

Guidelines for access to microdata from Statistics Norway

Applicable from 1 January 2021

Contents

1. Introduction.....	2
2. What can data from Statistics Norway be used for, and who can access them	2
2.1. What can data be used for	2
2.2. Who can access microdata	3
2.2.1. Research institutions	3
2.2.2. Public authorities and other bodies	3
3. Accessible data	3
3.1. Source data for official statistics	3
3.3. Timing of access to data	5
3.4. Data minimisation	5
3.5. Basic data.....	6
4. Special requirements applicable to public authorities.....	6
5. Requirements for application.....	7
6. Application processing at Statistics Norway	7
7. Requirements for processing by the recipient	8
8. Serial numbers.....	9
9. Key exchange and storage of scrambling key	9
9.1. Exchanging the scrambling key	9
9.2. Storing the scrambling key	9
10. Providing data internationally	9
11. Pricing.....	10
12. Technical aspects of access	11

1. Introduction

These guidelines supplement [Act no. 32 of 21 June 2019 relating to official statistics and Statistics Norway, Section 14](#) (the Statistics Act) and [Regulation no. 2731 of 11 December 2020 relating to official statistics and Statistics Norway](#) (the Statistics Regulation).

In addition to the Statistics Act and Statistics Regulation, these guidelines are also based on relevant preparatory works: Chapter 9 of [NOU 2018:7](#), and the Ministry's deliberations in Chapter 6.2.4 of the [Proposition to the Storting no. 72 \(2018-2019\)](#). In this document, the term 'preparatory work' refers to Proposition to the Storting no. 72 (2018-2019).

The guidelines apply to access to microdata (data on individual units), as well as access to statistics (aggregate data) that do not meet the statistical confidentiality requirement in the Statistics Act and are not therefore accessible in Statistics Norway's StatBank. In this document, *data* is used as a generic term.

Other relevant data sources from Statistics Norway, which are *not* covered by these guidelines, are as follows:

- [StatBank¹](#)
- [Open data API²](#)
- [Microdata.no³](#)

These are data sources that provide easy access to statistical results. Statistics Norway always recommends investigating whether one of these data sources can meet a user's specific data needs.

Section 14 of the Statistics Act, the associated regulations and these guidelines apply to access to non-anonymised data. The term 'non-anonymised data' covers a range of data; from virtually anonymous (anonymised) data to identifiable data.

2. What can data from Statistics Norway be used for, and who can access them

2.1. What can data be used for

Under Section 14 (1) of the Statistics Act, Statistics Norway is authorised to provide access to microdata for the production of statistical results and analyses, including for research purposes, when this does not entail a disproportionate disadvantage to other interests. The delimitation of the purpose of access is based on Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European Statistics.

The various statistical purposes of 'statistical results and analyses, including for research purposes' are defined below in more detail, and Chapter 2.2 describes who may be given access to microdata.

Research typically refers to active, thorough and systematic investigation using scientific methodology to generate new findings and add to the existing knowledge base.

Statistical⁴ results and analyses are defined as collated and processed data on individual statistical units in a group, or a selection of these units, or such data produced by means of systematic observation of the phenomenon, as well as systematic examination/assessment of such.

¹ StatBank contains detailed tables of time series. Users can select which figures they want to use and save them in various file formats.

² Statistics Norway offers an API where the user can quickly retrieve Statistics Norway's table data and integrate them with their own systems.

³ Through the microdata.no service, researchers and students have immediate and self-service access to anonymous data from Statistics Norway without the need to apply.

⁴ The Statistics Act, Section 3 (1) (a)

The use of data for non-statistical purposes, such as administrative, legal or fiscal purposes, or for checks on statistical units, is prohibited. The data cannot, therefore, be used to influence the legal position of the person to whom the data applies. In the preparatory works, this is further exemplified by the fact that data can be used to develop general systems and algorithms for use in tools aimed at optimal utilisation of the resources in a public authority. However, data on an individual or business cannot be used to adapt the case processing to that specific person or business.

2.2. Who can access microdata

As provided for in Section 1 (1) of the Statistics Regulation, access to data pursuant to Section 14 of the Statistics Act may be granted to research institutions for research purposes, and to public authorities and other defined bodies for the production of statistical results and analyses. The relevant bodies are named in Section 1 (1) (a)-(e) of the Statistics Regulation, and are discussed in more detail in Chapter 2.2.2. What constitutes a research institution is covered in Section 1 (2) of the Statistics Regulation, and is discussed in more detail in Chapter 2.2.1.

2.2.1. Research institutions

A research institution is defined as an organisation that has been recognised as a research entity by the [Research Council of Norway](#) or the statistical office of the European Union, [Eurostat](#).

Statistics Norway may recognise an organisation as a research entity even if it is not recognised as such by the Research Council of Norway or Eurostat. In processing the application, Statistics Norway will base its decision on the [Research Council of Norway's criteria](#), but may deviate from these, for example with regard to the requirement about dividends and the requirement to separate the research activity from financial activity. Applications for recognition as a research entity should be submitted to mikrodata@ssb.no. The following must be included with the application as a minimum:

- The entity's Articles of Association and the most recent annual report
- Certificate of registration from the Brønnøysund Register Centre
- An account of the independent research carried out by qualified research personnel, including details of such research within the last three years, specifying funding sources and amounts, as well as an overview of publishing activity from this research
- An overview of the R&D full-time equivalents and number of employees

2.2.2. Public authorities and other bodies

The bodies listed below may be granted access to data for the production of statistical results and analyses.

- a) Central government bodies,⁵ county authorities and municipal authorities
- b) Norges Bank
- c) The Office of the Auditor General of Norway
- d) National statistical offices in other countries⁶
- e) International organisations on the basis of an agreement under international law⁷

3. Accessible data

3.1. Source data for official statistics

Access can be given to the source data⁸ that Statistics Norway has obtained for the development, production and dissemination of official statistics, with the exception of personal health data as

⁵ For example, ministries and directorates. Statutory corporations and state-owned companies etc. are not included.

⁶ For example, Statistiska centralbyrån (Statistics Sweden) and Danmarks statistik (Statistics Denmark)

⁷ For example the UN, OECD and EU

⁸ The Statistics Act, Section 10

described in Chapter 3.2.

With the exception of personal health data, there are no absolute restrictions on what data can be accessed, but ***access presupposes that this does not entail a disproportionate disadvantage to other interests***, as provided for in Section 14 (1) of the Statistics Act.

The preparatory works emphasise that the assessment of disadvantage will be used to inform a specific and broad assessment prior to the decision on access. The preparatory works mention the following relevant interests that need to be taken into account in the assessment: privacy protection, maintaining public confidence in Statistics Norway, the recipient's ability to manage the data in a secure manner, protection of commercial interests and national defence and security interests.

The data that Statistics Norway may grant access to in individual cases will depend on a specific and broad assessment of the disadvantage of accessing the data. In this assessment, Statistics Norway will consider the nature of the requested data, and the relevant datasets will be assessed both in isolation and collectively. For example, data may not be considered to entail any significant disadvantage if accessed in isolation, but this may change when the data are assessed in conjunction with other data. The sensitivity of the data is one of the factors that must be taken into account in the access assessment, and here the preparatory works refer specifically to special categories of personal data as described in Article 9 of the General Data Protection Regulation (GDPR). Special categories include data concerning health, criminality and trade union membership.

The data's level of detail and identification are also relevant criteria for assessing whether access will entail a disproportionate disadvantage to other interests. In this context, it is also relevant to consider the requirement set out in Section 14 (2) of the Statistics Act, that the data must not contain a greater level of identification than is necessary for the purpose in question (data minimisation). The requirement for data minimisation in light of the specific purpose is covered in more detail in Chapter 3.4.

Based on statements in the preparatory works, access is not given to data if, for example:

- the data have commercial value and this value is diminished if the data become accessible
- access will violate someone's rights under the Copyright Act or other legislation
- access poses a risk to national security
- access does not meet the requirement for informed consent

Furthermore, the preparatory works to the Statistics Act indicate that Statistics Norway must exercise particular caution in relation to applications to access microdata that were obtained directly from individuals and enterprises in its own statistical surveys. Access to such data is generally only granted to researchers for research purposes. Statistics Norway can make exceptions to this main rule and give public authorities access to data that were obtained directly from individuals or enterprises if this is done in a way that will not diminish public confidence in Statistics Norway, for example by ensuring that the data are sufficiently anonymised. See also Chapter 4 *Special requirements for public authorities*.

If data have been retrieved from other authorities, the preparatory works indicate that Statistics Norway should, where necessary, consult with the relevant authority to establish whether there are restrictions on the use of the data that need to be taken into account.

An overview of the data in Statistics Norway that are often used for statistics, analysis and research is available on [Statistics Norway's website](#).

3.2. Personal health data⁹

As stipulated in Section 2 (1) of the Statistics Regulation, no access can be granted to data that Statistics Norway has retrieved from central health registers founded under Act no. 43 of 20 June 2014 on Personal Health Data Filing Systems and the Processing of Personal Health Data. Personal health data can be collated with data from Statistics Norway.

For access to personal health data, see www.helsedata.no.

3.3. Timing of access to data

The timing of access to data is regulated in Section 2 (2) of the Statistics Regulation, which indicates that access to the source material for official statistics can only be granted once the statistics have been published. This normally means the point at which the relevant official statistics become accessible to the public.

Data that Statistics Norway has collected exclusively for development or quality assurance purposes can be accessed once they have been produced in a statistical format and adapted for access. This will typically be in cases where Statistics Norway is giving consideration to establishing new statistics that are wholly or partly based on data that are not already included in any other source material. In such cases, new data will need to be collected in order to develop and test the new statistics. Another example is data that are used when editing the source material, i.e. for quality assuring the content of the source material. This typically relates to confirming/ruling out anomalies in individual variables in the source material.

3.4. Data minimisation

Pursuant to Section 14 (2) of the Statistics Act, the data must not entail a greater level of identification than is necessary for the purpose in question.¹⁰ The level of identification must also be considered in light of the risk of the data being used in contravention of the stated purpose and the risk to public confidence in Statistics Norway.

Furthermore, Section 14 (3) of the Act stipulates that, in general, the data must not contain any characteristics that allow direct identification of persons or enterprises. This means that data must be anonymised or deidentified.

‘Anonymised data’ is defined as data that cannot be traced to an identified or identifiable statistical unit. ‘Indirectly identifiable data’ is defined as data in which all directly identifiable variables have been deleted or replaced by serial numbers, see Chapter 8.

The provision also states that Statistics Norway can only make directly identifiable data available for collation purposes and for other processing if this is absolutely necessary for achieving the stated purpose and if there are no grounds for concern in relation to the data subject. Characteristics that allow direct identification of persons or enterprises must be deleted as soon as the processing has been completed. According to the preparatory works, access to directly identifiable data for collation purposes is contingent on the data being collated in accordance with relevant legislation.¹¹

Data from Statistics Norway will normally be provided in an indirectly identifiable format, which means that characteristics that allow direct identification have been removed and that the level of detail reflects the specific purpose of the processing. This will be in accordance with the principle of data minimisation, see Article 5 (1) (c) of the GDPR, which means that the data must be appropriate,

⁹ Section 2 (1) of the Statistics Regulation

¹⁰ As this appears from the application for access and other relevant documentation

¹¹ See for example, Section 3-1 of the Norwegian Patient Register Regulation, and Section 3-1 of the Norwegian Cardiovascular Disease Registry Regulation

relevant and limited to what is necessary for the purpose of the processing. When access is granted to data in an indirectly identifiable format, variables that allow direct identification are normally replaced by a serial number, making it possible to follow a particular unit over time. The use of serial numbers is discussed in more detail in Chapter 9.

A typical example of data minimisation is aggregation; such as age groups, income ranges, education level categories, etc.

3.5. Basic data

Identifiers that are a particularly important basis for economic and business statistics and which are registered and accessible in the Central Coordinating Register for Legal Entities are called 'basic data'.

Statistics Norway can grant access to basic data on legal persons (and subunits thereof) registered in the Central Coordinating Register for Legal Entities, including characteristics that allow direct identification (organisation number, name, etc.). Access to and use of basic data from Statistics Norway is also limited to the production of statistical results and analyses, including for research purposes.

Basic data consists of the following:

Enterprises:

Organisation number, registered name, edited name, postal address, business address, telephone number, fax number, email address, web address, date of incorporation/registration, industry classification, organisational structure, institutional sector code, status (deleted/not deleted), number of employees, corporate group affiliation, share capital/subordinated loan capital and roles.

Activity/unit in an enterprise with more than one activity:

Organisation number, registered name, edited name, unit name, postal address, location address, numerical address, basic statistical unit, coordinates, telephone number, fax number, start-up date, industry classification, status and number of employees.

4. Special requirements applicable to public authorities

The preparatory works to the Statistics Act stipulate that anonymous data will often be sufficient for public authorities' purposes. The main rule is therefore that public authorities can be given access to anonymised data for use in statistical results and analyses.

The principle of data having to be anonymised does not apply to the following data in the National Population Register: gender, age, place of residence, country of birth and date of death, or data that are freely accessible to the public. The exception from anonymisation of this data presupposes that the data are not collated with other data.

If a public authority applies for access to a type of non-anonymised data that is not included in the list above, it must explain how it will ensure that the data are not used for purposes other than the production of statistical results and analyses. This must include a description of safeguards to ensure that the data are not used to exercise authority, for management/control purposes or for other non-statistical purposes.

Statistics Norway's assessment of whether special conditions of access are required will be based on the nature of the information and the public authority's other tasks and organisation.

Where a public authority outsources the statistical work/analysis, the preparatory works allow for the public authority to access and forward data to the party performing the work on its behalf. The application for access must indicate that data will be forwarded to a third party, including to whom and by what means. In such cases, the party in receipt of the data will have the same duty of confidentiality as the public authority. The public authority must take responsibility for ensuring that

the party performing the work can process data in a secure manner and comply with the duty of confidentiality. The public authority will also be responsible for ensuring that the data are processed in accordance with the terms of use.

5. Requirements for application

Applicants can request extra assistance from Statistics Norway in the form of individual follow-up. This will help the applicant in the planning phase and help ensure that the application for access is robust in areas such as choice of data sources, variables and level of detail, delimitation of the study population/sample, description of data minimisation, increased awareness of data collation, etc. This assistance can help clarify whether applicants are likely to be granted access to the relevant data.

In order to assess whether the conditions for access have been met, Statistics Norway needs the following information from the applicant:

- The purpose of the access
 - o *research: research protocol or project description*
 - o *statistical results/analysis: project description or description of needs and purpose*
- Processing period (If the data need to be retained beyond the processing period, e.g. in the case of external documentation requirements for post-controls, Statistics Norway wishes to be informed of this.)
- Who is responsible (institution and person)
- What population, data and time periods are sought
- Specific reasons for requiring variables that are particularly identifiable (e.g. exact dates, coordinates/basic statistical unit, etc.)
- Where and how data will be stored (including use of external secure analytics platforms such as TSD, Safe, etc.)
- List of people (and their institutional affiliation) who will have access to and process the data
- The applicant's legal basis for the processing of personal data according to Articles 6 and 9 of the GDPR
- Submission of a Data Protection Impact Assessment (DPIA), or a reason why the DPIA is not necessary. (*These documents/assessments are requested because they will provide relevant information for use in Statistics Norway's assessment of the disadvantage to privacy protection, see Section 14 of the Statistics Act.*)
- Description of the linking process, as well as variable lists showing what data will be collated with Statistics Norway's data

Other requirements may also be set for the application and documentation, for example if the application concerns particularly identifiable variables and/or the collation of data from Statistics Norway that are subject to consent.

6. Application processing at Statistics Norway

Statistics Norway processes applications in accordance with the Public Administration Act, the Statistics Act, the Statistics Regulation and these guidelines. Each application is assessed on its own merits.

Access will be granted if the conditions for access are met. The access decision will set out the requirements for the recipient's processing of the data. The decision can be appealed to the Ministry of Finance in accordance with the rules of appeal in Chapter VI of the Public Administration Act.

According to the preparatory works, access to data shall, in principle, be granted if the conditions are met. However, it is conceivable in some cases that conditions may be met in principle, but that very specific reasons exist for not granting access. This acts as a safeguard.

If the conditions for access are not met, the application for access will be declined. This decision can also be appealed.

7. Requirements for processing by the recipient

Pursuant to Section 14 (5) of the Statistics Act, if access is granted to data that are subject to a duty of confidentiality pursuant to Section 8 of the Statistics Act, a corresponding duty of confidentiality is also imposed on the recipient of the data.

Under Section 14 (4) of the Statistics Act, Statistics Norway is also authorised to set conditions for access to and use of data. Such conditions must be proportionate and justifiable.

Statistics Norway normally sets the following conditions:

Data must only be used for the stated purpose and within the framework of the decision to grant access.

The person responsible for receiving the data shall ensure that the duty of confidentiality under the Statistics Act is complied with, that the data are processed and stored as described in the access decision, and that basic requirements for information security are met. This entails ensuring adequate access control, logging and subsequent control. Storage of data on mobile/local devices is not acceptable.

The person responsible for receiving the data must ensure that access is only given to employees with an official need for the data. Furthermore, before granting access to an employee, he/she must be made aware of the duty of confidentiality under the Statistics Act. A complete overview of who has access to data must be provided upon request from Statistics Norway.

Unwarranted identity disclosure and attempts to reconstruct the identity of a statistical unit (person/enterprise) in the data provided are not permitted.

Publication and other public disclosure of results, analyses etc. which are based on the data provided must not be in a format that enables individual units to be identified or information about such units to be revealed (anonymised).

All publications of results based on the data provided must cite Statistics Norway as the source. Statistics Norway is not responsible for analyses or interpretations based on the data in question.

The final report of the research project shall be published within a reasonable period of time after conclusion of the project.

Confirmation must be given to Statistics Norway that the data have been deleted/returned when access is no longer needed, and this must be no later than the expiry of the deletion deadline stipulated in the access decision.

Statistics Norway may impose additional justifiable and proportionate conditions in individual cases.

Before the data can be accessed, the applicant must accept the conditions of access as they appear in the decision. This is done by dating, signing and returning the decision. The signatory must be authorised to accept the conditions.

If the conditions of the decision are violated, an assessment may be made of the severity of the

violation to determine whether new conditions should be imposed. These can include a requirement to delete/return the data to Statistics Norway. Violations of the terms of the decision may also be taken into account when assessing future applications for access. Violation of the duty of confidentiality under Section 8 of the Statistics Act is a criminal offence, see Section 209 of the Penal Code.

8. Serial numbers

When access is given to data in an indirectly identifiable format, project-specific serial numbers are used. This means that all directly identifiable variables are deleted or replaced by a project-specific serial number.

Access to reuse *datasets with project-specific serial numbers* in new projects may, upon application, be permitted for projects where there is no temporal overlap.

Statistics Norway shall facilitate the reuse of datasets with new project-specific serial numbers to the greatest extent possible.

9. Key exchange and storage of scrambling key

A unique identifier will normally be needed to collate data from different sources at unit level. In line with the preparatory works, data collation should generally take place through the exchange of a unique identifier that is not directly identifiable. Alternatively, Statistics Norway can collate the data.

The unique identifier is referred to as a 'scrambling key' and Chapter 9.1 below explains when this is exchanged. Chapter 9.2 sets out the requirements for storing the scrambling key.

9.1. Exchanging the scrambling key

Distributed linking is generally used when publishing indirectly identifiable data from Statistics Norway that is linked to data from sources outside Statistics Norway.

This entails those responsible for the various data sources exchanging a data file containing a personal identification number or other unique identifier and a randomly generated serial number (scrambling key). The individual authorities then add serial numbers to their own data and send the data with the serial numbers to the project.

Distributed linking is not permitted if the applicant/institution is also responsible for data that will be collated with data from Statistics Norway, unless the institution can document that key management is kept completely separate from the research/analysis unit, typically through an independent computer department.

As a general rule, distributed linking is also not permitted if the result of the collated data is to remain anonymous, unless otherwise provided for by regulation.

In cases where distributed linking is not possible, the data normally have to be collated by Statistics Norway.

9.2. Storing the scrambling key

In cases where Statistics Norway defines the population, Statistics Norway will create and store the scrambling key. The key recipient then confirms that the key has been deleted once the project has received and collated the data.

10. Providing data internationally

Researchers at an approved research institution or body (see Section 1 (1) (d) and (e) of the

Statistics Regulation) outside Norway may, as a general rule, be granted access to anonymised data. This is based on the preparatory works, which stipulate that there will generally be a higher risk of data being disclosed to unauthorised parties when they are sent out of the country. Access to non-anonymised data from Statistics Norway outside Norway also poses some special challenges in relation to following up of any violations of the duty of confidentiality under the Statistics Act.

Following a specific overall assessment, researchers at an approved research institution or body within the EU/EEA (see Section 1 (1) (d) and (e) of the Statistics Regulation) may, as an exception, be granted access to indirectly identifiable data. In its assessment, Statistics Norway will place importance on measures to address the increased risk of data processed outside Norway's jurisdiction. In such cases, requirements are generally set for a specially adapted agreement with the foreign research institution/authority to ensure that Norwegian rules of law are applied and that a Norwegian legal venue is agreed.

Reference is also made to the international cooperation between the Nordic national statistical institutes on data access: the NordMAN project, which provides a unified access point to cross-Nordic microdata. The condition is that these data are stored and accessible via a remote access solution from one of the Nordic countries' national statistical offices.

11. Pricing

Under Section 3 of the Statistics Regulation, Statistics Norway may impose a charge for customising and providing access to data. This charge must not exceed the actual costs of processing the application, providing access and following up the individual application.

Statistics Norway can charge for the time spent processing the application, and customising and providing access to data in each case. A 25 per cent VAT charge is also payable.

Successful applicants will receive a quote from Statistics Norway, which will include the price and delivery date.

The price normally covers the following activities:

- Coordination, clarification, preparation of quote and agreement
- Creating a population, processing third-party data and customising data files
- Data extraction and de-identification
- Archiving, documentation and delivery

Coordination, clarification, preparation of quote and agreement

This item includes the preparatory work involved in producing quotes and agreements. The preparatory work will depend on the complexity and scope of the individual project.

Creating a population, processing third-party data and customising data files

The time spent creating a population depends on the complexity and number of populations included in the project. Before the data extraction work begins, the executive officer must also prepare the files for extraction. The work can be particularly time-consuming when third-party data are included in the project. Lack of documentation, difficult file formats, and work involved in exchanging data with other suppliers are some of the factors that can prolong the process and push up the price.

Data extraction and de-identification

After the quote has been accepted and data files have been prepared, most of the remaining work concerns extracting and de-identifying data.

- *Extractions from FD-Trygd*
The event history database for welfare benefits is an important data source in many research projects. The database is specially adapted for use by researchers, and a fixed price based on actual time spent is charged for extractions from individual tables.
- *Other extractions from variable lists*
A fixed price is applied to each statistical area in extractions of variables from Statistics Norway's variable lists. The extractions are generated from customised statistics files and the work is charged according to the number of statistical areas from which variables are required.
- *Extractions outside variable lists*
Extracting data and documenting content that is not covered in variable lists can be challenging and costly.

Archiving, documentation and delivery

During the process, and upon delivery of the finished product, the executive officer archives the correspondence and data and documents the data material included in the delivery.

12. Technical aspects of access

Data sent to/from Statistics is exchanged via fx.ssb.no (secure server). Statistics Norway sends recipients a link with instructions.

All files to and from Statistics Norway are encrypted and sent via Statistics Norway's secure area for file transfer. Encryption passwords are sent by text message. The files are automatically deleted from Statistics Norway's server after six months. Data from other suppliers are sent to Statistics Norway as agreed with the executive officer. The recipient is responsible for ensuring that data from Statistics Norway are transferred to secure areas at their institution, or to a secure analytics platform for third parties, before decryption. The University of Oslo's service for sensitive data (TSD) and the secure access to research data and e-infrastructure service at the University of Bergen (SAFE) are examples of secure analytics platforms.