Segmentation and Comparison of Neighborhood of New York City and Toronto

1. Introduction/Business Problem

Problem Background:

Both New York City and the city of Toronto are the financial capitals of their respective countries. They provide a lot of business opportunities and business friendly environment. Also, they have attracted many different players into the market. As global hubs of business and commerce, these two market are highly competitive. Thus, any new business venture or expansion needs to be analyzed carefully. The insights derived from comparison analysis will give good understanding of business environment which help in strategically targeting the market. This will help in reduction of risk and the return on investment will be reasonable.

Problem Description:

The company need to choose the correct location to start its first venture. If this is successful, they can replicate the same in other locations. First move is very important, thereby choice of location is very important. It's important to understand how New York City similar and dissimilar with Toronto. The average income and population density of boroughs of each city are needed to be explored. In order to learn about the competitors in New York City and Toronto, we need to explore the most common 20 venue categories of every neighborhood in each city.

Target Audience:

To recommend the correct location, the company has appointed me to lead of the Data Science team. The objective is to compare New York city and Toronto and recommend to the management which city and its neighborhood will be the best choice to start a restaurant. The management also expects to understand the rationale of the recommendations made.

This would interest anyone who wants to start a new restaurant in New York city or Toronto.

Success Criteria:

The success criteria of the project will be a good recommendation of borough/Neigborhood choice

to the company based on large population, average income, lack of such restaurants in that location and nearest suppliers of ingredients.

2. Data:

Data 1: Borough, Neighborhood, Latitude and Longitude of New York city data from https://cocl.us/new_york_dataset;

Borough, Neighborhood, Latitude and Longitude of Toronto data: merge data from Wikipedia https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M and geospatial coordinates of each neighborhood from http://cocl.us/Geospatial_data

Data 2: DOHMN Farmers Markets and Food Boxes dataset of the New York City.
https://data.cityofnewyork.us/dataset/DOHMH-Farmers-Markets-and-Food-Boxes/8vwk-6iz2

Data 3: Demographics of Toronto neighbourhoods:
https://en.wikipedia.org/wiki/Demographics_of_Toronto_neighbourhoods
Boroughs of New York City:
https://en.wikipedia.org/wiki/Boroughs_of_New_York_City

Data 4: d venues data from the Foursquare API were imported into two data frame for Toronto and New York city.