



REPUBLIC OF SLOVENIA



STATISTICAL OFFICE OF THE REPUBLIC OF SLOVENIA

Basic instructions for researchers concerning the access and the use of statistically protected microdata

In order to provide researchers with an insight into, and facilitate their understanding of, the issues of statistically protected microdata, we would like to convey some information which should be of help in understanding these issues. We have also presented the samples of **contracts** covering work with microdata in a secured room, the **form** for obtaining the statistically protected microdata, and a **list** of institutions/ individuals who obtained the approval of the Statistical Office of the Republic of Slovenia in 2005 for performing the work on the statistically protected microdata prepared for their use.

Researchers can obtain selected European microdata sets by [Eurostat](#).

The right to obtain microdata

In Slovenia, the entities entitled to obtain the statistically protected microdata include the registered research institutions, registered researchers, and the researchers of government offices (by definition, the statistically protected microdata have been modified in a way which prevents the identification of the unit to which they refer, while containing the variables on whose basis it is possible to carry out statistical analysis in the same way as this can be done with individual data).

Upon conveying the statistically protected microdata to researchers, the Statistical Office of the Republic of Slovenia consistently complies with the legal, ethical and professional criteria. The right of researchers to obtain the statistically protected microdata is not absolute; however, the researchers can obtain the statistically protected microdata only subject to their capacity to guarantee the protection of these data.

Intended use

The purposes for which the statistically protected microdata can be obtained are statistical-analytical purposes, and the scientific research purposes, where the analysis involves a public benefit or something positive for the people of Slovenia.

Statistically protected microdata cannot be obtained for non-statistical purposes. These include the administrative purposes (where dissemination of microdata would have a direct administrative consequence for a citizen or a company), a tax purpose, and the judicial or business purposes.

Dissemination of microdata and the preservation of public trust in the Statistical Office of the Republic of Slovenia

When providing the statistically protected microdata we are aware that it is essential to preserve the confidence of the public in the Statistical Office of the Republic of Slovenia. Mistrust in the Statistical Office of the Republic of Slovenia would result in inaccurate reporting of data by the reporting units to the Statistical Office of the Republic of Slovenia, the consequences of which would be a reduction in

the quality of statistical data. Even one serious incident, in relation to the disclosure of private information of individuals or the business interests of a commercial entity, if accompanied also by strong media coverage, might seriously jeopardize the quality of national statistics.

Three main factors which may increase the risk of disclosure of private information of individuals, and business secrets of business units when providing microdata, are present in Slovenia due to its development level and size. The first one is due to the advanced level and easy access to computer technology, both software and hardware, enabling the combination of data and the analyses, which in the past was solely the domain of the research institutions, but today can be performed on PCs at home without any major financial input. Another factor is the existence of an ever larger number of databases. These have been emerging in Slovenia in an organized and fast way as a response to the need of organized state governance. It is prescribed that they should also contain single identifiers in the field of administration (a single personal ID number, and a single registration number of an entity). This can serve as a basis for linking data from various bases. The third factor is the size of the Slovenian economy, and some prominent large companies which are so typical and therefore so recognizable that their data is hard to protect statistically.

The rules of the Statistical Office of the Republic of Slovenia when providing statistically protected microdata (and sensitive tables)

In the Statistical Office of the Republic of Slovenia, dissemination of statistically protected microdata and sensitive tables (from the point of view of statistical confidentiality) to researchers is organized through the function of the advisory body of the Director General of the Data Protection Committee, in compliance with the system of rules and procedures related to the dissemination of statistically protected microdata to researchers, and the use of software for the statistical protection of data.

The system of rules and procedures concerning the dissemination of statistically protected microdata

1. PROVISION OF INDIVIDUAL DATA

1.1 Provision of individual data within the national statistics system

Good practice: it is possible to exchange individual data for the implementation of the Annual Programme of Statistical Surveys between SURS (Statistical Office of RS) and authorized contractors of the Annual Programme of Statistical Surveys according to the prescribed procedure.

1.2 Provision of individual data, obtained through combined questionnaires, to partner institutions

In certain cases, SURS and one of the of government institutions send a combined questionnaire to the reporting units, thus relieving the reporting units from double reporting of the same or similar data to two different government institutions.

Good practice: SURS collects data by means of a combined questionnaire with a partner institution from within the government's offices only when the SURS and the partner institution have a legal basis to do so, and when the interest of SURS is not put at risk by this mode of collection of data.

Such legal basis for the collection of data by SURS and partner institutions shall be indicated on the questionnaire, and the circular or internal memo should also state the purpose of collecting data for the partner institutions. Data is sent to the institution according to a pre-defined procedure.

1.3 Provision of individual data to observation units requesting their own data

In some cases, observation units ask the SURS to provide them with data which they had previously submitted to SURS.

Good practice: if this data is available to SURS, it may convey it to observation units within the framework of its technical and financial possibilities. As a rule, SURS only provides data obtained through the questionnaire, according to a defined procedure. For the 2002 census, SURS provides a print of the scanned census form.

1.4 Provision of individual data to business and interest associations on their members

For reasons of economizing the collection of data, and in order to relieve the reporting units, certain business and interest associations do not collect data for the elaboration of various analyses but ask SURS to provide them with such data.

Good practice: SURS conveys individual data on the members of associations upon the presentation of a prior consent of its members. This is done in compliance with the procedure.

1.5 Provision of individual data for the purpose of interviews

For the purpose of interviews, SURS may provide registered scientific research organizations and registered researchers-individuals only the following personal data: name and surname of the individual, his/her residence, year of birth, sex and profession.

2. PROVISION OF DIRECTORIES – LISTS OF THE REPORTING UNITS

2.1 Provision of directories from statistical surveys

SURS creates directories for statistical surveys from administrative databases in compliance with various criteria, adding, as a rule, data from various statistical collections. Therefore directories are statistical databases, and their provision is treated in the same way as the dissemination of individual data.

Good practice: SURS does not convey directories from statistical surveys outside the system of state statistics, and authorized contractors can only obtain them for the purpose of implementing the Annual Programme of statistical surveys.

If the user does not fulfil the conditions for obtaining the directory, he/she is redirected to the appropriate administrative source (if there is one).

3. PROVISION OF STATISTICALLY PROTECTED MICRODATA

3.1 Provision of statistically protected microdata to scientific research institutions and registered researchers

Good practice: statistically protected microdata for scientific research and analytical purposes is conveyed only to scientific research organizations and researchers who can prove their registration (extract from the PRS for institutions, extract from the register of researchers for individual researchers). Data is provided according to a prescribed procedure.

3.2 Provision of statistically protected microdata to government bodies

SURS facilitates the work of government administration bodies, being the group of users of statistical data with huge and specific data needs, in forming various policies, by allowing them to use statistically protected microdata.

Good practice: statistically protected microdata are provided to government administration bodies if it is intended for research or analytical purposes. However, their request is rejected if the purpose of using microdata is the granting of administrative advantages or disadvantages to particular business

entities or natural persons. The procedure for conveying this data is the same as for the case of scientific research institutions.

3.3 Provision of statistically protected microdata to Social sciences data archives

By conveying statistically protected microdata to data archives, SURS enables analytical and research work to be carried out.

Good practice: statistically protected microdata forwarded to the social sciences archives enjoys the highest microdata protection level.

4. STATISTICAL PROTECTION OF TABLES

4.1 Statistical protection of tables for various sets of contents

The statistical protection of tables protects business and personal data. This means that not all the tables presented by SURS need to be automatically statistically protected but only those whose disclosure could induce invasion into the business secrecy of a business entity, or the privacy of information of an individual.

Good practice: all publication tables are statistically protected if their disclosure might interfere with the:

- Information privacy of individuals, or
- Interests of business entities.

If the person responsible for the statistical survey determines that tables need not be statistically protected, he/she submits this proposal for approval by the Sector for general methodology and standards. The Data Protection Committee rules in disputable cases.

4.2 Dissemination of tabular data in cases where the entry of only one, but important reporting unit in the table, would cause statistical protection of the table according to the (n, k) rule

Since the Slovenian economy is relatively small, it happens in some cases that only one unit, or smaller number of units, is entered under one particular item of standard of industrial classification, or that such a unit is strongly prevailing.

According to the (n, k) rule, such a cell should be statistically protected with a primary screening, while at the same time ensuring secondary protection for a certain number of other cells in the table. This would in some cases imply a huge loss of information and prevent quality economic analyses for certain segments of the Slovenian economy.

Good practice: SURS leadership held individual conversations with some companies and an agreement was reached on the presentation of this unit's data according to the standard of industrial classification.

5. ORGANIZATIONAL RULES

5.1 Submitting requests for microdata to the Data Protection Committee, and adoption of decisions on the dissemination of microdata for secondary analysis

Resolution of requests must comply with the prescribed procedure.

5.2 The »Need to know« principle and the dissemination of statistically protected microdata for the statistical analysis of data

Since there is always a theoretical possibility of disclosing the observation units while disseminating microdata, we are trying to reduce the probability of disclosure as much as possible, including with the so-called »Need to know« principle.

Persons performing the statistical analyses of data do not primarily think about the possibility of disclosure of observation units, and often submit a request for microdata with more variables and more categories of these variables than would be necessary for the analysis.

Good practice: a request for microdata submitted by a representative of an institution should state the purpose for what kind of research work they are needed. Then the representative and the person responsible for the statistical surveys define the smallest possible set of variables and categories of variables which allow a quality secondary analysis of data.

5.3 Several researchers work on the same project of the statistical analysis of data

Good practice: the care for the protection of microdata is at the same time the responsibility of the institution and the personal responsibility of an individual who undertook to safeguard microdata through a contract or a statement on data protection. Therefore each researcher receives his/her own base of statistically protected microdata and deals with it in line with the statement. This base may not be disseminated to another researcher, even when the same project is involved.

5.4 The use of the same base of statistically protected microdata for different statistical analyses

Good practice: if a researcher needs the same statistically protected microdata for different projects, he receives one copy of statistically protected microdata. The contract concluded by SURS and the researcher should specify all projects. The deadline for destructing data coincides with the termination of the last project.

Statistical protection of microdata

The Statistical Office of the Republic of Slovenia statistically protects data against direct or indirect identification. In the statistical protection of data against direct identification we remove name and surname, address, single ID number of the citizen and the business entity, tax number and similar identifiers. In statistical protection against indirect identification we review some rare features, or the combination of features, and we statistically protect the appropriate data. The statistically protected microdata are also protected by controlling the number of directly linked data in a data set, the accuracy of classifications, by signing the contracts on working with microdata, and familiarizing the researchers of the issues of data protection.

Statistical protection of data and other ways of protecting data are used for directly collected data from the reporting units and for data obtained from the administrative databases. Statistical protection is not used for public information if these are not connected with the statistical data.

Statistically protected microdata are divided into the so-called "public" statistically protected microdata where statistical protection consists of preventing disclosure and which are used only for rough analyses, and training and education of future researchers. Such an example is the Archive of social sciences data at the Faculty of Social Sciences, which also contains some series of statistically protected microdata obtained with the statistical surveys programme. In addition, we provide the researchers with the bases of statistically protected microdata which are protected only against direct identification (in spite of this, in the case of intended use it is not possible to recognize the analyzed units), but they can be obtained only by signing a special contract and a statement.

Methods of disseminating the statistically protected microdata to researchers

The statistically protected microdata are available for researchers in a "secure room" and by "distance access".

The Statistical Office of the Republic of Slovenia has established an operation of the so-called »secure room«, where the computer has no link with the external environment, no diskette unit and without the possibility of printing, where a researcher could analyse the statistically protected microdata. Entries to the secure room are recorded and allowed only to the entitled individuals, and the rules have been set as regards work in the secure room, which guarantees the highest possible degree of data protection.

In line with the protocol on implementing the joint project of researchers' access to statistically protected microdata from distance, signed on 10.10.2005 by the Statistical Office of the Republic of Slovenia and the Ministry of Higher Education, Science and Technology, the so-called »distant access of researchers to statistically protected microdata« was established. The objectives of the project were:

- provision of distance access of researchers to statistically protected microdata,
- increased transparency in disseminating the statistically protected microdata to researchers,
- provision of a high statistical protection of microdata,
- preservation of public trust in the Statistical Office of the Republic of Slovenia with a simultaneous increased dissemination of statistically protected microdata to researchers,
- offering the researchers a simple, comfortable and fast, and from the regional point of view, a fair system of analyzing statistical data,
- increased possibility for informed decision-making in preparing and evaluating legal acts in Slovenia.

Such an access to statistically protected microdata is simple from the point of view of the researcher, since s/he can, through the system of passwords, accede from his/her own computer (after a contract has been signed and the statistically protected microdata prepared) to statistically protected microdata prepared for him/her via the Internet connection to a special server of the Statistical Office of the Republic of Slovenia, separated from the production servers. From the point of view of data protection, this method is safer than dissemination of data on CDs. Since the statistically protected microdata never leave the Statistical Office of the Republic of Slovenia, the researcher receives only the tabular results on the Internet, while any access to data is recorded. The disadvantage of this system is the price which is essentially higher than in disseminating the statistically protected microdata on CD (for the Statistical Office of the Republic of Slovenia, not for the researcher, and a limited set of software for the analysis of data).

International cooperation of the Statistical Office of the Republic of Slovenia in the statistical protection of microdata

The Statistical Office of the Republic of Slovenia is very active in setting international recommended standards to statistical offices in disseminating the statistically protected microdata to researchers, since it participates in the project of the United States Economic Commission for Europe and the Conference of European Statistics. The operation of Data Protection Committee at the Statistical Office of the Republic of Slovenia is the example of such good practice in the recommendations.

The Statistical Office of the Republic of Slovenia participates as one of the seven statistical offices of the EU Member States in the Eurostat Working Group for the Statistical Protection of Data.