Data processing

This project will compile data frames for each financial center in the city of Toronto, Canada, by combining data scraped from several online sources. The data that I am going to use contains the neighborhoods, boroughs, and their respective longitude and latitude for each city.

First, I will compile the data containing the top 5 most popular venues for each neighborhood which is used in a KMeans algorithm to cluster the neighborhoods together based on the similarity of the categories for their top 5 venues by using Foursquare location data to generate the map.

This project only utilizes venues when grouping similarities. I plan to incorporate more data in the future such as proximity to water, population in each neighborhood, prices of houses, etc. to further create unique clusters to help identify similar neighborhoods.