Center of Excellence on Statistical Disclosure Control

User support for and maintenance of SDC tools
The CoE on SDC was created in October 2014 for 4 years with 7 countries:
- Netherlands, Austria, Germany, Finland, France, Hungary and Slovenia.

It has five main objectives:
- **Methods**: Develop methodological solutions to SDC issues common in the whole European Statistical System (ESS) and for different statistical domains;
- **Tools**: Improve and develop SDC tools;
- **Guidelines**: Develop guidelines and requirements that are compatible with the current legal frameworks and practices;
- **Support**: Support and/or guide National Statistical Institutes (NSIs) to set up and adapt their infrastructure in order to comply with minimum requirements for protection of confidential data;
- **Training**: Provide training and courses on particular SDC issues.
Three projects launched so far:

- Public use files (PUFs) for ESS micro data
  - In 2015 → publication of PUFs for LFS *Labour Force Survey* and SILC *Statistics on Income and Living Conditions*

- User support for and maintenance of SDC tools
  - Apr. 2016 – Mar. 2018

- Harmonised protection of Census data
  - Sept. 2016 – Aug. 2017 → recommendations to help census experts tackle the issue of disclosure control for Census 2021
User support for and maintenance of SDC tools

- Partners of the project: Netherlands, Austria, Germany, France and Slovenia.
- Main aim is to ensure support, testing, maintenance and development of SDC software of interest to the NSIs in the ESS:
  - ARGUS (μ-ARGUS, τ-ARGUS)
  - R-packages (sdcMicro, sdcTable, SimPop)
Argus and R-packages are different UIs to integrate the underlying methods (modular, optimal, CTA, local suppression, PRAM ...). They coexist because they are in different environments:

- Integrating SDC in R is a strong feature for users familiar of R,
- whereas Argus is already in use in several NSIs and has been integrated in the statistical processes.

In these interfaces the underlying methods are sometimes similar but not exactly identical.

- Another aim of the project: identify such cases of parallelism and treat the important ones.
A first inventory of functionalities was completed with two questionnaires to identify the functionalities to be supported in the SGA.

- One on microdata, the other on tabulated data.
- Sent to the members of the Working Group on Methodology

Important core functionalities were defined with the questionnaires and sorted in three groups:

- A-functionality: each UI should provide it and the results should be identical with each UI
- B-functionality: can be supported by one UI (rare)
- C-functionality: less relevant, not to be supported but may remain in the UIs
User group (1)

- Informal: no official representatives of the different Member States.
  - Anybody using SDC tools can join by registering on https://joinup.ec.europa.eu/software/sdctools/home
  - A mailing list provides information on workshops, meetings, new releases.

- A first meeting with ~20 participants was held in Luxembourg on 2 December 2016. The developers presented the inventory of supported features.
The main goals of the user group are to:

- **test** the software. This is done indirectly by using the SDC tools in daily practice.
- **report** bugs. Users may encounter strange or unexpected behaviour and are encouraged to report those at the Help desk portal so that developers can assess the issue.
- **exchange** information and good practices between users. Making use of the Help desk portal, users can interact and provide answers to issues raised when they already found a work around.
- **suggest** improvements and desired new functionalities. A real production environment may yield different ideas of improvements than the developers environment. Users can suggest improvements and new features.
Help desk portal (1)

- Key aspect of the user group. Set up on github:
- Users can report bugs or anything strange they encounter while using software.
- Users can also suggest improvements and new features.
Help desk portal (2)

1. Detecting inconsistent status :enhancement, :feature, :tau-Argus
   #67 opened on 29 Jun by ppdewolf

2. SAS2ARGUS: provide example inputs :enhancement, :help-wanted
   #56 opened on 12 May by bernhard-da

3. High dimensional table gives strange error message :enhancement, :tau-Argus
   #52 opened on 11 Apr by ppdewolf

4. Skipping complete table :enhancement, :tau-Argus
   #51 opened on 11 Apr by ppdewolf

5. Do not deselected the cell after setting to protected/unsafe or setting its cost :enhancement, :feature
   #50 opened on 22 Mar by junosukan
- intended to be as backward compatible as possible

The code repositories for the different UIs and the underlying methods were put on Github at the following link: https://github.com/sdcTools.

The major part of the actual development is done by CBS (Argus) and Statistics Austria (R-packages).

Developers set up unit testing with a collection of automated tests defined to check the software at each release and guarantee stable versions.
Center of Excellence on Statistical Disclosure Control: User support for and maintenance of SDC tools

Thanks for your attention

Maël Buron
mael-luc.buron@insee.fr

French National Institute for Statistics and Economic Studies (Insee)

Insee