Mobile Network Data / 6-month Datathon project

ONS and the recently launched [ONS Data Science Campus](#) ONS is launching an accelerated 6 months project to build on early UK progress made with mobile network data (primarily understanding commuter patterns) and better engage the UK's mobile network operators (MNO's) to use these data to better inform national policy making. We are open to and welcome input and ideas from the international Stats community but will be focussed on the Challenge themes below.

Toward the end of the six months project ONS will run a 'Datathon' style event to bring together public and private sector teams of data scientists and engineers applying collaborative intensive effort in a fixed duration research '*Datathon'* style event (likely 2-3 days in Sep - Oct 2017). Ideally the recently launched Data Science Campus will host this event. The format of the event will be structured to focus the very best data science methods and expertise on pre-prepared datasets against carefully curated policy 'Challenges.' These Challenges will be curated and run in an approach similar to the Nesta and ODI Open Data Challenges series.[1] This project will balance the strategic need for ONS to improve timeliness and authority of existing outputs with the imperative of driving down data collection costs and the need to formulate new statistics. The initial six month project goal will focus on working together to apply mobile network data to tackle a shortlist of prequalified policy questions as well as to lay a foundation to better secure meaningful longer-term data access. Currently the policy questions we are exploring span two primary themes:

1.0   Population and migration

Daytime populations (DTP) by location are reported by CBS Netherlands to be easier to produce. Such an output also underpins other applications such as commuting so consideration to the overlapping utility of this use case should be given. We welcome input from ESSnet and CBS Netherlands.

1.1   Better understanding 'workplace zones'
Better understanding population movements and density in the urban/built environment and in particular in business districts is a sub category we are exploring as wider interest in the UK statistical community for developing this has been received.

Population by time of day can be also be reported by applying methods to estimate population density from call or subscriber volumes using a variety of data formats from aggregated and individual level call detail records (CDR). Policy questions include:
- How might mobile network data provide an alternative approach to the Annual Population Survey and or augment/improve it?
- Can ONS access network cell counts over time to understand better the movement of people over time and in different locations?
- Can these data support/augment or replace insights developed from the UK/Passenger surveys?

---

[1] http://www.nesta.org.uk/project/open-data-challenge-series

2.0   Developing a 'national cell count' reference dataset

One of the primary challenges in working effectively with these data is that each operator collects, processes and applies methods differently to these data. Another challenge is to work towards a defensible sample size across a national geography, and this could mean ONS will require access to standardised data from multiple operators. So these methods and technological challenges are amplified. Our approach would lay a foundation to overcome these challenges, define a standard for collecting, processing and potentially combining from multiple operators and will be a great way to better build confidence in working with these data in a way which supports aggregate statistical work without undermining MNO's commercial interests. A quick win goal could be to have operators collaborate to develop a 'national cell count' dataset. Such an approach could not only benefit ONS but could also support the UK and international statistical community to better utilise these data.

Data privacy, confidentiality and communications

Public perception is a critical consideration in this work. A communications led approach to ONS, operator and external stakeholders will be taken from the start and we shall resource independent external data privacy expertise to ensure a rigorous approach to personal data privacy. It is generally accepted that accessing anonymised individual level data versus aggregated counts has the potential to support broader and more in depth research (e.g. a wider range of mobility applications) but the trade off is greater ethical and privacy implications. This project will further explore these challenges and look at solutions to overcoming them (e.g. taking methodology/algorithms *to* the data instead of accessing the data could be one such approach).

One of the primary challenges in working effectively with these data is that each operator collects, processes and applies methods differently to these data. Another challenge is to work towards a defensible sample size, and that ONS will require access to standardised data from multiple operators – so these methods and technology challenges are amplified. This approach will lay the foundation to overcome the challenges, define a standard for collecting, processing and potentially combining data from multiple operators.

6 month goal

This initial 6 month engagement is intended to provide a platform upon which MNOs mobile operators and their key commercial partners can understand, support and help implement a new sustainable model for commercial Data as a Service and that ONS is able to better compare and benchmark industry expertise and capability.