# Statistics Canada CEWS and API Explorations

This project was completed by Eric Baxter, Vicens Paneque, and Bohan Gao in collaboration with the Statistics Canada Centre for Special Business Projects. The project includes two distinct components: an exploration of Canada Emergency Wage Subsidy data, and an exploration into the use of APIs for the development of novel data sources.

In the first half of the project, we explore the impact of the CEWS program. CEWS is a subsidy program put in place by the Canadian government to help businesses manage the higher costs associated with operating during the Covid-19 pandemic. By providing wage subsidies, the government's goal was to help businesses maintain and rehire employees during the period of reduced income. Data on the CEWS program is compiled in the CEWS Regional and Community-Level Database, which provides geographic and industry breakdowns of the money given out in the CEWS program. In a data exploration done in Python, we examined the effects of geographic regions and industry on given subsidies. We primarily focused on developing insights into the effects of the program on rural Canada and reported on some of the more interesting results. One of our primary findings was that regions in rural Alberta received disproportionately large subsidies, primarily in the hospitality sector and the oil and gas sector. Furthermore, an interactive visualization dashboard was developed in Microsoft PowerBI to allow users to visually explore the CEWS database. This dashboard includes several tools for exploring all dimensions of the data.

During the second half of the project, an inventory of application-programming interfaces (APIs) was taken, with the intent of researching novel sources of economic and social data to be explored at Statistics Canada. Eighteen APIs were researched and reported on. We found that there are major limitations to the useability of most APIs as economic data sources, due to the inaccessibility of aggregate data and the need for user authentication. Of our researched APIs, we considered four APIs “feasible” (usable for our goals), five APIs “somewhat feasible” (limited use), and 9 APIs “practically infeasible” (no obvious use). The most promising use of APIs we found was in social analytics. Through some APIs, we were able to retrieve user posts and comments. Using natural language processing and sentiment analysis methods, we were then able to parse these messages and gain insight into public opinion on certain topics. Preliminary NLP software was developed for three of APIs: Twitter, Steam, and Reddit.