

README – Swedish Contextual Database

The Swedish Contextual Database is part of the Generations and Gender Programme (GGP), an international social science infrastructure for family dynamics and relationships (see: <http://www.ggp-i.org>). This research programme provides comparative, longitudinal individual-level survey data (GGS) for a large number of European and some non-European countries to which a contextual database of macro-level indicators is matched. The Swedish Contextual Database (Swedish CDB) accompanies the Swedish Generations and Gender Survey (GGS).

The construction of the Swedish CDB and the data collection followed a template developed within the GGP. The Swedish CDB offers a rich and unique set of time-series indicators at the national and regional level. Many of the indicators were calculated using Swedish Register Data (see the description below). These indicators are not available in publicly accessible statistics and the Swedish CDB is thus currently the only database to provide them.

The Swedish CDB is open access and can be downloaded from the Stockholm University Demography Unit, SUDA website: www.suda.su.se and from the website of the GGP: www.ggp-i.org. At the SUDA website, the data are available in xls format. At the GGP website, the Swedish CDB is also available in xls format in the Contextual Data Collection of the GGP. In addition to that, the internationally comparative data will be integrated into the Contextual Database (CDB) of the GGP. These data can be exported in other formats, as well (e.g. CSV, XML). The indicators can also be accessed in a single file in STATA or SPSS format. The data can be matched with the Swedish GGS. Other regional coding schemes are also supported, such as NUTS, OECD. The Swedish CDB can thus be used in research with other surveys, e.g., the European Social Survey (ESS) or the Survey on Health, Ageing and Retirement (SHARE). The CDB-data are also a unique source for macro-level studies (See: Caporali et al. 2016 for a detailed description of the CDB and its potentials).

The establishment of the Swedish Contextual Database was supported by a grant from the Riksbankens Jubileumsfond, RJ Infrastructure for Research, Grant IN15-0511:1.

If you use data from the Swedish CDB, we kindly request the following acknowledgement: *Data from the Swedish Contextual Database (SCD) were provided by Stockholm University Demography Unit (SUDA) in connection with the Generations and Gender Programme (GGP). The collection of the data was supported by a grant from the Riksbankens Jubileumsfond, RJ Infrastructure for Research, Grant IN15-0511:1.*

Contact person for the Swedish Contextual Database: Johan Dahlberg and Gerda Neyer (SUDA)

Johan.Dahlberg@sociology.su.se, Gerda.Neyer@sociology.su.se

Contact person for the GGP-Contextual Database at INED:

Arianna Caporali (INED), Arianna.Caporali@ined.fr

SWEDISH REGISTER DATA

Most of the statistical items in the Swedish Contextual Database were taken from Swedish register data. Only in a few cases were the indicators calculated from data from surveys. Additionally, data were retrieved from internationally available sources. All sources are provided as meta-information in the data. The following description concerns only the use of Swedish register data.

Three sources of register data have been used – Statistics Sweden's (SCB) online Statistics database, The Swedish Medical Birth Register (MBR) provided by The National Board of Health and Welfare (Socialstyrelsen), and a collection of registers from Statistics Sweden organized and accessed at Stockholm University.

Statistics Sweden (SCB) online Statistical database (SCB 2016) provides access to most of the **Official Statistics of Sweden**. The Official Statistics of Sweden comprise statistical products and statistics that the authorities of Sweden have decided should comprise the official statistics. The system for the official statistics includes the statistics, metadata, the production systems, final observation registers, publications, separate tables and databases.

Stockholm University organizes the **STAR** data collection (Sweden in Time: Activities and Relations), stored by Statistics Sweden, and accessed through remote desktop. This data collection contains information on all Swedes born in 1932 or later and who have been registered as living in Sweden any year after 1960. STAR data are based on registers such as the total population register (RTB; see below), income and taxations registers, the multigenerational register, the cause of death register, educational registers, conscription register, etc. The STAR data also include a collection called LISA (see below). The STAR data have especially been used to calculate regional-level statistics (NUTS2). The Stockholm University Demography Unit (SUDA), responsible for the construction of the Swedish contextual database had access to register data with information up to 2012, which is our last year of observation.

Longitudinal integration database for health insurance and labour market studies (LISA)

LISA presently holds annual registers since 1990 and includes all individuals 16 years of age and older that were registered in Sweden as of December 31 for any calendar year. Since 2010, data for individuals 15 years of age is available in the database. Statistics Sweden, the Social Insurance Agency, and the Swedish Agency for Innovative Systems have participated in the construction. LISA integrates existing data from the labor market, educational and social sectors and is updated each year with a new annual register. The individual is the primary object in LISA, but connections to family, companies and places of employment are also available (SCB 2017). When creating the Swedish

Contextual Database information taken from LISA was vital to create longitudinal statistics at the regional level. Especially regarding:

- Gainful employment as of November each year (according to register-based labour market statistics),
- Income from gainful employment (distinguished as largest, second largest, and third largest),
- Sector, economic and professional (employed or entrepreneur) for largest, second largest, and third largest gainful employment,
- Work for gainful employment in each November,
- Incomes or compensations from employment, entrepreneurial activities, studies, national military service, illness, parental leave, unemployment, labour market activity, rehabilitation, partial retirement, early retirement, retirement, occupational pension, annuities, social assistance, private pensions, etc.
- Disposable income,
- Place of residence (county, municipality, parish and property).

The Total Population Register (RTB) contains data that is essential for information about Sweden's population. The Total Population Register was created in 1968 and contains information about the population and its changes, and to a large extent reflects the content of the population register of the Swedish Tax Agency. The Total Population Register is available for every year from 1968 onwards (e.g., Ludvigsson et al. 2016).

The Swedish Medical Birth Register (MBR), which is maintained by the Swedish National Board of Health and Welfare, was established in 1973 and contains information on all live births from the 22nd gestational week, and stillbirths. The MBR is an exceptionally rich population-based data source. The numbers of missing values are very low. The MBR is compared every year to the output of infants reported to the Birth Register at Statistics Sweden, and missing cases that are discovered are included by obtaining copies of medical records from the hospitals involved (Källén and Källén 2014). In many cases, MBR has been the only source of information. When possible, MBR has been used, before other registers, as a source of fertility measurements. This because the accuracy is slightly higher in the MBR than in other registers. In a few cases, the STAR data do not contain information on births if the newborn did not survive to the end of the calendar year. This does not really cause any substantial discrepancy. However, when possible, we have chosen the source with the slightly higher accuracy.

Aggregated data

Regional subdivision is made in accordance with the NUTS system (Nomenclature des Unités Territoriales Statistics). Regional data in the Swedish Contextual Database are reported at the NUTS2 level. For Sweden, NUTS2 consists of 8 regions and NUTS 3 of 21 administrative counties (Län) (SCB 2008). In some cases publicly available data are only offered on the NUTS3 level (county). When possible, such items were recalculated and aggregated by using information on NUTS 3 and weighted by relevant population size.

For example: Item 108b (mean age at birth – reg) was calculated using information on mean age at birth and number of births at NUTS3 level. For item 108b the MBR only provides information about the mean age at birth at the NUTS 3 level. NUTS2 region “East Middle Sweden” consists of NUTS3 regions “Uppsala County”, “Södermanland County”, “Östergötland County”, “Örebro County”, and “Västmanland County”. Thus, mean age at birth in NUTS2 region “East Middle Sweden” was calculated by using mean ages at birth and number of births in the five NUTS3 regions.

The calculations using information at NUTS 3 (county) and weighting by relevant population size at that level were made by Johan Dahlberg at The Stockholm University Demography Unit (SUDA).

References:

- Caporali, A., Klüsener, S., Neyer, G., Krapf, S., Grigorieva, O., and Kostova, D. (2016). The Contextual Database of the Generations and Gender Programme: Concept, Content, and Research Examples. *Demographic Research* 35(9): 229-252. <http://www.demographic-research.org/volumes/vol35/9/35-9.pdf> DOI: 10.4054/DemRes.2016.35.9
- Källén, B., & Källén, K. (2014). The Swedish Medical Birth Registry: a summary of content and quality. The Swedish Centre for Epidemiology. The National Board of Health and Welfare, Stockholm, Sweden.
- Ludvigsson, J. F., Almqvist, C., Bonamy, A. K. E., Ljung, R., Michaëlsson, K., Neovius, M., ... & Ye, W. (2016). Registers of the Swedish total population and their use in medical research. *European Journal of Epidemiology* 31(2), 125-136. doi: 10.1007/s10654-016-0117-y
- SCB (2008). *Karta över NUTS-indelningen i Sverige (fr.o.m. 2008-01-01)* [Map of NUTS Division in Sweden (as of 01/01/2008)]. Retrieved from https://www.scb.se/Grupp/Hitta_statistik/Regional%20statistik/Kartor/_Dokument/NUTS_1_2_3_20080101.pdf
- SCB (2016). *Official Statistics of Sweden – Annual Report 2016*. Retrieved from https://www.scb.se/contentassets/c1dc2a4307cb4431b6ef4becd184b569/ov9999_2016a01_br_x43br1701eng.pdf

SCB (2017). *Longitudinal integration database for health insurance and labour market studies (LISA)*. Retrieved from <https://www.scb.se/en/services/guidance-for-researchers-and-universities/vilka-mikrodata-finns/longitudinella-register/longitudinal-integration-database-for-health-insurance-and-labour-market-studies-lisa/>

SUDA (2017). *Register data*. Retrieved from <https://www.suda.su.se/research/demographic-data/register-data>