**Coursera IBM data science – Capstone Project**

**Introduction**

Toronto is the most popular city in Canada (the largest country in North America). Different ethnic neighborhood throughout the city focuses on specific cuisines such as Italian cuisine in Little Italy, Chinese in Chinatown. Numerous world cuisines are available throughout the city including Portuguese, Japanese, Indian, Greek and Caribbean.

For this capstone project, I am going to analyze all available restaurants in Toronto which may help to an entrepreneur to decide which cuisine restaurant will be best to open in a specific location in Toronto. Analyzing the available restaurant categories in different neighborhoods would give us an idea of the distribution of the restaurant types across Toronto.

**Business Problem**

The objective of this capstone project is to determine which places in Toronto has which type of restaurant more. By using machine learning algorithm such as clustering this project aims to find which cuisine restaurants dominate in which part of Toronto.

**Data**

To solve this problem, I am going use the below data sources.

* List of neighborhoods in Toronto, Canada.
* Latitude and longitude of these neighborhoods.
* Venues related to food category around different neighborhoods.

**Extracting the data**

* Web scrapping of Toronto neighborhood data from Wikipedia (‘https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M’).
* Fetch location details (latitude and longitude) of these neighborhoods using google geocoder package or use the existing csv file with the neighborhood location details (Geospatial\_Coordinates.csv).
* Use Foursquare API to fetch food venue details for neighborhoods in Toronto using category id ‘4d4b7105d754a06374d81259’.