

**UNITED NATIONS STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS**

**EUROPEAN COMMISSION
STATISTICAL OFFICE OF THE
EUROPEAN COMMUNITIES (EUROSTAT)**

Joint UNECE/Eurostat work session on statistical data confidentiality
(Manchester, United Kingdom, 17-19 December 2007)

Topic (iii): Applications (including practical implementation of SDC methods, actual issues within NSIs and software)

THE AVAILABILITY OF DUTCH CENSUS MICRODATA

Invited Paper

Prepared by Eric Schulte Nordholt, Statistics Netherlands

THE AVAILABILITY OF DUTCH CENSUS MICRODATA¹

Summary: Data from many different sources were combined to produce the census 2001 tables for the Netherlands. Statistics Netherlands conducted a virtual census, using registers and surveys already available. The virtual census is cheaper, comparable to earlier Dutch censuses and more socially acceptable. In this paper the availability of Dutch census microdata from 1960, 1971 and 2001 is described in more detail. In the last section of this paper some background information is given about the applied Statistical Disclosure Control (SDC) methods. The aim is to release as much census information as possible. However, the privacy of individual respondents should be respected. Therefore, SDC techniques have been developed to protect sensitive information that can be attributed to individual respondents. SDC is thus relevant to be able to decide properly what kind of census tables and microdata can be released.

Keywords: census; micro datasets; micro linking; Statistical Disclosure Control (SDC)

1. Introduction

In 2003 data were combined to produce the Dutch 2001 census tables. In the Netherlands this was not done by interviewing inhabitants in a complete enumeration, but by using data that Statistics Netherlands already had available. This way, the Dutch tax payer got a much lower census bill. The costs for a traditional census would be about three hundred million euros, while the costs made now are ‘only’ about three million. The estimate includes the costs for all preparatory work such as developing a new methodology and accompanying software. The costs of the registers are not included, but the analyses of the results are. Registers are not kept up-to-date for censuses but for other purposes. Saving money on census costs is only possible in countries that have sufficient register information.

The 2001 census relates to forty extensive frequency tables. Twenty-eight are about the Netherlands as a whole, nine are at the regional level (NUTS 3) and three at municipal level (NUTS 5). The forty tables fall into a number of groups. Eight tables concern housing, two tables concern commuting and the other thirty tables are demographic tables, relating to occupation, level of education and economic activity. Additionally, demographic, housing and labour figures are compiled at sub-city district level for ten large cities that participate in Urban Audit II (Statistics

¹ The views expressed in this paper are those of the author and do not necessarily reflect the policies of Statistics Netherlands.

Netherlands, 2003). These ten large cities are Amsterdam, Rotterdam, The Hague, Utrecht, Eindhoven, Tilburg, Groningen, Enschede, Arnhem and Heerlen.

The virtual census in the Netherlands was off to a later start than in other countries where a traditional census was conducted. It did not make sense to really start the 2001 Census Project until all sources were available; some registers were available relatively late. Nevertheless, the Netherlands was quicker with the compilation of the forty census tables than most of the other countries that participated in the 2001 Census Round. In fact, the Netherlands was one of the first to send the complete set of forty tables to Eurostat, which co-ordinated the contributions of all European Union (EU) member states, accession countries and European Free Trade Association (EFTA) member states. The Netherlands had the advantage that the incoming census forms did not need to be checked and corrected. However, one must realise that for some variables only sample information is available, which implies that it was impossible to meet the level of detail required in some Dutch tables.

The reason why Statistics Netherlands has compiled the set of tables is a gentlemen's agreement. In 1991 the Census Act was rescinded, officially cancelling Statistics Netherlands obligation to hold a census once every ten years (Corbey, 1994). There was no European obligation to supply 2001 Census data, but it is almost inconceivable that the Netherlands would not compile census data for the international organisations just like all other European countries do. Eurostat has a co-ordinating role in collecting harmonised data on the EU and a duty to make international comparisons of the outcome.

It was the third time that the Netherlands conducted a virtual census. However, the Dutch data that have been compiled for 1981 and 1991 were of a much more limited character than the set of tables of the 2001 census. Moreover, they were largely based on a register count of the population in combination with the then existing surveys on the labour force and housing conditions.

In sections 2, 3 and 4 the method of compiling the 2001 Census is explained, the micro linkage aspects are illuminated and information is given about recent publicity about censuses in the Netherlands. In section 5 some more detailed information can be found how microdata of the 1960, 1971 and 2001 Censuses were released. The applied rules for these public use microdata files can be found in section 6. In that section also other methods that allow use of census data are discussed.

2. Method of compiling

The current virtual census relates to 2001. The backbone of the census is the central Population Register (PR), which is the combination of all municipal population registers. The Population Register (PR) contains demographic information on every inhabitant of the Netherlands (Prins, 2000).

A number of integrated surveys and registers were linked to the PR. For this linking process only exact matches were used based on the unique so-called Social security and Fiscal (SoFi) number. The integrated system is called the Social Statistical Database (SSD) system. It is developed originally to conduct virtual censuses, but now it is also used for many social statistics.

The Nordic countries (Denmark, Finland, Iceland, Norway and Sweden) have more variables available in registers than the Netherlands (United Nations, 2007). Moreover, some Nordic countries conduct a (limited) enumeration for variables missing in the registers. Most of the other countries are in a similar position as the Netherlands where some variables relevant for the census can be found in registers, while other variables are available on a sample basis only. That's why much interest exists in the Dutch approach to combine registers and surveys and to use modern statistical techniques and accompanying software to compile the tables.

To be able to estimate every table as accurately as possible, each estimate is based on the largest possible number of records. Tables that contain register variables only are counted from the registers. Tables that contain at least one variable from a survey are estimated from the largest possible combination of registers and surveys.

We guaranteed consistency among the tables by using the technique of repeated weighting. It generates a new set of weights for each estimated table and is based on the repeated application of the regression estimator. When using repeated weighting, the weights of the records in the microdata are adapted in such a way that a new table estimate is consistent with all earlier table estimates.

It is possible to use the technique of repeated weighting in other countries as well. However, first it should be possible to use registers for statistical purposes. In most countries, not all census variables can be derived from register information. Additional surveying then remains a necessity, but a consistent set of census tables can be produced using the technique of repeated weighting.

3. Micro linkage

Most of the present administrative registers are provided with a unique linkage key. This SoFi-number is a personal identifier for every (registered) Dutch inhabitant and those abroad who receive an income from the Netherlands and have to pay tax over it to the Dutch fiscal authorities.

To prevent misuse of the SoFi-number, Statistics Netherlands recodes it for statistical processing into a so-called Record Identification Number (RIN-person). Personal identifiers, such as date of birth and address, are replaced by age at the reference date and RIN-address. This is all done in accordance with regulations of the Dutch Data Protection Authority to protect the privacy of the citizens.

All social statistics data files can be linked exactly to the PR. In practice this means that these data files are all indirectly linked to each other via the PR. Therefore the

PR can be considered the backbone in the set of social data sources. When linking the PR and the jobs register, or the PR and a register of social benefits, it is a linkage between different statistical units (persons, jobs, benefits). In that case multiple linkage relationships can exist because someone can have more than one job or can benefit from several social benefits.

In household sample surveys, like the Labour Force Survey (LFS), records do not have a SoFi-number. For those surveys an alternative linkage key is used, which is often built up by a combination of the following personal identifiers:

- sex;
- date of birth;
- address².

This sort of linkage key will usually be successful in distinguishing people. However, it is not a 100 percent unique combination of identifiers. Linking may result in a mismatch in the case of twins of the same sex. False matches may also occur when part of the date of birth or the postal code and house number is unknown or wrong. Another drawback is that the linkage key is not person but address related, which may cause linkage problems if someone has recently moved. When linking the PR and the LFS with this alternative key, and tolerating a variation between sources in a maximum of one of the variables sex, year of birth, month of birth or day of birth, the result is that close to 100 percent of the LFS records will be linked.

In its linkage strategy, Statistics Netherlands tries to maximize the number of matches and to minimize the number of mismatches. So, in order to achieve a higher linkage rate, more efforts are made to link the remaining unlinked records by means of different variants of the linkage key. For example, leaving out the house number and tolerating variations in the numeric characters of the postal code. To keep the probability of a mismatch as small as possible, some 'safety' devices are built in the linkage process. This last linking attempt accomplishes an extra one percent matches.

4. Publicity about censuses in the Netherlands

At the end of 2003 the complete set of forty census tables for the Netherlands was sent to Eurostat. The book 'The Dutch Virtual Census of 2001, Analysis and Methodology' was written afterwards (Schulte Nordholt et al., 2004). This book provides a wide-ranging description of the socio-demographic and socio-economic state of the Netherlands based on the 2001 census results. It discusses differences in size and composition among households, economic activity of households, individual activity status by region, age, education level and branch of economic

² In fact, the combination of a postal code (mostly related to the street) and house number is used as substitute for the address. The postal code in the Netherlands consists of four figures, followed by two letters.

activity. There are separate chapters on the economic activities of young people and people of retirement age. The economic activities, levels of education and occupation of foreigners from various countries of origin are compared with each other and with the native Dutch population. Regional aspects are also examined, including commuting. The results of the 2001 census are compared with the census results of some other European countries and with earlier Dutch censuses. Lastly, the virtual census methodology used is described in some detail.

The PDF version of the book can be found at the Statistics Netherlands website, at page <http://www.cbs.nl/en-GB/menu/themas/dossiers/volkstellingen/publicaties/default.htm>. An extra Chapter (number 15) is available at page <http://www.cbs.nl/en-GB/menu/themas/dossiers/volkstellingen/methoden/default.htm> with an overview of the used data sources, methods and definitions. Hard copies of the book were sent to all authors of the book, to the management of Statistics Netherlands and to several libraries. The book was also offered to the Prime Minister, the Minister of Economic Affairs and the Minister of Education, Cultural Affairs and Science of the Netherlands and to Director-Generals of statistical offices in several countries. In August 2004, the book was publicly released at an official presentation in the Statistics Netherlands' office in Voorburg. The research process and the main findings were then presented to an audience of academics, press representatives, government officials, as well as Statistics Netherlands' employees. Several articles were written in national and regional newspapers about the Dutch virtual census of 2001 and its results. Announcements, book reviews of Schulte Nordholt et al. (2004) and interviews appeared in several journals, mailing lists and newsletters. The methodology and key results of the virtual census of 2001 were also published as Schulte Nordholt (2005).

The set of forty standard tables for the Netherlands (in Excel format) can be found at page <http://www.cbs.nl/en-GB/menu/themas/dossiers/volkstellingen/publicaties/artikelen/archief/2005/default.htm>

5. Producing microdata

5.1 Introduction

Protected 1 percent samples of the microdata of the Dutch censuses of 1960, 1971 and 2001 were in 2005 disseminated via the IPUMS (Integrated Public Use Microdata Series) project, see <http://www.ipums.org/international>. These micro datasets contain a number of demographic and economic variables and can also be analysed via the institute DANS (Data Archiving and Networked Services), see <http://www.dans.knaw.nl/en/>. Bona fide researchers who want to make more detailed studies on these three censuses can work on-site at the premises of Statistics Netherlands. More information about this last option can be obtained via Statistics

Netherlands' Centre for Policy Related Statistics (<http://www.cbs.nl/nl-NL/menu/informatie/beleid/centrum-voor-beleidsstatistiek/diensten/default.htm>).

The Dutch censuses of 1960, 1971 and 2001 have been selected to be part of the IPUMS project. The censuses of 1960 and 1971 are traditional censuses, of which most of the micro data records have been recovered. The 2001 census is a virtual census, which means that it is composed of available register data and existing surveys. Unfortunately, this results in not having all variables available for all individual records. As a consequence we have not released the complete set of micro data, but an anonymised balanced sample of the individual personal records for which we have all demographic and economic variables. The sample fraction is a little bit over 1 percent of the total population.

The first stage in our cooperation in the IPUMS project has been the release of the 2001 census micro data. The selection of variables of the 2001 census has been leading in the selection of the variables of the censuses of 1960 and 1971. Due to differences in variable definitions, classifications and variable availability over time, differences among the three micro data sets remain. Also for 1960 and 1971 we release anonymised balanced 1 percent samples of the total population.

More information about the variables selected of the three censuses can be found in the next subsections. Links to more information about the 2001 census can be found in the previous section. For some more background and documentation of the 1960 and 1971 censuses we refer to the following web site: <http://www.volkstellingen.nl/en/documentatie/>.

5.2 The variable selection of the 2001 census

5.2.1 The sample

The sample is composed in three stages.

Of the persons 0-14 years a 1 percent sample, stratified to age (in years) and sex is drawn from the combined register data.

Out of the records from persons 15-74 years all complete records are selected. (Remember that we only had existing surveys and register data at our disposal). These records sum up to about 1 percent of the population in this age group. Record weights have been provided.

Of the records of persons 75 years or older all complete records are selected. Record weights have been provided.

Persons in institutional households are not included in the sample because they are not included in the surveys used for the virtual census of 2001.

5.2.2 The variables and their categories

Because of disclosure reasons we have to limit the detail. This results in the impossibility of providing regional detail. Moreover, some rare combinations of

identifying variables will lead to a limited number of suppressions of variable scores. The list of variables of the 2001 census sample is as follows.

1. *Sex*
2. *Age*
3. *Household position*
4. *Household size*
5. *Place of residence one year prior to the census*
6. *Country of citizenship*
7. *Country of birth*
8. *Level of educational attainment (ISCED level 4)*
9. *Economic status*
10. *Occupation (ISCO-COM 1 digit)*
11. *Branch of current economic activity (NACE, 1 letter)*
12. *Marital status*

5.3 The variable selection of the 1971 census

5.3.1 The sample

For the census year 1971 the gross sample of 1.25 % of the total population is randomly drawn, stratified to sex, 17 age groups (16 5-year groups and 80+) and 12 regions (11 provinces and 1 region consisting of newly made land (polders) and the centrally registered population). After removing incomplete and other problematic records a net sample of over 1 percent remained. The records have been weighted to the published census combined totals of sex times age in years, 11 provinces (the newly made land and the centrally registered travelling population are added to one of the provinces) times sex times age in 5-year classes and simple totals of most of the published variables.

The total population includes persons in institutional households, this in contrast to the IPUMS dataset of the virtual census of 2001.

5.3.2 The variables and their categories

Because of disclosure reasons we have to limit the detail. This results in the impossibility of providing regional detail. Moreover, some rare combinations of identifying variables will lead to a limited number of suppressions of variable scores. The list of variables of the 1971 census sample is as follows.

1. *Sex*
2. *Age*
3. *Country of citizenship*
4. *Marital status*
5. *Household position*
6. *Religious denomination*
7. *Country of birth*
8. *Household size*

- 9. *Economic status*
- 10. *Level of educational attainment*
- 11. *Occupation (ISCO 1 digit)*
- 12. *Branch of current economic activity (SBI 1970)*

5.4 The variable selection of the 1960 census

5.4.1 The sample

For the census year 1960 the gross sample of 1.25 percent of the total population has been drawn randomly, stratified to sex, 17 age groups (16 5-year groups and 80+) and 12 regions (11 provinces and 1 region consisting of newly made land (polders) and the centrally registered population). After removing incomplete and other problematic records a net sample of over 1 percent remained. The records have been weighted to the published census combined totals of sex times age in years and region times sex times age in 5-year classes.

The total population includes persons in institutional households, this in contrast to the virtual census of 2001.

The source material is the over 11 million original punch cards, which have been reread and digitised from 1973 onwards. A report in English on the reconstruction of the dataset is available.

5.4.2 The variables and their categories

Because of disclosure reasons we have to limit the detail. This results in the impossibility of providing regional detail. Moreover, some rare combinations of identifying variables will lead to a limited number of suppressions of variable scores. The list of variables of the 1960 census sample is as follows.

- 1. *Sex*
- 2. *Age*
- 3. *Marital status*
- 4. *Household position*
- 5. *Religious denomination*
- 6. *Country of birth*
- 7. *Economic status*
- 8. *Level of educational attainment*
- 9. *Occupation*
- 10. *Branch of current economic activity (SITC)*

6. Applied Statistical Disclosure Control methods

6.1 Introduction

The information from statistics becomes available for the public in tabular and microdata form. Historically, only tabular data were available and National Statistical Institutes (NSIs) had a monopoly on the microdata. Since the eighties the PC revolution led to the end of this monopoly. Now also other users of statistics have the possibility of using microdata. These microdata can be conveyed with floppies, CD-ROMs, USB sticks and other means. Recently also other possibilities of getting statistical information have become more popular as remote access and remote execution. With these techniques researchers can get access to data that remain in a statistical office or can execute set-ups without having the data on their own PC. For very sensitive information some NSIs have the possibility to let bona fide researchers work on-site within the premises of the NSI.

The task of statistical offices is to produce and publish statistical information about society. The data collected are ultimately released in a suitable form to policy makers, researchers and the general public for statistical purposes. The release of such information may have the undesirable effect that information on individual entities instead of on sufficiently large groups of individuals is disclosed. The question then arises how the information available can be modified in such a way that the data released can be considered statistically useful and do not jeopardize the privacy of the entities concerned. The Statistical Disclosure Control theory is used to solve the problem of how to publish and release as much detail in these data as possible without disclosing individual information (Willenborg and De Waal, 1996 and 2001).

6.2 The release of public use census microdata files

Many users of statistics are satisfied with the safe tables released by statistical offices. However, some users require more information. For many surveys microdata for researchers are released. In the case of census data it was preferred to release public use microdata files. Public use microdata files contain less detailed information than microdata for research. However, the audience for these files is much larger. The software package μ -ARGUS (Hundepool et al, 2007a) is of help in producing all kinds of protected microdata files. For the public use microdata files Statistics Netherlands uses the following set of rules:

1. The microdata must be at least one year old before they may be released.
2. Direct identifiers should not be released. Also direct regional variables, nationality, country of birth and ethnicity should not be released.
3. Only one kind of indirect regional variables (e.g. the size class of the place of residence) may be released. The combinations of values of the indirect regional

variables should be sufficiently scattered, i.e. each area that can be distinguished should contain at least 200 000 persons in the target population and, moreover, should consist of municipalities from at least six of the twelve current provinces in the Netherlands. The number of inhabitants of a municipality in an area that can be distinguished should be less than 50 % of the total number of inhabitants in that area.

4. The number of identifying variables in the microdata is at most 15.
5. Sensitive variables should not be released.
6. It should be impossible to derive additional identifying information from the sampling weights.
7. At least 200 000 persons in the population should score on each value of an identifying variable.
8. At least 1 000 persons in the population should score on each value of the crossing of two identifying variables.
9. For each household from which more than one person participated in the survey we demand that the total number of households that correspond to any particular combination of values of household variables is at least five in the microdata.
10. The records of the microdata should be released in random order.

According to this set of rules the public use microdata files are protected more severely than the microdata for research. For public use files it is not allowed to release direct regional variables. This is not considered to be a big problem for the census microdata files as the aim is to make international comparisons.

The software package μ -ARGUS is of help to identify and protect the unsafe combinations in the desired microdata file. Thus the rules 7 and 8 can be checked with μ -ARGUS. Global recoding and local suppression are two data protection techniques used to produce safe microdata files. In the case of global recoding several categories of an identifying variable are collapsed into a single one. This technique is applied to the entire data set, not only to the unsafe part of the set, so that a uniform categorisation of each identifying variable is obtained.

Both global recoding and local suppression lead to information loss, because either less detailed information is provided or some information is not given at all. A balance between global recoding and local suppression should always be found in order to make the information loss due to the Statistical Disclosure Control measures as low as possible. It is recommended to start by recoding some variables globally until the number of unsafe combinations that has to be protected is sufficiently low. Then the remaining unsafe combinations have to be protected by local suppressions.

6.3 Other methods that allow use of census data

All Statistical Disclosure Control techniques necessarily involve data manipulation or suppression and are likely to reduce the quality of estimates to be produced from the data. As a result, NSIs have begun to investigate other methods that allow use of data while protecting confidentiality of sensitive information given by respondents. These methods allow the data to be used in an environment controlled by the NSI and require that its use be subject to the same legal and ethical protections placed on the NSI itself.

Probably the most important access modality developed in the past decade is that of restricted access sites. These sites permit NSIs to respond to the microdata needs of researchers. Some researchers need namely more information than is available in the released public use census microdata files. As the releasing of richer data is not allowed, it is then possible for individual researchers to perform their research on richer microdata on the premises of the NSIs. Statistics Netherlands is one of the NSIs that has such a facility. Bona fide researchers have the opportunity to work on-site in a secure area within Statistics Netherlands. Researchers can choose at will between the two locations of Statistics Netherlands: Voorburg in the west of the Netherlands and Heerlen in the south of the Netherlands. The possibility to export any information is however only possible with the permission of the responsible statistical officer. They can apply standard statistical software packages and also bring their own programmes. Like all employees of Statistics Netherlands, these people who work on-site have to swear an oath to the effect that they will not disclose the individual information of respondents (Kooiman, Nobel and Willenborg, 1999).

Finally, an option is to allow remote access. This access modality combines the advantage that researchers can stay in their own institute and the advantage that the data stay in the NSI. Normally, researchers get access through an intermediary controlled by the NSI that guarantees that all use conforms to the law. One step further goes the option of remote execution. An intermediary is then no longer placed between the researcher and the NSI. With remote execution researchers can execute set-ups without having the data on their own PC. Although remote execution is a more efficient option than remote access the question is whether the security systems are strong enough to let this technique become an often used modality. Currently, Statistics Netherlands has a Centre for Policy Related Statistics that is running the on-site and remote execution and access facilities for researchers.

6.4 Discussion and conclusions

In this paper methods have been described that have been developed to protect confidentiality, while at the same time providing access to data, through various means that either alter the data or restrict access to them. The balance between data confidentiality and data access is a delicate one. Hopefully, the new research methods and software for Statistical Disclosure Control can help in keeping the right balance for the Dutch census microdata.

The software packages μ -ARGUS and τ -ARGUS have emerged from the Statistical Disclosure Control (SDC) project that was carried out under the Fourth Framework Programme of the European Union. The Computational Aspects of Statistical Confidentiality (CASC) project in the Fifth Framework Programme of the European Union can be seen as a follow-up of the SDC project.

New manuals for μ -ARGUS and τ -ARGUS become regularly available (Hundepool et al, 2007a and b). The ARGUS packages have moved towards interfaces with several state of the art engines produced by statisticians from many different countries. The most recent information is published at the CASC website: <http://neon.vb.cbs.nl/casc>.

References

- Corbey, P., 1994, "Exit the population Census", *Netherlands Official Statistics*, 9, summer 1994, pp. 41-44.
- Hundepool, A., A. van de Wetering, R. Ramaswamy, L. Franconi, S. Poletti, A. Capobianchi, P.P. de Wolf, J. Domingo, V. Torra, R. Brand and S. Giessing, 2007a, *μ -ARGUS, user's manual, version 4.1*, Voorburg, The Netherlands: Statistics Netherlands.
- Hundepool, A., A. van de Wetering, R. Ramaswamy, P.P. de Wolf, S. Giessing, M. Fischetti, J.J. Salazar, J. Castro and P. Lowthian, 2007b, *τ -ARGUS, user's manual, version 3.2*, Voorburg, The Netherlands: Statistics Netherlands.
- Kooiman, P., J.R. Nobel and L.C.R.J. Willenborg, 1999, "Statistical data protection at Statistics Netherlands", *Netherlands Official Statistics*, 14, pp. 21-25.
- Prins, C.J.M., 2000, "Dutch population statistics based on population register data", *Monthly Bulletin of Population Statistics*. Vol. 2000/02 (February 2000), pp. 9-15.
- Schulte Nordholt, E., 2005, "The Dutch virtual Census 2001: A new approach by combining different sources", *Statistical Journal of the United Nations Economic Commission for Europe*, 22, 2005, pp. 25-37.
- Schulte Nordholt, E., M. Hartgers and R. Gircour (Eds.), 2004, *The Dutch Virtual Census of 2001, Analysis and Methodology*, Statistics Netherlands, Voorburg / Heerlen, July, 2004. <http://www.cbs.nl/NR/rdonlyres/D1716A60-0D13-4281-BED6-3607514888AD/0/b572001.pdf>
- Statistics Netherlands, 2003, "Urban Audit II, the implementation in the Netherlands", *Report, BPA no. 2192-03-SAV/II*, Statistics Netherlands, Voorburg. <http://www.cbs.nl/nr/rdonlyres/8c6e4c9d-4338-4e32-848b-8d43b9b3242d/0/urbanauditiiinetherlands.pdf>
- United Nations, 2007, "Register-based statistics in the Nordic countries – Review of best practices with focus on population and social statistics". *Report*, United

Nations Economic Commission for Europe (UNECE) Statistical Division,
Geneva.

Willenborg, L.C.R.J. and T. de Waal, 1996, *Statistical Disclosure Control in practice, Lecture Notes in Statistics 111*, New York: Springer-Verlag.

Willenborg, L.C.R.J. and T. de Waal, 2001, *Elements of Statistical Disclosure Control, Lecture Notes in Statistics 155*, New York: Springer-Verlag.