CBS Internships

Arnout van Delden, Bart F.M. Bakker

13 December 2019
Innovation in Official Statistics

– Output
  - for central policy makers and policy evaluation
  - tailor-made publication for all kinds of institutions

– Data sources widened
  - surveys, admin data, unstructured data and sensor data
  - from single source to multisource statistics

– Challenge: innovative use of data while still
  - reliable estimates, CBS = quality stamp
  - transparent methods
Bèta producten in ontwikkeling

Sociale spanningen en emoties in de samenleving
gisteren 11:50

Innovatievormen bepalen met behulp van tekstanalyse
gisteren 11:49

CBS berekent toegevoegde waarde van de infrastructuur
gisteren 11:48

Hooikoortsindex
gisteren 11:47

Cybercrime achterhalen in aangiften
gisteren 11:46

Door schatting actueel beeld verhuizingen mogelijk
gisteren 11:46
Assume you aim to measure crime in the Netherlands

- What data sources?
  - survey on victims of crimes
  - police registration data
  - text mining

- What estimation problems?
  - memory effects
  - not everybody reports to police

Statistics is not about counting but about estimation
Projects at CBS

- How to estimate changes with text mining, in situation of concept drift?
  - I. How to semi-automatically construct an ontology network to stabilise the concepts?
  - II. How to retrain models, to obtain accurate change estimates?

- Anonymising input files (GDPR):
  - III. How can we reliably recognise names, locations and so on in descriptions of input files? We wish to replace them by generic labels like #name, #location and so on.
Furthermore, and contacts

- Requirements
  - Programming in Python and/or R
  - An admission interview is part of the procedure

- CBS provides
  - A stimulating environment with team of methodologists and students
  - Monthly internship allowance

- Contacts
  - Visit the website www.cbs.nl
  - To learn about the projects: a.vandelden@cbs.nl