

Procedure for dissemination of confidential data for scientific purposes

1. Purpose

- 1.1. This Procedure governs the dissemination of all confidential data at the disposal of Statistics Estonia, regardless of the source from where 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 unless otherwise provided by the terms and conditions of the relevant contract.
- 1.3. All actions that Statistics Estonia has to perform to prepare the confidential data for their transfer for scientific purposes are classified as statistical actions not included in the statistical programme.

2. Definitions

- 2.1. For the purposes of the Official Statistics Act, 'confidential data' are the 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. '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 which 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 involve the restriction of the quantity of disseminated data, e.g. 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 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 Information and Marketing Service of 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 or employee of the Information and Marketing Service 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 evaluation of the application.
- 3.5. The Confidentiality Council shall consider every data set 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 or employee of the Information and Marketing Service 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 was satisfied, Statistics Estonia shall, before dissemination of the data concerned, enter into a contract with the legal person or agency that submitted the application.
- 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 Department, the Head of Population and Social Statistics Department, the Head of Enterprise Statistics Department, the Head of Price and Wages Statistics Department, the Head of Agricultural Statistics Department, the Head of Information and Marketing Service of the Marketing and Dissemination Department, the person responsible for personal data protection and the lawyer.
- 4.2. The Head of Methodology 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 them are present. Decisions shall be taken by the simple majority of votes. 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 a civil servant with the same capacity.

5. Statistical processing, dissemination and storage of confidential data

- 5.1. Statistics Estonia shall prepare the data and meta-data 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 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 meta-data contain the description of variables (name of variable in the database, explanation of content, type of variable, interval in between the values, applied classification or codification, calculation process in case of derived variables), the description of survey methodology and the questionnaires used for data collection. Meta-data can be downloaded from Statistics Estonia's web site.
- 5.4. If the disclosure risk and/or sensitivity are rated as high, the confidential data can be used only in a safe centre.
- 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 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 publish on its web site a list of official statistical actions, whereas the confidential data collected in the framework of the listed official statistical actions have been systematized and furnished with meta-data 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 are stored longer than a year, he/she shall cover the costs of their storage.
- 5.10. The Marketing and Dissemination Department of Statistics Estonia shall publish on the web site a reference to the publication site of the research (an article, monograph, etc.).

6. Safe centre

- 6.1. Safe centres are located in the offices of Statistics Estonia at Endla 15 in Tallinn and at Ülikooli 1 in Tartu.
- 6.2. Subject to advance registration, safe centres are available for use from 9 a.m. to 4 p.m. on working days, or round the clock through the VPN-tunnel.
- 6.3. On a safe centre computer, every 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 such 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 or employee of the Methodology Department of Statistics Estonia.
- 6.6. After the disclosure control performed in Statistics Estonia, a public servant or employee of the Methodology 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 civil servant or employee of the Methodology Department of Statistics Estonia may also deliver the research results on a removable device.

7. Supervision over implementation

- 7.1. Supervision over this Procedure shall be exercised by the Methodology Department of Statistics Estonia.

Prepared by: Tuulikki Sillajõe, Kaja Sõstra, Katre Viljalo, Ilvi Reitel

Sample of the application for confidential data

The application shall set out the following information:

- 1) name of legal person or agency;
- 2) registration code of legal person or agency;
- 3) title of the research;
- 4) objective of the research;
- 5) name of the statistical action or a list of data necessary for the research;
- 6) a list of data which the applicant has obtained from other sources and which the applicant wishes to link with the data applied for;
- 7) a list of persons wishing to use the relevant data during the research (given and surname, personal identification code, email address);
- 8) in case the use of personal data is involved, a confirmation issued by the Data Protection Inspectorate to prove that the applied organisational, physical and information technology-related security measures are sufficient and, if an ethics committee has been founded, also the opinion of such committee;
- 9) in case the use of sensitive personal data is involved, a confirmation issued by the Data Protection Inspectorate proving that the processing of sensitive personal data has been registered.

(Signature)

Given and surname of the person with the right of representation

Position

Name of legal person or agency

Place

Statistics Estonia's Procedure for processing applications for confidential data

	Action	Responsible person	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 Information and Marketing Service of Statistics Estonia checks whether the application contains all the necessary information, asks specifying questions if needed, and advises how to prepare an application.	The Information and Marketing Service of the Marketing and Dissemination Department in cooperation with the Methodology Department	T+2 working days
3.	The Confidentiality Council reviews the application and decides whether to approve or refuse it.	The Confidentiality Council	T+10 working days
4.	Delivery of a positive decision or a grounded decision of refusal.	The Information and Marketing Service of the Marketing and Dissemination Department	T+15 working days
5.	Entry into a contract for the provision of confidential data.	The Methodology Department in cooperation with the General Department	Within reasonable time
6.	Preparation of data files (only in case the data required for the research are data not used before), incl. linking.	The Methodology Department in cooperation with respective subject area 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 contract).	Agency conducting scientific research	
8.	Receipt and registration of the application.	The Information and Marketing Service of the Marketing and Dissemination Department	
9.	Processing of the application and making of amendments to the contract.	The Methodology Department	
10.	Preparation of supplementary data if necessary.	The Methodology Department in cooperation with respective subject area departments	

	Action	Responsible person	Implementation date
11.	Making changes in the system, incl. blocking users and archiving data upon termination of contract.	The Data Processing Systems Department	
12.	Notification of users.	The Methodology Department	
Use of safe centre			
13.	The user schedules a time and place for use of a safe centre (Tallinn or Tartu).	The Information and Marketing Service of the Marketing and Dissemination Department	
14.	Meeting the user at the agreed time and showing him/her to a safe centre.	The Information and Marketing Service of the Marketing and Dissemination Department	
15.	Installation of the researcher's own software in a safe centre (if necessary).	The Data Processing Systems Department	
16.	Review and delivery of results.	The Methodology Department	
17.	Answers to questions and resolution of problems.	The Methodology Department and the Data Processing Systems Department	

*T denotes the day when the application is submitted.