

Procedure for dissemination of confidential data for scientific purposes

The Procedure is established based on subsection 38 (6) of the Official Statistics Act.

1. Purpose

- 1.1. This Procedure governs the dissemination of all confidential data at the disposal of Statistics Estonia, regardless of the source from which they were collected, for their use for scientific purposes, pursuant to section 38 of the Official Statistics Act.
- 1.2. Confidential data collected in the course of a statistical action which has not been included in the statistical programme shall be disseminated by Statistics Estonia pursuant to the same Procedure.
- 1.3. All actions that Statistics Estonia has to perform to prepare the confidential data for their transfer for scientific purposes and to check the results are classified as statistical actions outside the statistical programme.
- 1.4. Pursuant to section 20 of the Official Statistics Act, statistical actions outside the programme are subject to a fee.

2. Definitions

- 2.1. For the purposes of the Official Statistics Act, **confidential data** are data that permit either the direct or indirect identification of a statistical unit and thereby disclose micro-data.
- 2.2. **A statistical unit is directly identifiable** if the data include the name, address or personal identification code or registration code of the statistical unit.
- 2.3. **A statistical unit is indirectly identifiable** if, in the absence of characteristics that permit direct identification, the statistical unit can be identified on the basis of other data. In order to determine whether a statistical unit is identifiable, all possibilities that a third person may presumably use to identify the said statistical unit shall be taken into consideration.
- 2.4. **A statistical unit** is an object or subject described by the data collected, statistically processed or disseminated during the production of official statistics.
- 2.5. **Micro-data** are detailed data characterising a specific statistical unit.
- 2.6. **Statistical disclosure control** is a set of techniques that reduces the risk of identifying persons, enterprises or other organisations on the basis of disseminated micro- or macro-data and thereby reduces the risk of disclosing the data attributable to them. The main methods include the restriction of the quantity of disseminated data, e.g. the removal of the characteristics with higher detection risk from the data, and the modification of data, e.g. the recoding of data or micro-aggregation.
- 2.7. **The risk of disclosure of a statistical unit** indicates the probability of identifying an individual object or subject within a dataset, and of thereby disclosing the data of the identified unit.
- 2.8. **The impact of disclosure of a statistical unit** indicates the potential consequences for the identified unit as a result of the disclosure of data, e.g. the disclosure of sensitive personal data or

business secrets; and the potential consequences for Statistics Estonia, e.g. damage to its reputation.

3. Submission and processing of an application

- 3.1. A legal person or agency in need of confidential data for scientific purposes shall submit to Statistics Estonia a written application (sample in Annex 1).
- 3.2. The processing of applications for the dissemination of confidential data for scientific purposes and the counselling of applicants are organised by the Marketing and Dissemination Department of Statistics Estonia pursuant to the stipulated processing procedure (Annex 2).
- 3.3. Applications for the dissemination of confidential data for scientific purposes shall be reviewed according to their order of arrival by the Confidentiality Council called by a public servant of the Marketing and Dissemination Department.
- 3.4. The Confidentiality Council shall consider the substance of the application and make a decision whether the confidential data can be used for scientific purposes within ten working days from the receipt of all documents necessary for the review of the application.
- 3.5. The Confidentiality Council shall consider every dataset applied for, taking into account two aspects: the risk of identification of a statistical unit and the impact of identification of a statistical unit.
- 3.6. Based on the decision of the Confidentiality Council, the Director General of Statistics Estonia shall grant approval for the dissemination of data.
- 3.7. A public servant of the Marketing and Dissemination Department of Statistics Estonia shall notify the applicant in writing of the decision taken.
- 3.8. In case the application is approved, Statistics Estonia shall, before the dissemination of the data, enter into a contract with the legal person or agency that submitted the application. Pursuant to subsection 38 (2) of the Official Statistics Act, Statistics Estonia shall not enter into such contracts with natural persons.
- 3.9. Statistics Estonia shall not enter into a contract with a legal person or agency that has repeatedly violated the terms and conditions of the contract on dissemination of confidential data.

4. Confidentiality Council

- 4.1. The Confidentiality Council consists of the following public servants of Statistics Estonia: the Deputy Director General, the Head of Methodology and Analysis Department, the Head of Data Processing and Registers Department, the Head of Enterprise and Agricultural Statistics Department, the Head of Economic and Environmental Statistics Department, the Head of Marketing and Dissemination Department, the Leading Data Protection Specialist, and the Legal Specialist.
- 4.2. The Head of Marketing and Dissemination Department of Statistics Estonia holds the position of the Chair of Confidentiality Council.
- 4.3. The Confidentiality Council has a quorum if all members of the Council or persons substituting for them are present. Decisions shall be taken by a simple majority vote. Dissenting opinions shall be recorded in the minutes. If the votes in favour and the votes against are equal, the Chair has the deciding vote.
- 4.4. If a member of the Confidentiality Council is absent from work, he/she shall be substituted by the person specified in his/her job description.

5. Statistical processing, dissemination and storage of confidential data

- 5.1. Statistics Estonia shall prepare the data and metadata to be disseminated for scientific purposes.
- 5.2. The disseminated data are presented in the form of a matrix as SAS, SPSS, STATA or text files, where one record presents the data of one statistical unit. For every survey year, there is a separate data file. The data include the code of the survey, the survey year, the unique code attributed to the statistical unit and the variables collected and derived in the course of the survey.
- 5.3. The disseminated metadata contain the description of variables (name of variable in the database, explanation of content, type of variable, value range, classification or codification used, calculation process in the case of derived variables), the description of survey methodology and the questionnaires used for data collection. Metadata are made available together with the data.
- 5.4. If the disclosure risk and/or sensitivity are rated as high, the confidential data can be used only in a safe centre. Such datasets are, among others, the data of the population and housing census, the data of the Structure of Earnings survey, the data on success on the labour market.
- 5.5. If the disclosure risk and sensitivity are rated as low, the confidential data can also be delivered on removable devices.
- 5.6. The Methodology and Analysis Department of Statistics Estonia shall apply disclosure control to all results that the user wishes to receive at his/her email address.
- 5.7. If the results contain indirectly identifiable data, the user shall be warned on the first occasion, and on the second occasion the respective agency's contact person shall also be notified thereof.
- 5.8. Statistics Estonia shall specify on its website, under the ESMS metadata, those official statistical actions in the case of which the confidential data collected as part of the action have been systematised and furnished with metadata with a view to using them for scientific purposes.
- 5.9. Statistics Estonia shall store all confidential data disseminated to users for scientific purposes for 25 years in the form they were made available in for the particular research. Statistics Estonia shall store log files, users' intermediate results and other working files for one year after the expiry of the contract. If a user wishes that his/her intermediate results or working files be stored longer than a year, he/she shall cover the cost of their storage.
- 5.10. The Marketing and Dissemination Department of Statistics Estonia shall publish on its website a link or reference to the published research (an article, monograph, etc.), whereas this link or reference shall be provided by the author of the scientific research.

6. Safe centre

- 6.1. Safe centres are located in the offices of Statistics Estonia at Tatari 51 in Tallinn and at Ülikooli 1 in Tartu.
- 6.2. Subject to advance registration, the safe centres are available for use from 9 a.m. to 4 p.m. on working days, or round the clock through a VPN tunnel.
- 6.3. On a safe-centre computer, each user has a separate folder (20 GB) for every piece of research, wherein the respective user can save the intermediate results of his/her research, whereas the backup of these folders is ensured by Statistics Estonia. Each user has access only to the data specifically needed for his/her research.
- 6.4. Users have no possibility to print from the safe centre computers, copy files to external data media or use the Internet.

6.5. In case the user needs to save his/her own files on the safe-centre computer for research purposes, the given files shall be saved by a public servant of the Marketing and Dissemination Department of Statistics Estonia.

6.6. After the disclosure control performed at Statistics Estonia, a public servant of the Methodology and Analysis Department of Statistics Estonia shall send the research results to the user's email address. Upon the request of the user of a safe centre, a public servant of the Methodology and Analysis Department of Statistics Estonia may also deliver the research results on a removable device.

7. Supervision over implementation

7.1. Supervision over the implementation of this Procedure shall be exercised by the Marketing and Dissemination Department of Statistics Estonia.

Sample of the application for confidential data

The application shall set out the following information:

1. name of the legal person or agency;
2. registration code of the legal person or agency;
3. title of the research;
4. objective of the research;
5. name of the statistical action required for the research or list of the necessary data;
6. list of the data which the applicant has obtained from other sources and which the applicant wishes to link with the data applied for;
7. list of persons wishing to use the relevant data during the research (name, personal identification code, email address);
8. name and position of the person signing the contract;
9. form of publication of the research results (article, monograph etc.);
10. period of using the data (i.e. term of contract);
11. preferred format of the data (SPSS, STATA, text file).

(signature)

Name of the person with right of representation

Position

Name of the legal person or agency

Location

Statistics Estonia's procedure for processing applications for confidential data

	Action	Person responsible	Implementation date
Initiation of new research			
1.	A legal person or agency submits an application for confidential data to be used in the framework of a specific piece of research.	Agency conducting scientific research	T*
2.	The Marketing and Dissemination Department checks whether the application contains all the necessary information, specifies some details if necessary, and gives advice on the preparation of applications.	Marketing and Dissemination Department in cooperation with the Methodology and Analysis Department	T+2 working days
3.	The Confidentiality Council reviews the application and decides whether to approve or refuse it.	Confidentiality Council	T+10 working days
4.	The applicant is notified of approval or refusal (incl. reason for refusal).	Marketing and Dissemination Department	T+15 working days
5.	Entry into a contract for the provision of confidential data.	Marketing and Dissemination Department in cooperation with the Methodology and Analysis Department	Within reasonable time
6.	Preparation of data files (only if the required dataset has not been used before), incl. linking.	Methodology and Analysis Department in cooperation with other departments	Within reasonable time
Amendment or termination of contract			
7.	The legal person or agency sends an application for amendment of the contract (e.g. for addition or removal of persons, addition of data, extension or termination of the contract).	Agency conducting scientific research	
8.	Receipt and registration of the application.	Marketing and Dissemination Department	
9.	Processing of the application and making of amendments to the contract.	Marketing and Dissemination Department	
10.	Preparation of supplementary data, if necessary.	Methodology and Analysis Department in cooperation with other departments	
11.	Making changes in the system, incl. blocking users and archiving data upon the termination of the contract.	Marketing and Dissemination Department in cooperation with the Methodology and Analysis Department	
12.	Notification of users.	Methodology and Analysis Department	
Use of safe centre			
13.	The user schedules a time and place for the use of a safe centre (in Tallinn or in Tartu).	Marketing and Dissemination Department	
14.	Meeting the user at the agreed time and showing him/her to a safe centre.	Marketing and Dissemination Department	
15.	Installation of the researcher's own software in a safe centre (if necessary).	Marketing and Dissemination Department	

16.	Review and delivery of the results.	Methodology and Analysis Department	
17.	Answers to questions and resolution of problems.	Marketing and Dissemination Department in cooperation with the Methodology and Analysis Department	

*T denotes the day on which the application was submitted.