

# **Visualization of the Canada Emergency Wage Subsidy and Feasibility Study of API use for the Prediction of Economic Activity and Recovery**

## **Proposal and Statement of Work**

### **1. Executive Summary**

This proposed project is in collaboration with Statistics Canada and the University of British Columbia's (UBC) Master of Data Science Program.

The first half of the project, Visualization of the Canada Emergency Wage Subsidy will involve the development and deployment of an interactive visualization dashboard for the Statistics Canada website, as well as an exploratory report detailing some interesting findings in the data. The Canada Emergency Wage Subsidy (CEWS) is a wage subsidy program implemented by the Government of Canada with the intent of helping employers to maintain staff and rehire employees laid off due to the Covid-19 global pandemic. The CEWS program alleviates the strain on businesses caused by reduced profitability during the pandemic, and stimulates economic activity while positioning businesses for a smooth reopening following the end of the crisis. A database has been made public on the Statistics Canada website which details all of the subsidies released through this program. The interactive dashboard will allow users of the Statistics Canada website to explore how this program has affected their regions of interest, at the level of province, census agglomeration, or census subdivision. They will be able to view the subsidies given to each industry within their region of choice. The report will primarily focus on how rural and urban communities have been affected differently, with visualizations to demonstrate some of the more interesting findings.

The second half of the project would be a feasibility study, with the intent of examining the potential of available application programming interfaces (APIs) as real-time predictors of economic recovery and activity, and for incorporating quality of life measurements into economic decision-making. In this study, we would examine in detail a set of promising APIs and the potential data that could be gained through them. Through evaluations of the cost, efficiency, and usefulness of data, we would then make recommendations to Statistics Canada on which, if any, APIs should be used. We would also develop some preliminary functionality for one of the studied APIs.

This project would be conducted in the six weeks between May 10 2021 and June 22 2021, and would come at an estimated cost of \$0 to Statistics Canada. The primary development would be conducted by BoHan Gao, Eric Baxter, and Vicens Paneque of UBC. The project would be directly managed by Scott P. Bacon of Statistics Canada. Other involved parties would be Scott Fazackerley and Irene Vrbik of UBC, and Haaris Jafri, Ashe Monroe, Nick Newstead, and Mahamat Hamit-Hagggar of Statistics Canada.

## **2. Background Research**

### **2.1. Overview of Canada's COVID-19 Economic Response Plan**

Through Canada's COVID-19 Economic Response Plan (the Plan), the government has taken strong, immediate, and effective action to protect Canadians from the impacts of the pandemic.

On the advice of public health officials, jurisdictions across Canada took measures to protect the health and safety of Canadians through physical distancing, temporary restrictions on non-essential businesses, and school closures. These measures had a significant impact on the daily lives of Canadians and their families and led to job losses and reduced employment income.

The Plan supports workplaces of all sizes across sectors, protecting jobs and helping Canadians return to work from coast to coast. The priority was getting support out to Canadians as quickly as possible, which the government did through a mix of existing systems and the rapid design and delivery of brand-new programs.

The Plan is focused on three broad areas of support:

#### **1. Health Care System and Safety of Canadians**

Includes provinces and territories, industry and research organizations, international partners, and Canadians abroad.

#### **2. Support to Individuals**

Includes workers, students, families, seniors, and vulnerable groups.

#### **3. Support to Business**

Includes small, medium, and large enterprises, primarily in heavily affected sectors.

Protecting Canadian jobs has been a priority throughout the COVID-19 pandemic. As businesses have reduced productivity and slowed down in accordance with public health guidance, finding solutions to preserve the connection between employees and employers was critical to Canada's continued economic resilience.

The government has provided support for businesses of all sizes, across sectors, from coast to coast to coast. Through a variety of programs, tailored to the varying needs of Canada's diverse economy, the federal government is making sure that businesses can get flexible support, so that Canada's economy is stable and so that Canadians' jobs are protected [1].

The decisive and substantial support provided by the government helped prevent further damage to the economy by replacing lost income and avoiding even higher unemployment rates. Programs like the Canada Emergency Wage Subsidy have helped keep money in the pockets of Canadians so they can pay for housing, groceries, and medicine, while also helping businesses survive through the pandemic [2].

### **2.2. Canada Emergency Wage Subsidy (CEWS)**

One of the programs designed to support businesses and individuals through the pandemic is the Canada Emergency Wage Subsidy – a wage subsidy of 75 per cent for qualifying employers, up to \$847 per week per employee, which began March 15, 2020.

CEWS provides financial support to employers to help cover the costs of maintaining and rehiring their workers through periods of reduced profitability. To date, CEWS has ensured that millions of Canadian workers continue to receive paychecks from their employer, and are ready to return to work as soon as public health measures allow. By protecting the employee-employer relationship, this program also ensures supply chains can rebound from the crisis in a strong position [1]. CEWS has also played a critical role in bolstering consumer spending [3]. By providing affected Canadians with a source of income, CEWS has given many Canadians the confidence to spend money, which would otherwise be kept in savings.

With some exceptions, employers of all sizes and across all sectors of the economy are eligible for CEWS benefits. When first launched, the program was intended to last 12 weeks, beginning March 15, 2020. To receive the subsidy, eligible employers had to have seen a drop of at least 15 percent of their revenue in March 2020, or 30 percent in April or May [1].

The government has since extended the wage subsidy to June 30, 2021, for a total of 64 weeks. The timeline is broken down into sixteen 4-week periods [4].

- From March 15th, 2020 to July 4th, 2020, eligible employers that saw a drop of at least 15% of their qualifying revenue in March 2020 and 30% for the following months of April, May and June, when compared to their qualifying revenue for the same period in 2019, qualify for the wage subsidy.
- From July 5th, 2020 to December 19th, 2020, the wage subsidy was available for all eligible employers that experienced a decline in revenue for a claim period. A larger subsidy was available for those employers that were most adversely affected by the COVID-19 crisis.
- From December 20th, 2020 to March 13th, 2021 the wage subsidy was raised to 75% for eligible employers [5].
- From March 13, 2021 to June 5, 2021, the rules related to the wage subsidy will remain essentially the same as the previous 12 weeks, with the addition of an elective alternative baseline remuneration period for this 4-week period [4].

The wage subsidy program is currently set to expire in June 2021, but it is proposed in Budget 2021 to extend the wage subsidy until September 25, 2021 [6]. If this extension comes to pass, there will be a gradual decrease in the subsidy rate to ensure a smooth phase-out of the program as vaccinations are completed and the economy reopens.

### **2.3. API Feasibility Study**

In Canada, and globally, data and the digital economy are driving innovation, along with economic and social opportunity. Statistics Canada's data strategy is

embedded in capitalizing on this data revolution and leveraging the explosion in the volume of data and data analytics technologies by actively researching, developing, and implementing solutions to provide Canadians and decision-makers with increased access to high-quality, relevant and timely information on issues affecting the society.

Statistics Canada is currently modernizing its statistical programs to respond to a rapidly changing and increasingly complex economy and society, a proliferation of data and providers, and increased user expectations for real-time and micro-detailed data. This includes creating strategic partnerships with other organizations and researching and discovering data inputs that can be used by statistical programs, such as administrative data, open data, found data, commercial data, crowdsourced data and web-scraped data, while respecting privacy and maintaining public trust.

Social media platforms, sharing-economy companies and crowdsourced data through APIs enable promising new approaches to measuring economic activity and analyzing economic behavior at high frequency and in real time using information independent from standard survey and administrative sources.

APIs are tools which allow individuals and businesses to access and use data collected by other applications and companies. Many data companies have APIs available for public use, often with an associated cost. The data accessible from these APIs is often real-time, and learning how to harness these programs for Statistics Canada would allow for more efficient policy and decision making. APIs can be used to learn about the spending habits and quality of life of Canadians as the country recovers from the pandemic. APIs could allow for the study at far more granular levels than traditional measures allow, by allowing the study of individual neighborhoods, streets, or businesses.

### **3. Data Sources**

The following open data sources will be used:

#### **3.1. Canada Emergency Wage Subsidy Regional and Community-level Database**

##### **3.1.1. Structure**

The Canada Emergency Wage Subsidy Regional and Community-level Database (the database) is a custom dataset constructed with CEWS microdata and other administrative data sources available at Statistics Canada. This dataset contains information on every subsidy given out through the CEWS program, aggregated to different levels of geography and industry. Geographic aggregation is available at the levels of country, province, census metropolitan area/ census agglomeration (CMA/CA), census subdivision (CSD), and urban/rural splits by province/ canadawide. These geographic classifications were done according to the Standard Geographic Classification 2016 (SGC). This system provides a hierarchical code to each region, with provinces receiving a 2-digit code, census agglomerations

receiving a 5-digit code, and so on. Industry aggregation is available at the level of all industries, as well as by specific industry according to the 2-or-3-digit North American Industry Classification System (NAICS). 2-digit NIACS are general industries, with 3-digit NIACS more specific subdivisions of these. As such, the most general information available is the total subsidies given in Canada, and the most specific information available is the subsidies given in a particular CSD, within a particular 3-digit NIACS industry. Each row contains data for one claim period.

In each row, data is given for the total CEWS supported employees, total supported businesses, and total CEWS claim dollar amounts. An example of the dataset is shown in Appendix A.

### **3.1.2. Metadata**

The database contains data for the first 8 four-week claim periods of the CEWS program, between March 15 and October 24 2020. The file contains all approved CEWS claims as of January 4 2021.

To generate this regional and community-level database, CEWS microdata were first aggregated to the business enterprise level, and subsequently allocated to the business locations of the enterprise [8]. For multi-location businesses, allocation of CEWS data was done to individual locations to more accurately reflect the use of CEWS by these businesses.

Given the high level of granularity provided in the database, data suppression was implemented to preserve confidentiality. Data suppression occurred due to one of two reasons:

- Primary suppression: There are too few businesses contributing to the value of a cell, or one business dominates the value of the cell.
- Secondary suppression: Another cell is suppressed to protect the value of other primary suppressed data during the publication of a higher-level aggregate. For example, if only two cells make up the higher level aggregate and one is primarily suppressed, its value could be found using the value of the aggregate and the value of the other cell.

There is a planned update to extend the data for claim periods up until the end of the CEWS program.

## **3.2. Public Data Tables, Micro-data, and Data Visualizations**

Where possible, additional data may be sourced from Statistics Canada, in order to accent our analysis with other variables. For example, a breakdown of GDP by geographic region would allow us to compare total subsidies to GDP. We would also like to source data on the number of workers within industries and geographic regions to allow for analysis on a per-capita basis. This data will be merged with the CEWS database, giving us additional columns of information to work with.

### **3.2. Data Hosting and Availability**

All of the data we plan to use is publicly available on the StatCan website, and as such there is no need for us to take measures to ensure the privacy of our data.

Furthermore, the dataset is small enough to be accessed locally. We plan to upload our wrangled dataset to our GitHub repository. When the second version of the CEWS data is made available, we will run the revised data through our wrangling code and store it on GitHub as well.

## **4. Objectives**

### **4.1. CEWS Report**

The primary objective of the CEWS report is to explore how the CEWS program has affected rural and urban communities differently. Some interesting questions to be answered are:

- What proportion of subsidy money went to urban vs. rural businesses?
- Which industries were most affected in each?
- Are there significant differences between provinces in how urban and rural communities were affected?

After answering these general questions, the goal is to shift the focus of the report on to rural communities, and explore CEWS at the level of census subdivision. The goal in this part of the report will be to examine the rural communities that were most affected by the pandemic, and determine any trends or relationships.

- Are rural communities disproportionately affected in particular provinces?
- Are rural communities which are heavily dependent on specific industries more subsidized than others?
- What is the most subsidized rural community in each province, in general and per capita?

The report will finish with an examination of some census subdivisions which received particularly large subsidies in an industry which was in general not heavily subsidized, with the goal of exploring the breadth of different communities which required significant assistance during the pandemic. For example, if there exists a CSD which received its largest subsidy in the single industry which was least subsidized throughout Canada, what caused that particular region to be affected so differently than the country as a whole?

### **4.2. CEWS Dashboard**

The primary purpose of the dashboard is to allow users to explore the effects of the CEWS program on whichever region or industry interests them. The main page of the dashboard will be an interactive map of Canada, with a heatmap cover

The proposed project would be written with the intent to analyze CEWS data and identify trends across a number of dimensions. The authors would study the effects of geographic location, industry, and week-periods of claims on the extent and effectiveness of the CEWS program. The primary focus would be on determining the CEWS outcomes in avoiding layoffs, rehiring employees, and creating new jobs. However, as there would be a substantial exploratory aspect to this project, a number of other insights may be drawn as well.

The evaluation of a wage subsidy program can be divided into the microeconomic evaluation and macroeconomic evaluation. The microeconomic evaluation would focus on the effects of the program on employment and/or earnings of the individual participants. The macroeconomic evaluation would examine the effect on the economy as a whole, such as unemployment rates and total income in the economy [7].

The CEWS analysis component of the project would include two parts. The first would be an interactive visualization dashboard, to be hosted on the Statistics Canada website, which would allow users to explore the CEWS dataset along any dimension or subset they are interested in. The second part would be a report detailing any findings the authors make while exploring and visualizing the CEWS data. This report would be intended for a general audience, with sections of the report included in articles for the Statistics Canada website.

### **4.3. API Study**

The second half of the project would seek to research the feasibility of using various APIs to study economic activity and recovery in real-time, with more granularity than traditional metrics allow. The authors intend to write a report detailing the potential effectiveness of various promising APIs toward this goal, along with the potential costs associated with their use. The end goal is to pave the way for the development of tools which allow more efficient and granular measures of economic activity, recovery, and growth. This would aid the government in quickly and efficiently allocating financial assistance in the event of future pandemic or similar economic disaster.

This component of the project would assess the use of social media platforms, sharing-economy companies and/or crowdsourced data through APIs to create a real-time business indicator for economic recovery in the context of COVID-19.

Social media represents an excellent opportunity for the construction of timely socio-economic indicators. The Twitter platform, in particular, facilitates an online sphere where users can publicly discuss on economical and political issues, attracting politicians, journalists, strategists, and citizens. The Twitter API includes a number of endpoints and features designed to help understand the current or historical situation, discover insights, listen for events, and more that could be used in developing a real-time business indicator for economic-recovery.

Depending on the type of data extracted, numerical analysis, or text-based sentiment analysis and content extraction could be used in order to parse meaningful data from the API output.

Issues of data access, such as interoperability, ethics, privacy, and governance, as well as methodological issues would be analyzed to make sure that we can make sense of all these data in a meaningful way.

Statistical and machine learning methods may be applied in the development of the real-time business indicator.

## **5. Methods**

### **5.1. Summary of Approach**

#### **5.1.1. Team Assignments**

Eric will be in charge of all email communication with the client, and will work with the team and the client to set up meeting times. He will be in charge of data wrangling and exploration, with a primary focus on the CEWS report. He will look at the data through the lens of geography. When this is finished, he will move onto the dashboard, with a focus on the development of functionality to compare different geographic regions. During the API component, he plans to examine the use of Reddit by examining posts on subreddits dedicated to particular regions and industries in Canada.

Vicens will be the primary developer of weekly presentations. He will work with Eric on data exploration and the CEWS report, with a focus on the industry dimension. Once the report is complete, Vicens will begin preliminary exploration of APIs. He will focus on the use of the Twitter API for sentiment analysis to predict economic activity and recovery in specific regions.

Ben will spend the first few weeks entirely focused on the dashboard. He will develop the map functionality using ArcGIS. During the API component, his focus will be on the use of rideshare and food delivery apps such as Uber, Lyft, Skipthedishes, and Doordash as indicators of economic recovery within specific regions.

#### **5.1.2. Tools and Technology**

Most collaborative written work will be done using Google Docs, so all team members can work and contribute in real-time. Coding and numerical analysis will be done in Python, in the form of Jupyter Notebooks. The visualization dashboard will be built using Microsoft PowerBI, to accommodate the preferences of the StatCan team. Virtual meetings and conferences will be done using Microsoft Teams with Statistics Canada and Zoom for team and school meetings. All work will be version controlled and shared through Github. Hours worked will be tracked manually or through the use of Toggl, depending on the preferences of each individual.

### **5.2. Task Breakdown and Time Estimates**

The first week of work will primarily be dedicated to the development of this project proposal. Weeks 2 through 4 will be focused on the development of the CEWS dashboard and report, with one team member spending some time on the initial development of the API study. Weeks 5 and 6 will be primarily focused on the API feasibility study, although depending on progress a team member may use this time to continue work on the dashboard. The final week will be entirely focused on the creation of the final report and presentation for the DATA599 course. A tentative schedule can be seen in Appendix B

### **5.3. Project Deliverables**



At the end of this project we expect to have produced:

- An interactive visualization dashboard for the CEWS database
- A report detailing our analyses of the CEWS data
- A feasibility study for the use of APIs to predict economic activity, and potentially some proof-of-concept code
- A final report and presentation detailing all aspects of the project, as a final deliverable for the DATA599 course.

## **6. Deliverables and Schedule/Timeline**

A schedule detailing important project dates and milestones can be found in Appendix C.

## **7. Risk Assessment**

### **7.1. Confidentiality Risk**

While all of version 1 CEWS data is publicly available on the Statistics Canada website, the proposal to include version 2 data in the analysis before it is made public necessitates a confidentiality agreement. All involved parties have signed standard non-disclosure agreements (NDAs), which prohibit the disclosure of privileged information under threat of academic or legal consequences. Under NDA terms, any and all confidential information must be kept confidential by project members until published or made available to the general public. Confidential data will be released to project members on a need-to-know basis. Furthermore, it will be ensured that all privileged information and data is shared over secure communication channels in order to negate the risk of interception by malicious third parties.

### **7.2. Project Risk**

As this project has been given a hard deadline of June 22 2021, there is no risk of significant delays in the final product. Furthermore, as the work will be conducted in an academic setting by the MDS team, there is little risk in the project going over budget.

In the situation that the project requires more time than expected, the scope of the feasibility study will be reduced, and resources will be redistributed in order to ensure a high quality, finished visualization dashboard and analysis paper are completed to an acceptable quality standard. As the feasibility study is very open-ended and of limitless depth, if the project is completed with time to spare, all additional resources will be devoted to further work on this aspect.

The complete dashboard and analysis are dependent on the release of a second set of CEWS data. As such, there is some risk of delay to the release of this data holding back production. In order to counter this, all development of the dashboard will be done in an algorithmic way, such that upon release the dashboard can be easily and immediately updated to include the new data. If there are severe delays

to the data, it is possible that the analysis paper will need to be written with only the earlier data in mind. However, if the new data is released early enough, all efforts will be made to update the analysis to include this data.

### **7.3. Other Risks**

Due to the prevalence of suppressed data in the CEWS database, it is possible that any analysis may be misleading if there is any systematic method to data suppression. Care will be taken to verify with Statistics Canada that any insights gained from the suppressed dataset will generalize to the full data. It will also be made clear both in the dashboard and analysis paper that all results are generated from a sample of data, and may not perfectly correspond to the population as a whole. Error statistics will be included to ensure this is evident.

The final known risk for this project is the possible accrument of cost from the use of paid APIs during research in the feasibility study. As this component of the project is a feasibility study, there will not be large-scale use of the APIs, and any exploratory tests will be conducted using small data samples. Most APIs will allow for free use up to a point, and we expect that our usage will fall in the free tier of any API we study. Any accidental costs generated through API usage will be the sole responsibility of the UBC students.

## **8. References**

- [1] <https://www.canada.ca/en/department-finance/services/publications/economic-fiscal-snapshot/overview-economic-response-plan.html>
- [2] <https://www.canada.ca/en/department-finance/services/publications/economic-fiscal-snapshot/summary.html>
- [3] <https://www.canada.ca/en/department-finance/services/publications/economic-fiscal-snapshot/economic-fiscal-profile.html>
- [4] <https://www.canada.ca/en/revenue-agency/services/subsidy/emergency-wage-subsidy/cews-frequently-asked-questions.html>
- [5] <https://covid.smallbusinessbc.ca/hc/en-us/articles/360046425413-Canada-Emergency-Wage-Subsidy-CEWS->
- [6] <https://www.budget.gc.ca/2021/report-rapport/p1-en.html#17>
- [7] <https://www.kli.re.kr/downloadBbsFile.do?atchmnflNo=10152>
- [8] <https://www150.statcan.gc.ca/n1/pub/11-26-0003/112600032021001-eng.htm>

## Appendices

### Appendix A. Description of CEWS Dataset

An example of the dataset:

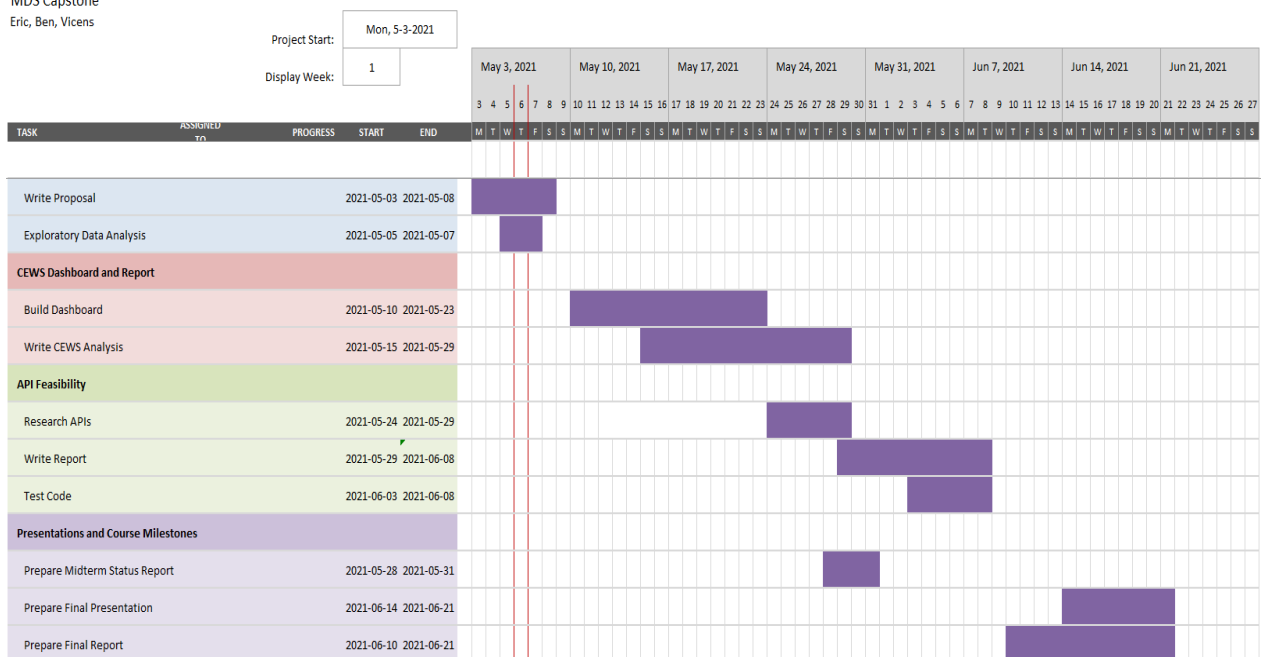
	Start_date_of_CEWS_period	RegionCode	RegionName	RuralUrbanFlag	CMACAFIag	IndustryCode	IndustryName	Number_business_locations	Subsidy_amount	Supported_employees	CEWS_rehire_count
511860	2020-09-27	1315015	Beresford	URBAN	Not applicable	624	Social assistance	0	X	X	0
548431	2020-09-27	4806001	Foothills No. 31	RURAL	Not applicable	56	Administrative and support, waste management a...	10	51,000	74	0
21313	2020-03-15	2457025	McMasterville	URBAN	Not applicable	333	Machinery manufacturing	0	X	X	0
298901	2020-07-05	1006009	Gander	URBAN	Not applicable	711	Performing arts, spectator sports and related ...	0	X	X	0
208267	2020-05-10	4807014	Castor	RURAL	Not applicable	447	Gasoline stations	0	0	0	0

A full description of the dataset can be found in the CEWS metadata, available here:

<https://www150.statcan.gc.ca/n1/en/catalogue/11260003>

### Appendix B. Tentative Time Allocations for Project Tasks

MDS Capstone  
Eric, Ben, Vicens



## Appendix C. Schedule of Important Dates

Week	Date	Topic	Client
1 (May 3 - May 9)	May 4	Professor meeting	UBC
	May 7	Project proposal - Client's approval	StatCan
	May 9	Project Proposal - Due	UBC
2 (May 10 - May 16)	May 10	Client meeting & Group meeting	StatCan
	May 11	Professor meeting	UBC
	May 16	Complete EDA & Data Visualizations	
3 (May 17 - May 23)	May 17	Client meeting & Group meeting	StatCan
	May 18	Professor meeting	UBC
	May 23	Complete Milestone 1 - Data dashboard	
4 (May 24 - May 30)	May 24	Client meeting & Group meeting	StatCan
	May 25	Professor meeting	UBC
	May 30	Prepare for the Presentation(slides, dashboard, practice, etc.)	
	May 30	Complete the feasibility study.	
5 (May 31 - June 6)	May 31	Client meeting & Group meeting	StatCan
	June 1	Status presentation	UBC
6 (June 7 - June 13)	June 7	Client meeting & Group meeting	StatCan
	June 8	Professor meeting	UBC
7 (June 14 - June 20)	June 14	Client meeting & Group meeting	StatCan
	June 15	Professor meeting	UBC
	June 20	Complete the project (data wrangling, dashboard, feasibility analysis, final report, etc.)	
8 (June 21 - June 22)	June 21	Prepare for the final presentation(slides, dashboard, practice, etc.)	

	June 22	Final presentation and report due	UBC
--	---------	-----------------------------------	-----