

How-To

Maël-Luc Buron, Peter-Paul de Wolf, Bernhard Meindl, Tobias Enderle

2021-06-02

protoTestCensus

This repository was created in order to guide census experts to test prototype software versions of targeted record swapping and the cell-key method for Census tables. Over time, the information in this repository became outdated as the developed methods have been included in different software packages/tools. So this document now refers to the most up-to-date documentation for the respective tools, as well as to the latest releases of these tools.

How to use the open-source versions of the recommended census protection methods

(This text is also available as a [pdf](#).)

In the Specific Grant Agreement (SGA) called *Harmonised protection of CENSUS data in the ESS* two SDC methods have been developed and tested. In deliverables 3.3 and 3.4 recommendations were given on how to protect the hypercubes and the grid data in the next Census. The two methods (record swapping and the cell key method) can be applied independently from each other.

During that SGA some initial SAS code was used to test the methods. For reference: the SAS code for the cell key method can be found [here](#).

In a follow-up SGA called *Open source tools for perturbative confidentiality methods* the record swapping and the cell key method were implemented in the open-source software R and Argus. These implementations can be downloaded on the respective Github-repositories which will be linked from this document.

Part 1 is related to targeted record swapping while part 2 focuses on the cell-key method. The methods can be used independently: it is either possible to compute unperturbed hypercubes with swapped microdata (part 2 can be ignored), perturbed hypercubes with original microdata (part 1 can be skipped), or perturbed hypercubes with swapped microdata (part 1 then part 2).

Part 1: Targeted record swapping

To apply record swapping to original microdata it is possible to use either the R package `recordSwapping` (paragraph 1.1) or `MuARGUS` (paragraph 1.2).

Both implementation use the same underlying new C++ code which is optimized to be incredibly fast. This new implementation of the targeted record swapping method is intended to be very close to the SAS codes tested in the previous SGA based on the ONS codes. More information regarding the small differences can be found in the vignette of the `recordSwapping` package.

1.1 R-package `recordSwapping` The package is developed at github.com/sdcTools/recordSwapping where also installation instructions and example usage as well as an introductory vignette can be found.

1.2 MuARGUS The latest release of **MuARGUS** that includes the implementation of record swapping is available at github.com/sdcTools/muargus/releases where also a quick reference documentation with explanations and screenshots can be downloaded.

Part 2: Cell-key method

The cell-key method can be applied either with the R package **cellKey** (paragraph 2.2) or **tauArgus** (paragraph 2.3). Both implementation use the **ptable** package (paragraph 2.1).

2.1 R-package ptable The package is developed at github.com/sdcTools/ptable where also installation instructions and example usage can be found. The package includes a vignette and a graphical user interface and it can draw plots that describe the perturbations that will be applied to the hypercubes based on a specific pTable with the **cellKey** package, **tauArgus** or SAS macros.

2.2 R-package cellKey The package is developed at github.com/sdcTools/cellKey where also installation instructions and example usage can be found. The package includes also a detailed vignette that helps users getting started with the functionality of the package.

2.3 TauARGUS The latest release of **TauArgus** that includes the implementation of the cellKey method is available at github.com/sdcTools/tauargus/releases where also a quick reference documentation with explanations and screenshots can be downloaded.