

Big Data and Official Statistics: Challenges and Applications at Statistics Netherlands

Piet J.H. Daas^{1,2}

¹Department of Data Services, Research and Innovation, Statistics Netherlands, the Netherlands, ²Department of Mathematics and Computer Science, Eindhoven University of Technology, the Netherlands.

Abstract

The use and application of Big Data in official statistics has made considerable progress at Statistics Netherlands. The major contributors are the increased attention for Big Data in the methodological research program, in the creation of experimental statistics and in its use for regular statistics production. To stimulate this the Center for Big Data Statistics was setup in 2016. The most important research challenges identified are:

- 1. Concept: What (derived) concept is measured in Big Data?*
- 2. Population: What part of the target population is included in Big Data?*
- 3. Methods: What new methods (or new ways of thinking) are needed?*
- 4. Infra: What infrastructural requirements are needed?*

The infrastructural (IT) challenge is ignored here. The fact that there is a steady increase in the application of Big Data at the office indicates the need (and progress made) in the study of the research challenges identified. The statistics that make use of Big Data and are either in production or for which an implementation process has started at Statistics Netherlands are:

- 1. Using scanner data and scraped prices for the Consumer Price Index*
- 2. Using road sensor data for Traffic Intensity statistics*
- 3. Using website texts for Online Platform Economy statistics*
- 4. Using social media for the Social Tension indicator*
- 5. Using texts of online job advertisement for Vacancy statistics*
- 6. Using solar panel output and weather data for Solar Energy production*

The presentation will discuss the research challenges and how this has affected the use of Big Data for official statistics at the office.

Keywords: *Key topics; Applications; Challenges; Methodology.*
