

Topic 3: Confidentiality of big data and special types of data

Peter-Paul de Wolf and Josep Domingo-Ferrer



Contents

- 16:10 Mirjana Boshnjak (*Stats Macedonia*)
Data confidentiality and statistical registers in the Macedonian statistical system
- 16:25 Said Khadraoui (*Secure Data Access Center*)
Tabular Data anonymization with Big Data tools. (Secondary Cell Suppression via Tabu search)
- 16:40 Peter-Paul de Wolf and Edwin de Jonge (*Statistics Netherlands*)
Location related risk and utility
- :-- Eduard Suñé, Cristina Rovira, Frederic Udina, Mireia Farré, Mònica Gasulla and Daniel Ibáñez (*Idescat*)
Statistical Disclosure control in visualising geocoded population data using quadrees
- 16:50 Questions to authors
- 17:05 General discussion and conclusions

Contents

Eduard Suñé, Cristina Rovira, Frederic Udina, Mireia Farré,
Mònica Gasulla and Daniel Ibáñez (Idescat)

Employees of Institut d'Estadística de Catalunya (Idescat)

Title:

Statistical Disclosure control in visualising geocoded population data using quadtrees

Summary:

Displaying geo-coded population data on a map by aggregating a layer of points into a layer of polygons, using a quadtree structure based on the European standard grid. Using a Monte Carlo simulation to estimate the errors in the calculation of the population when different quadtrees were used.

Contents

Mirjana Boshnjak (Stats Macedonia)

Head of sector for business statistics, agricultural statistics and statistics on environment

Title:

Data confidentiality and statistical registers in the Macedonian statistical system

Summary:

Overview of measures taken at the State Statistical Office to ensure confidentiality of statistical data in registers.

Contents

Said Khadraoui (Secure Data Access Center)

Data scientist at TeraLab/CASD

Title:

CASD-TERALAB, a secure remote access system to confidential big data: tabula data anonymization using big data tools

Summary:

Protection of high dimensional tables with cell suppression, using the Tabu-search algorithm

Contents

Peter-Paul de Wolf (Statistics Netherlands)

Senior methodologist at Statistics Netherlands

Title:

Location related risk and utility

Summary:

Defining disclosure risk and utility measures attached to locations
(geo-coded information displayed on a map)