

BIG DATA METHODS AND TECHNIQUES

Presenter(s): e-mail(s): Date:	Piet Daas & Marco Puts, Statistics Netherlands pjh.daas@cbs.nl, m.puts@cbs.nl 21 March 2018
EMOS learning outcome addressed	Statistical methods
Webinar aims	The webinar aims to show statisticians how to deal with big data in real world statistical applications.
Webinar learning outcomes	To be aware of the need for a paradigm shift when using big data. To be familiar with methods that enable extraction of information from big data in a reliable way. To understand the importance of the need for good quality data and well thought of checks and controls.
Webinar content	In the webinar, the current state of art of using big data for official statistics is discussed. These are illustrated by walking through a big data based production process. The observations are generalised to other big data based applications.
Difficulty level	Advanced
Prerequisites for the webinar	The webinar will build on the content presented in: 2017 EMOS Webinars: Big Data I and II
Further readings and resources	Daas, P.J.H., Puts, M.J.H. (2014). Big data as a Source of Statistical Information. The Survey Statistician 69, 22-31. Singh, D., Reddy, C.K. (2014). A survey on platforms for big data analytics. Journal of Big Data, 1-8. Daas, P.J.H., Puts, M.J., Buelens, B., van den Hurk, P.A.M. (2015). Big Data and Official Statistics. Journal of Official Statistics 31(2), 249-262. Daas, P.J.H., Burger, J., Quan, L., ten Bosch, O., Puts, M. (2016). Profiling of Twitter Users: a big data selectivity study. Discussion paper 201606, Statistics Netherlands, The Hague/Heerlen, The Netherlands. Puts, M., Daas, P., de Waal, T. (2017). Finding Errors in Big Data. In: The Best Writing on Mathematics 2016, Princeton, USA. (Pitici, M., ed.), pp. 291-299, Princeton University Press, USA.