

# Master of Data Science - University of British Columbia Okanagan

Minutes for April 29, 2020 - 11:00 AM - 12:00 PM PST

**Present:** Bruno St-Aubin (Statistics Canada), Marian Radulescu (Statistics Canada), Sofia Bahmutsky, Ngan Lyle, Kaitlyn Hobbs (Minutes), Shreeram Murali

## Agenda

### *Discuss Timeline*

- 1 — Share MDS expectations in regard to timing as well as our team's discussed working hours (PST)
- 2 — Determine whether to organize a shared calendar and on which applications
- 3 — Determine frequency of regular meetings

### *Discuss Project Data and Proposals*

- 1 — Clarify if we are able to upload data to Github
- 2 — Inquire if we are able to integrate other data sources, such as economic data, depending on focus

## Meeting Notes

1. Statistics Canada representatives introductions
2. **Reviewed skills:** SQL, python, R, C++ (Shreeram), tensorflow, Julia optimization, JavaScript (Shreeram), GIS (Sofia)
3. Proposals from Statistics Canada:
  - (a) **Open database of buildings - LODE** - enrich with more data attributes. In collaboration with Microsoft, only shows geometry of buildings. **Purpose:** use machine learning and other databases (such as Open StreetMap) to assign addresses, building heights, building types. **Drawbacks:** database structure to relate with other open databases (open addresses, recreational facilities, open database of health facilities) for on-the-fly querying, structure of addresses, assess quality of database and linkages, dealing with missing data.
  - (b) **Visualizations of Statistics Canada datasets** - interactive product that showcases data (potentially Covid-19). Data includes case counts and provincial data. Proximity database may provide insight for long-term care facilities. Attainment of time series data may be possible.

### *Questions and Answers —*

**Question from Ngan:** Where are your data sources?

**Answer from Statistics Canada:** Other open data sources on provincial level - standardized and integrated, webscrape and geocoding of dispersed data (e.g. for proximity measures).

**Question from Ngan:** In terms of application piece are you thinking more HTML?

**Answer from Statistics Canada:** Example of application of international cases across time (dashboard). The simpler the better.

**Question from Ngan:** In terms of programming, what language did you use?

**Answer from Statistics Canada:** HTML JavaScript, observable HQ D3 - library for development.

**Question from Kaitlyn:** Is there a database for hospital equipment need?

**Answer from Statistics Canada:** Not on the hospital level for privacy reasons. It may be interesting to consider population density for the regions but we have to be careful not to use predictive modelling.

**Question from Ngan:** It is interesting to look at mobility of individuals, is that something that might be a privacy concern or is there data available for that?

**Answer from Statistics Canada:** It may be worth looking at proximity to resources.

**Question from Statistics Canada:** Do you need to produce any reports or final products?

**Answer from Shreeram:** It may be worth looking at proximity to resources.

**Question from Statistics Canada:** Do you have flexibility in what you work on?

**Answer from Ngan and Sofia:** Our goal is to solve a client's problem, you being the client, so as long as you have flexibility then the program does not limit us. The only set thing is a report and presentation.

**Question from Statistics Canada:** What are your thoughts, are you expecting more or less?

**Answer from Shreeram, Sofia, Kaitlyn, Ngan:** Shreeram - was hoping for something more analytical but is happy with both projects. Sofia, Kaitlyn - prefer second project for timely relevance and personal interest in content. Ngan - both projects are interesting, the first being more technical and the second having a lot of unknowns.

Wednesdays - 11:00 AM PST (Kelowna time) using Webex.

## Action Items

1. Share Trello team boards with Marian and Bruno
2. Discuss as team which project we would like to work on and begin research and proposal

**Next Meeting:** Wednesday, May 6, at 11:00 AM PST with Bruno and Marian via Webex