

Conditions for granting access to confidential statistical data for scientific purposes

Posledná aktualizácia:13.12.2018 | Počet zobrazení:null

Research institutions, especially

1. Universities and other higher education organizations
2. Organizations or research institutions

may have granted access to the confidential statistical data for scientific purposes. PhD students as well as researchers should request for the data via the research institution whose employees they are.

Statistical Office of the SR shall provide access to the confidential statistical data either by sending data to research institutions or allowing access to Safe centre of the SO SR.

SO SR shall grant access when

1. [Research proposal application form doc](#) (175 kB), [pdf](#) (266 kB) has been correctly filled in and submitted;
2. Contract on providing confidential statistical data for scientific purposes between the research institution and the SO SR, which specifies conditions for access to the data, obligations of the contract parties, measures to ensure statistical confidentiality and penalties for violation of these obligations, was signed;

Every individual researcher working with the data shall comply with the [Instructions for handling confidential statistical data for scientific purposes](#) (pdf - 40,0 kB) ([version for access to the data via the Safe centre of the SO SR](#) (pdf - 39,9 kB)).

Research proposal application form may be sent to info@statistics.sk or to the [postal address of the SO SR](#).

Just direct identifiers are removed from microdata available via Safe centre; other statistical disclosure control methods are applied to the researchers' outputs. The information loss in microdata is therefore equal to zero, in the outputs it is proportionate to the re-identification risk. Researchers working with the microdata in the Safe centre are led to make such outputs that do not need to be protected (due to no risk of re-identification).

Microdata are sent to research institutions in form of anonymized dataset, i.e. protected such that re-identification risk is minimized and maximum possible information value is preserved. Especially some of these methods are applied:

- global recode – several categories of a grouping variable are aggregated;
- suppression – identifying value is suppressed;
- top/bottom coding – extreme values are replaced by some open interval;
- swapping – values between respondents are exchanged;

- microaggregation – numerical value is changed by a mean value.

Information loss is therefore always non-zero and depends on the detail of requested microdata.

Microdata available via Safe centre has then less information loss than microdata that are sent to research institutions.