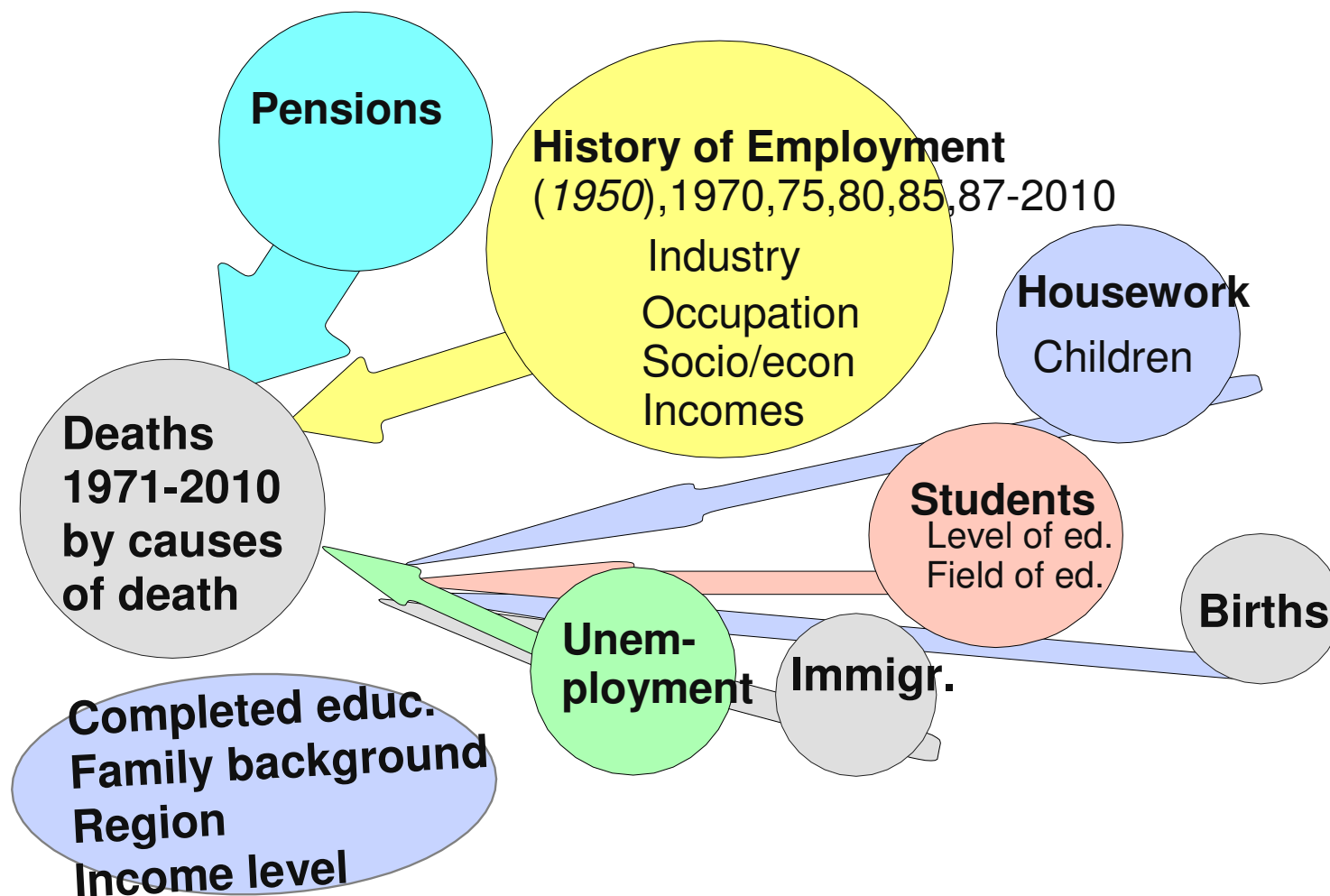


Statistics Finland

Proportion of persons 20-24 years in university studies 2010 according to the level of education of father



Deaths and the history of different activities



Related to the topic

Use of Registers and Administrative Data Sources for Statistical Purposes

Best Practices of Statistics Finland, Statistics Finland, 2004

Can be downloaded in pdf-format at:

<http://tilastokeskus.fi/censusbyregisters>

Register-based statistics in the Nordic countries

Review of the best practices with the focus on population and social statistics

United Nations Economic Commission for Europe, 2007

Can be downloaded in pdf-format at:

<http://www.unece.org/stats.pub.htm>