

FCSM Interest Group on Computational Statistics and the Production of Official Statistics Proposal

October 2017

1. Objective

The purpose of this interest group is to share information on computational statistics and the production of official statistics across federal agencies, to identify areas of common research, to coordinate collaborative research projects in those areas, and to disseminate information related to computational statistical methods and the implementation of these methods to produce official statistics.

2. Background

Computing has assisted in the production of official statistics since its use in tabulating the 1890 Census¹. In the subsequent 120 years, computing needs have grown considerably; impacting all facets of production, from data collection to dissemination. Given the importance of computing and the challenges of Big Data, it is important that agencies have a forum to collaborate and learn from each other. An example of topics germane to this group are:

- New and emerging computational methods in statistics such as Hierarchical Bayesian model integration through Gaussian fields and parallel implementations of machine learning methods.
- Training statistical and information technology staff on emerging programming languages such as R and python.
- Sharing best practices for implementation and maintenance of statistical software.
- Identifying opportunities for collaboration on statistical computing projects.
- Creating a dialog between information technology staff and statisticians.
- Creating a dialog between information technology staff at different agencies.

The proposed interest group is a product of conversations between Wendy Martinez (BLS) and Jonathan Lisic (formerly of NASS) about challenges faced in the

¹ 1890 Overview, U.S. Census:

https://www.census.gov/history/www/through_the_decades/overview/1890.html

implementation of statistical methods. Further support and identification of interagency interest was identified through meetings of the Interagency R-User's Group in 2016 and 2017.

It was determined that this interest group would provide real value for the participating agencies through the sharing of information on ways to share knowledge and resources on statistical computing and the production of official statistics.

3. Scope and approach

The main focus of this group will be on computational statistics and the production of official statistics. In particular, this includes the challenges listed in section 2, but are not limited to only those. Any topics related to computational statistical methods or the implementation of statistical methods to produce official statistics are in-scope.

The group will hold monthly meetings at the Bureau of Labor Statistics through December 2018. This meeting schedule will allow the group to determine the most promising areas to explore, to identify the list of activities and to plan the deliverables. We expect to move to bi-monthly or quarterly meetings after December 2018.

4. Proposed Deliverables and Dates

The team will provide the following deliverables to the sponsor in 2018:

- Proposed list of activities for the group – July 2018
- A report that could include a description of member agencies' current research efforts and challenges – December 2018
- One workshop in the DC area – Fall 2018

We expect this group to continue beyond 2018, if there continues to be interest and value-added to the agencies.

5. Sponsor

The Federal Committee on Statistical Methodology (FCSM)

6. Communications

- The team will provide regular informal status reports to the sponsor.
- The team will use a Wiki or other means, e.g. collaboration site through FCSM
- As needed, the group will plan presentations to other FCSM or federal groups.

7. Group members

The members of this group must be federal civil servants. However, contractors and interested parties can provide briefings and participate in specific meetings and workshops.

The members do not have to be managers or statisticians. Members should be those who are familiar with information technology and how it is used in implementing statistical methods.

Agencies can have more than one member in the group.

Wendy Martinez (BLS) will serve as liaison with FCSM.