News on R packages for SDC

Johannes Gussenbauer

Wien, 29.02.2024

Overview packages

- `sdcTable`

  — cell suppression, primary and secondary

  — protect / attack cross tables

  — fast heuristic against exact disclosure

  — allows to prepare inputs for Argus and call it

  — on CRAN `install.packages("sdcTable")`

Overview packages

- `sdcMicro`

  — Methods to protect categorical and numerical key variables using

  — suppression, perturbation, swapping, recoding, …

  — provides risk and utility measures for micro data

  — interactive GUI (`` `sdcMicro::sdcApp() ``) for `R` novices

  — on CRAN `install.packages("sdcMicro")`

  — Targeted Record Swapping

Overview packages (2)

- `cellKey`

  — applies the cellkey-method on frequency- and magnitude tables

- `ptable`

  — creates required lookup-tables for perturbation

- `sdcHierarchies`

  — allows to define (nested) hierarchies for cross tables

  — can also be used in `sdcTable`

  — on CRAN `install.packages("sdcHierarchies")`

Overview packages (3)

- synthetic micro data

  — create micro data using ML models

  — → packages `simPop`, `synthpop`

- disclosure control for spatial (raster-based) data

  — sdcSpatial

  — on CRAN `install.packages("sdcSpatial")`

sdcMicro

https://github.com/sdcTools/sdcMicro

- Currently not on CRAN due to stricter CRAN-Checks (C++ related);

- Version 5.7.8 is on github;

- Many updates for recordSwap() (only R-related not C++-related)
  — better documentation, improved tests and updated vignette
  — custom risk tested

- Bugfix in sdcApp()

sdcTable

https://github.com/sdcTools/sdcTable

- new version on CRAN (0.32.6)

- Improvements in *SIMPLEHEURISTIC* when adding additional suppressions

pTable, cellKey

https://github.com/sdcTools/ptable,
https://github.com/sdcTools/cellKey

- pTable: on CRAN since 2023-03-01 in version 1.0.0 (https://cran.r-project.org/web/packages/ptable/index.html)

- cellKey: on CRAN since 2023-03-17 in version 1.0.2 (https://cran.r-project.org/web/packages/cellKey/index.html)

  — Performance Fix: Do not compute contributing indices (relevant only for the perturbation of numerical variables) if no such variables were specified in ck_setup()

sdcHierarchies, sdcSpatial

https://github.com/bernhard-da/sdcHierarchies,
https://github.com/edwindj/sdcSpatial/tree/master

- sdcHierarchies and sdcSpatial on CRAN

sdcHierarchies:

- massive performance improvement in hier_convert()
- allow to keep original-order of nodes in hier_import() via argument keep_order

sdcSpatial:

- protect_quadtree: stopping criterion + `max_zoom` bug-fixed