**Introduction**

In every age and time, people have been moving around to the world for one reason or the other. When they move to a new city, they need to select a place where they want to buy new house or where they want to reside. While selecting a place they have several factors in mind depending on their scenarios. Such as a single person moving due to his job might look for place close to his job but at the same time some other factors like coffee shops, restaurants, banks, pharmacy, gym and more. On the other hand, if a family is moving with their children one of the many important things for them will be a school nearby.

So to find a place of their choice they have to physically go in different neighborhoods in the city or search thoroughly on the internet, which could be cumbersome. And it is also difficult to have a comparison of all neighborhoods in a big city. The aim of this project is to cluster neighborhoods based on the factors mentioned above and make it easy for the people to decide which neighborhood to choose to buy a house.

**Data**

The solution is specifically provided for the Toronto City that is for the people who are moving to Toronto. For this data for neighborhood along with their postal codes has been collected from Wikipedia page: 'https://en.wikipedia.org/w/index.php?title=List\_of\_postal\_codes\_of\_Canada:\_M&oldid=1011037969' as shown in the figure 1. From the data, the postal code that has not been assigned any neighborhoods was deleted. After that this data has been merged with geospatial data gathered from 'https://cocl.us/Geospatial\_data'. The figure shows… two columns Longitude and Latitude are added to the table for every corresponding neighborhood.

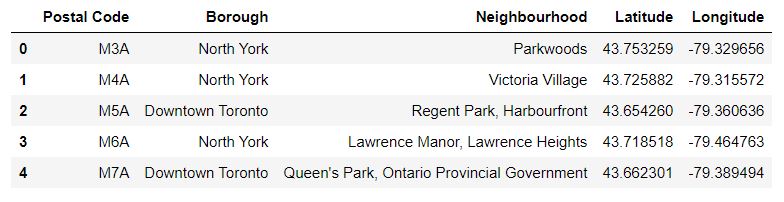
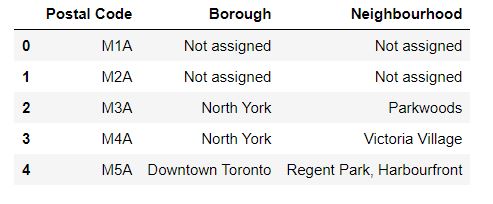


Figure Figure 2

After gathering Toronto city neighborhoods with their Longitude and Latitude, Foursquare API has been used to get nearby venues for each neighborhood. Among those, venues with top 10 frequencies are being selected for calculating similarity among neighborhoods. The below figure 3 shows those top 10 venues with their respective mean against each neighborhood.



Figure