**Battle of The Neighborhoods**

Rui Luo

July 25 2019

1. **Introduction**

1.1 Business Problem:

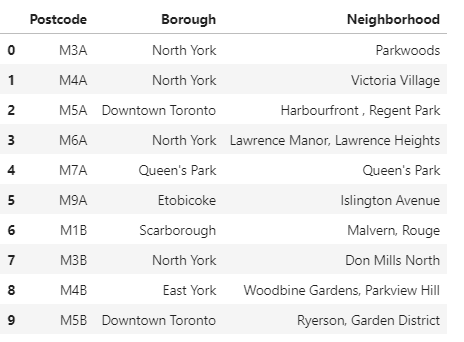
As an Asian who is planning to move to Toronto from another city, how can I find the most Asian-friendly neighborhoods to settle down in Toronto?

1.2 Interest:

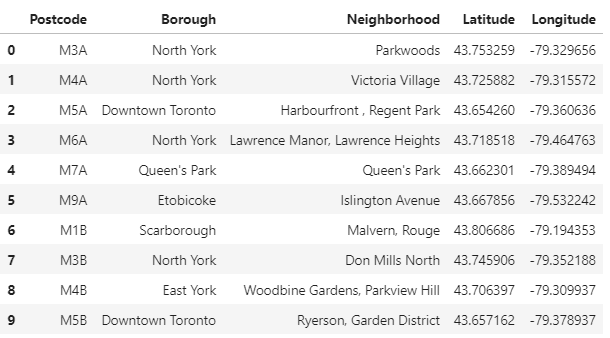
Asians who are planning to move to Toronto but are not familiar with Toronto.

1. **Data acquisition**

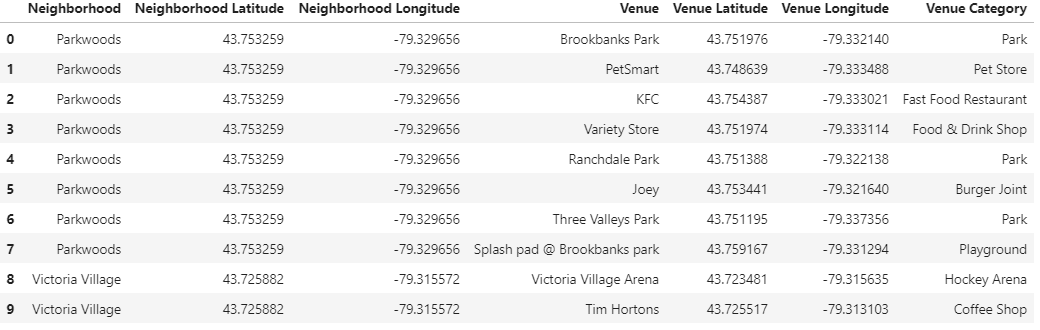
**2.1** Postcode, Borough and Neighborhood information of Toronto is retrieved from Wikipedia webpage <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>. BeautifulSoup is used to do scraping and some data clean is done to get only the Postcodes, Borough and Neighborhood information.



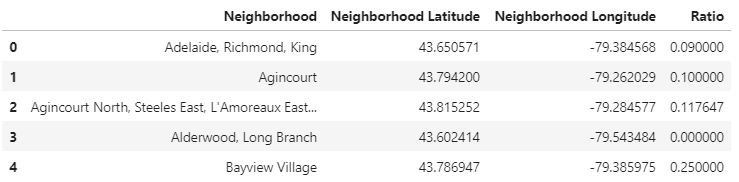
**2.2** Geo-spatial data is retrieved from <http://cocl.us/Geospatial_data> to get Latitude and Longitude for all the Neighborhoods in Toronto



**2.3** Venue information is retrieved from FourSquare



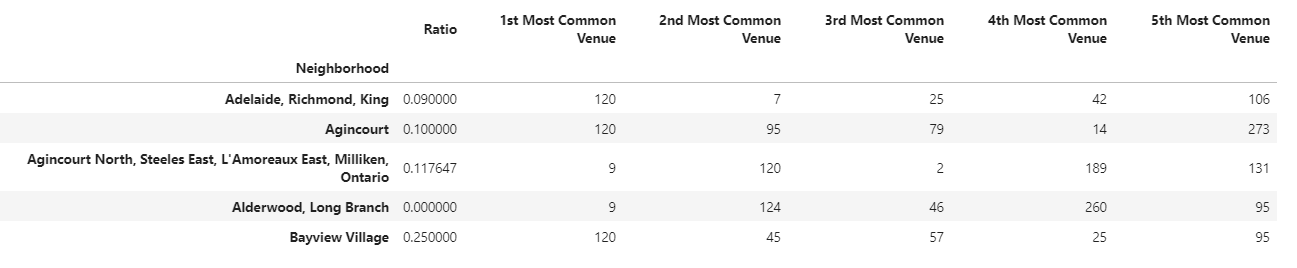
**2.4** Ratio of Asian Restaurant among all venue categories is calculated for all the neighborhoods since it is used as a factor in k-means.



**2.5** The other factor for the k-means algorithm is the Top 5 venue categories for each neighborhood.



**2.6** The two data frames from 2.4 and 2.5 are merged to get the final data frame for the k-means algorithm. And since k-means can’t be applied to process float and string data at the same time, the categories are mapped into corresponding numbers,



1. **Methodology**

Use k-means algorithm to cluster the neighborhoods based on 1) ratio of Asian Restaurant among all venue categories, and 2) top 5 venue categories in each neighborhood.

1. **Result**

The neighborhoods in Toronto are clustered into 10 clusters. And Cluster #7 (the yellow circles shown below) has the highest ratio of Asian restaurant and the most numbers in top 5 venue categories. So neighborhoods in cluster 7 are recommended.

