Introduction/Business Problem
New York and Toronto are the financial capitals of their respective countries and are very diverse. Both cities offer a wide variety of world's delicious food.
Target Audience: The task is to compare them and provides insights to people who would like to open a restaurant in these two cities.

Data Set
In order to compare New York City and Toronto, data such as geographical data and venue data were collected. Geographical data include boroughs, neighbourhoods, latitude and longitude for both cities to draw maps.

To compare different public venues in the neighbourhoods of both cities, data were collected from Foursquare api and appended into the geographical data.

• A .json file to fetch data of New York city: https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs/newyork_data.json
• A Wikipedia Page to fetch data of Toronto city: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
• Foursquare api to fetch different public venues in the vicinity of the neighbourhood

Methodology section:
Data downloaded or scraped from multiple sources were combined into one table for New York City and Toronto respectively. After data cleaning, New York City were divided into 5 boroughs with 306 neighbourhoods while Toronto were divided into 15 boroughs with 103 neighbourhoods. Geopy library were used to get the latitude and longitude values of New York City and Toronto.

Foursquare API was applied to explore the neighbourhoods and segment them by accessing and acquiring the venue data such as venue name, venue unique ID, venue category, venue location (latitude and longitude) etc. A Foursquare Developer account was registered to obtain credentials (ie. client ID and client Secret key).

Each Neighbourhood was analysed and top 10 most common venues with their frequency for each neighbourhood were displayed.

Performing K-means to cluster the neighbourhoods into 5 clusters. Then each cluster was examined to discriminate venue categories that distinguish each cluster. Folium was used to help create maps of New York and Toronto with clustered neighbourhoods on top for visualisation purpose.

Results:
The results allow the target audience to identify which neighbourhoods have higher concentration of restaurants while which have fewer.