Mobile phone data for daytime populations

Peter Struijs

International Conference on Big Data for Official Statistics, Beijing, China, 28 – 30 October 2014
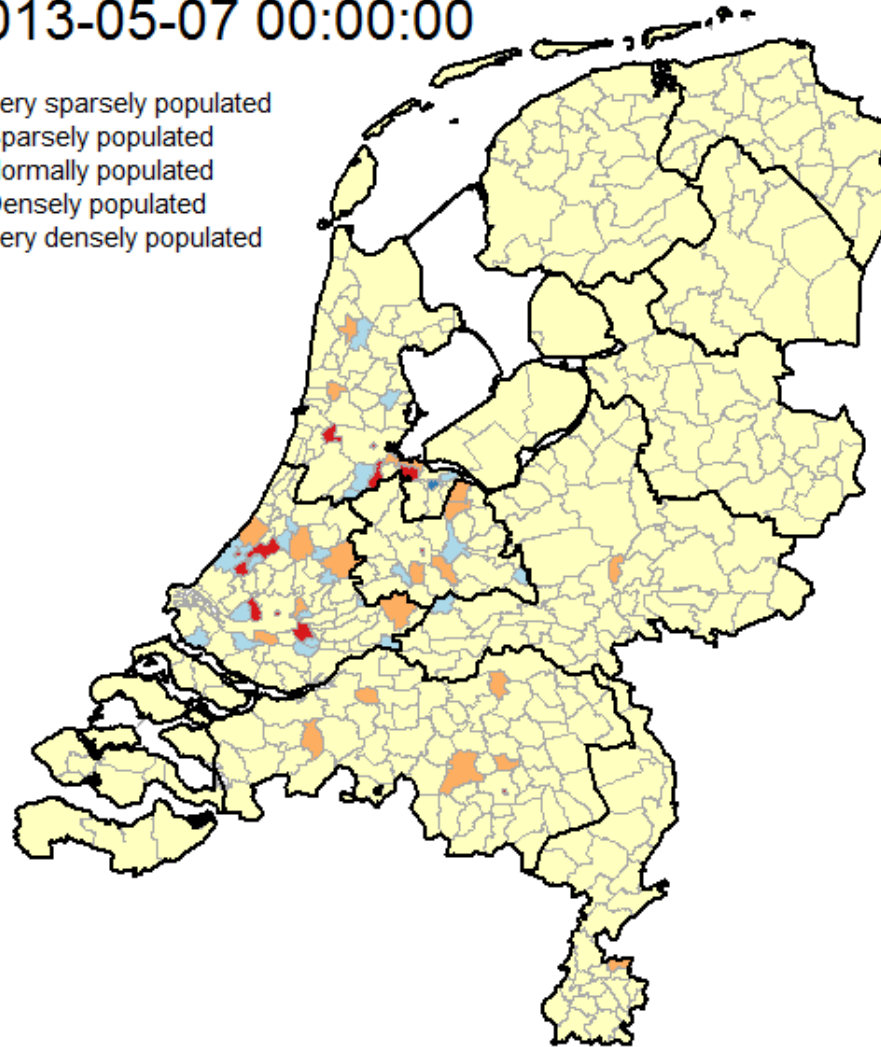
Statistics Netherlands

Topics covered

➢ Why use mobile phone data for official statistics?

➢ Characteristics of the data source

➢ Issues when using mobile phone data

➢ Solutions to issues

➢ Results

Daytime population based on mobile phone data

2013-05-07 00:00:00

- Very sparsely populated
- Sparsely populated
- Normally populated
- Densely populated
- Very densely populated

Possible uses of mobile phone data

- ➤ Daytime population statistics

- ➤ Mobility statistics

- ➤ Tourism statistics

- ➤ Other uses

Mobile phone activity as a data source

- – Nearly every person in the Netherlands has a mobile phone
  - - Usually on them
  - - Almost always switched on
  - - Many people are very active during the day

- – There is a grid of antennas with good coverage

- – Data of a single mobile company was used
  - - Hourly aggregates per area
  - - Threshold of 15 events

5

Issues when using mobile phone data

➢ Privacy

➢ Data acquisition

➢ Methodology

- Representativeness

- Selectivity

- Quality

➢ Other issues

- Infrastructure

- Skills needed

Solutions

➤ Agreement with data provider to provide only aggregates and apply a threshold

➤ Data provider performed analysis of mobile phone ownership characteristics

➤ A large number of analyses were made, with the regular population registration data as a reference

➤ A number of assumptions had to be made

7

Lessons learned

➢ Use only aggregate data if possible

➢ Invest in partnerships with data providers

➢ Invest in methodological research and play with the data to get a grip on quality