Introduction/Business Problem

A multi-national company wants to set up the manufacturing facility for Restaurant Equipment and Supply in the neighborhood of Toronto. The company's management wants the perfect location which is closer to restaurant of diverse types. They want diverse customers to get more orders.

To meet their requirements, they gave the data scientist a task to find the neighborhood with diverse restaurant types.

Data Scientist team gathered the data and perform K-means Clustering to give prediction on the diversity of the neighborhood according to the Postal Codes of each neighborhood.

Data Acquisition and Cleaning

Data downloaded or scraped from Wikipedia were combined into one table. There were missing values in Borough and Neighborhood. Those rows were dropped. Second data which was possessed through Foursquare location search, was used to explore the restaurants in neighborhood of Toronto.

Data feature of 'Postal Code' was used to combine the two table from Wikipedia and Foursquare location to get diverse types of restaurant according to the 'Postal Code' of the neighborhood. Top ten nearest restaurant were listed according to Longitudinal and Latitudinal axis of the location of the Neighborhood.