

---

**HLG-MOS Machine Learning Project: May F2F Sprint Update**

Eric Deeben; Oliver Mahoney (Office for National Statistics, United Kingdom)

Eric.Deeben@ons.gov.uk; Oliver.Mahoney@statistics.gov.uk

***Abstract***

Between the 13 to 15 May 2019 the UK's Office for National Statistics (ONS) will be hosting the first face-to-face sprint for the UNECE HLG-MOS Machine Learning Project at both its London and Newport offices. Group members from various countries are expected to attend. The ONS would like to present the outputs of the sprint at the ModernStats World Workshop 2019, hoping that it will give the CES insight into projects progress and direction. During the sprint the group currently expects to cover the following topics:

**I. Address outstanding questions regarding scope.**

This was initially defined by members completing questions and discussing at the first virtual sprint on the 4 April 2019.

**II. Determine a high-level plan for the (WP1) Pilot Studies including themes.**

Groups members have been asked to submit proposals for the machine learning pilots. These proposals are either existing domestic projects which are exploring ML or common surveys which have elements which are suitable for ML application. With a full list of submitted proposals at hand, the sprint will determine three key pilot study themes which will be conducted. A lead and corresponding members will be assigned under each theme. We proposed for WP 1 the have four common groups objectives and how the work packages seek to achieve them.

Pilot studies:

- Classification and coding on traditional sources of data
- Edit and imputation on traditional sources of data
- Imagery
- Sentiment and Web data

**III. Create an approach method for the pilots within WP1**

Outputs from each pilot study and any other business will be presented to delegates at the ModernStats World Workshop.

***Keywords***

Machine Learning; Pilot Study; Satellite; Imagery; Classification; Coding; Edit; Imputation; Web Scrape; Sentiment; Data