**Final Report**

* Introduction where you discuss the business problem and who would be interested in this project.

In this project, I will compare two cities, New York City (NYC) and Toronto. Both of these cities are the financial capitals of their countries. The most popular Boroughs of these two cities were selected and compared, Downtown Toronto and Manhattan. Here, we intended to compare the venues of these two boroughs. This report can provide useful information to companies who want to open new business, for them to evaluate the potential success. Since these two Boroughs are both located in North America, with similar cultures, the successes of certain business in one Borough may indicate it’s also needed in another similar borough.

* Data where you describe the data that will be used to solve the problem and the source of the data.

In order to complete the report, we need the following dataset:

The list of postal code of Canada from Wikipedia (<https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>)

The geographical data of Toronto provided by IBM

The data for NYC provided by IBM (https://cocl.us/new\_york\_dataset)

The data of neighborhoods provided by Foursquare API

* Methodology section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, if any, and what machine learnings were used and why.

Step 1: import all necessary libraries and tools, including numpy, pandas, matplotlib, seaborn, json, sklearn, geocoder and Folium.

Step 2: The boroughs and neighborhood data of Toronto were imported from Wikipedia, and the geographical data of Toronto were imported from IBM service. These two tables were cleaned, renamed and merged into one table, and only the Borough of Downtown Toronto were analyzed later, and thus its data were used to create a new dataframe named “downtown\_toronto\_data”.

Step 3: The data for NYC were provided by IBM in json format and were directly used. Since only Manhattan is the interest of this research, its data were extracted and saved as a new dataframe named Manhattan\_data.

Step 4: After two dataframe were created for these two boroughs, Foursquare API was used to explore their neighborhoods. Top 5 most common venues were extracted for their neighborhoods, and maps were created.

* Results section where you discuss the results.

The results indicated that the most popular venue in the neighborhoods of Downtown Toronto is Coffee shop, but for Manhattan, the 1st most common venues were different dependent on neighborhoods. For the 39 neighborhoods of Manhattan, Italian Restaurant is the 1st most common venue for 12 neighborhoods, and other popular venues include coffee shop, café, Deli, part, etc. But overall, these two cities share similarities in that coffee shop, café and restaurants are the top venues. People at Downtown Toronto more prefer Vegetarian / Vegan Restaurant, Asian (Vietnamese, Chinese and Japanese, while people in Manhattan more like Italian and Mexican restaurants.

* Discussion section where you discuss any observations you noted and any recommendations you can make based on the results.

From the results section, for business runners at both boroughs, opening a coffee shop is always a good idea. Opening a restaurant is also good, but the styles vary depending on boroughs and their neighborhoods. Other businesses that can be considered include baker, boutique and bars.

* Conclusion section where you conclude the report.

The downtown Toronto and Manhattan neighborhoods are very similar in their most common venues, but they are also different in details, such that although restaurants are popular in both boroughs, the styles people prefer are different.