

Final Report - Capstone Project - The Battle of Neighborhoods

Introduction:

The purpose of this project is to assist people in exploring better facilities around their neighborhood. It'll help people making smart and efficient decision on selecting great neighborhood out of all neighborhoods in Scarborough, Toronto. Lots of people are migrating to varied states of Canada and needed many research permanently housing prices and reputed schools for his or her children. This project is for those people that are trying to find better neighborhoods. For simple accessing to Cafe, School, Supermarket, medical shops, grocery shops, mall, theatre, hospital, compatible people, etc.

This project's aim is to make an analysis of features for nation migrating to Scarborough to look-up a be neighborhood after a comparative analysis between neighborhoods. The features include median housing price and better school consistent with ratings, crime rates of that specific area, road connectivity, weather, good management for emergency, water resources both fresh and wastewater and excrement conveyed in sewers and recreational facilities. The major purpose of this project is to suggest a far better neighborhood for the one that are shifting there. Social presence in society in terms of compatible people. Connectivity to the airport, bus stand, city center, markets, and other daily needs things nearby.

It will help people to urge awareness of the world and neighborhood before moving to a replacement city, state, country, or place for his or her work or to start out a replacement fresh life.

Data Description

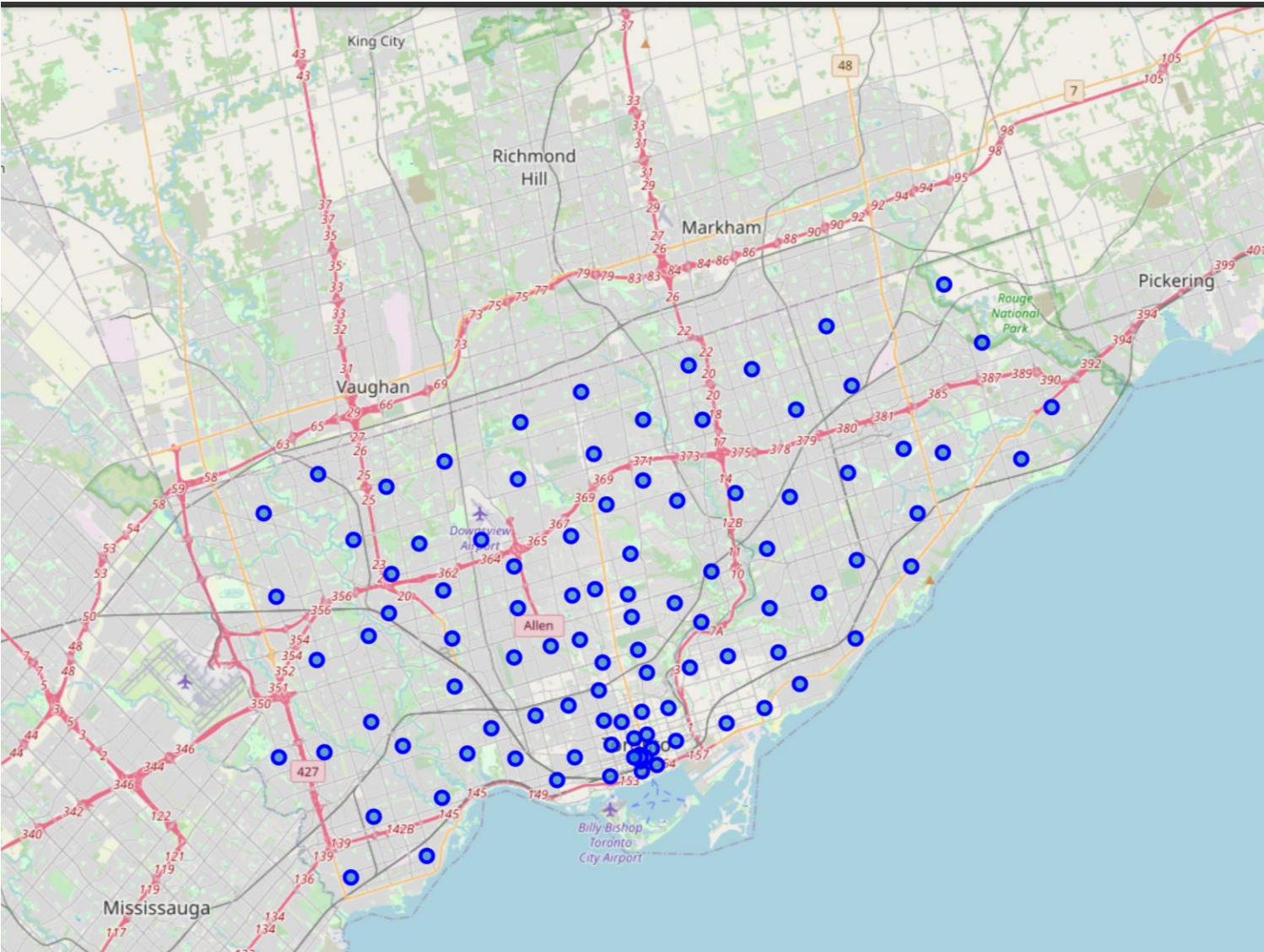
Link: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

Dataset consisting of latitude and longitude, zip codes of Scarborough, Canada after cleaning in the 3rd week of Capstone. We will need data about different venues in different neighborhoods of that specific borough. In order to gain this information, we are using "Foursquare" locational information. Foursquare is a very gigantic location data provider with information about all sorts of venues and events within any area. Such information includes venue names, locations, menus, photos and even reviews. As such, the foursquare location platform will be used as the sole data source, As all information we need is provided by foursquare there will be no need to rely on other data sources.

After connecting to the Foursquare API to gather information about venues inside each and every neighborhood. For each neighborhood, the data retrieved from Foursquare contains information of venues within a specified distance of the longitude and latitude of the postcodes. The data for every venue is as follows:

- Neighborhood
- Neighborhood Latitude
- Neighborhood Longitude
- Venue
- Name of the venue
- Venue Latitude
- Venue Longitude
- Venue Category

Map of Scarborough



Methodology

I used K-means clustering to segment the big city into smaller clusters that basically group neighborhoods that have similar properties in terms of venues present near them. I formed 3 clusters. To be able to do that I ran K-means clustering algorithm and assigned cluster labels to each row so as to classify that neighborhood. I also managed to find out the Top 10 most common venues in every neighborhood. This data was definitely helpful in forming clusters.

```
neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)

Scarborough_merged = df_2.iloc[:16,:]

# merge toronto_grouped with toronto_data to add latitude/longitude for each neighborhood
Scarborough_merged = Scarborough_merged.join(neighborhoods_venues_sorted.set_index('Neighborhood'), on='Neighborhood')

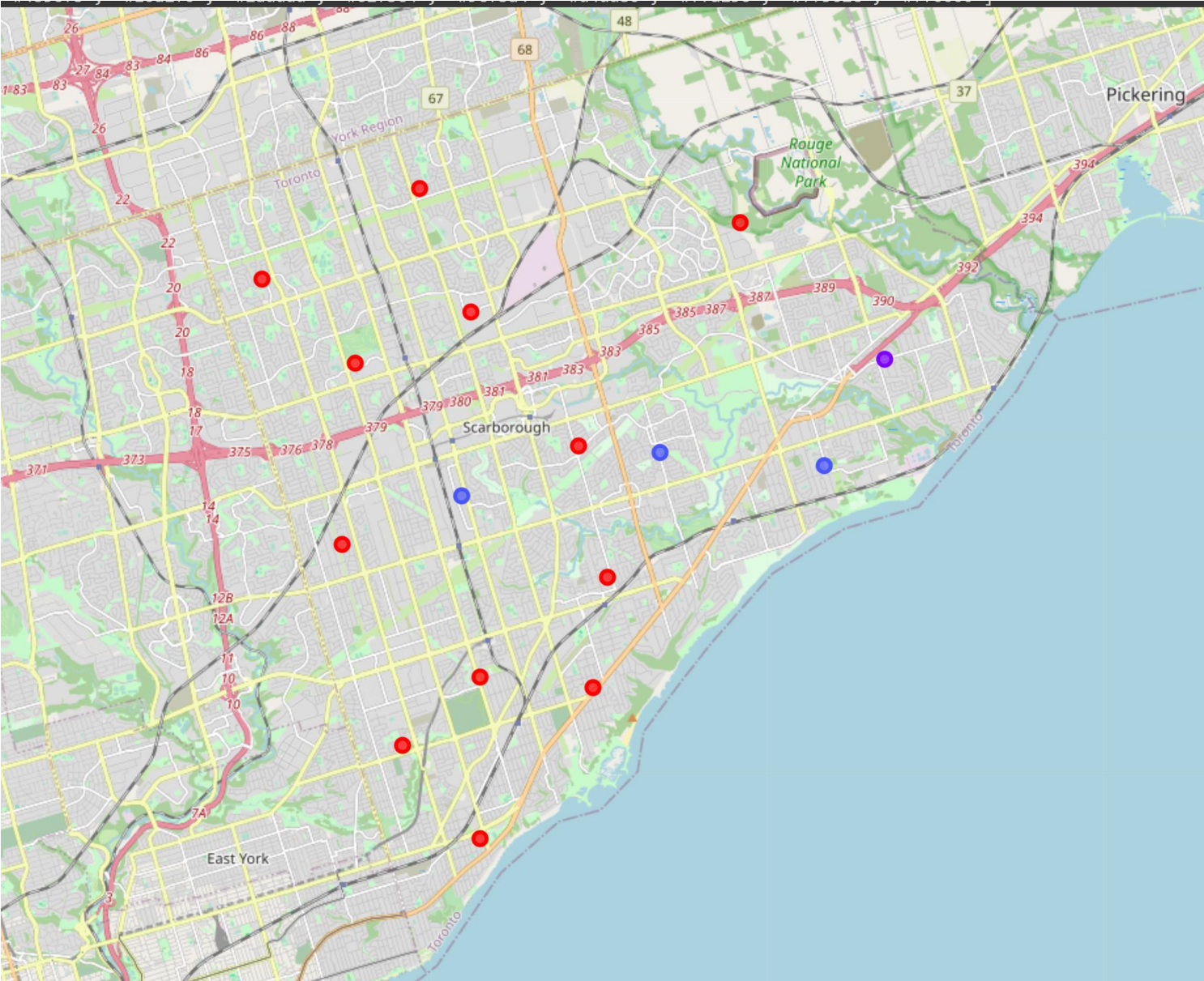
Scarborough_merged.head()# check the last columns!
```

	PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	M1B	Scarborough	Malvern, Rouge	43.81139	-79.19662	0	Zoo Exhibit	Construction & Landscaping	Fast Food Restaurant	Paintball Field	Escape Room	Doner Restaurant	Donut Shop	Dumpling Restaurant	Eastern European Restaurant	Electronics Store
1	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek	43.78574	-79.15875	1	Bar	Yoga Studio	Ethiopian Restaurant	Donut Shop	Dumpling Restaurant	Eastern European Restaurant	Electronics Store	Elementary School	Escape Room	Event Space
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.76575	-79.17470	2	Park	Gym / Fitness Center	Athletics & Sports	Gymnastics Gym	Yoga Studio	Dog Run	Doner Restaurant	Donut Shop	Dumpling Restaurant	Eastern European Restaurant
3	M1G	Scarborough	Woburn	43.76812	-79.21761	2	Fast Food Restaurant	Chinese Restaurant	Park	Coffee Shop	Yoga Studio	Elementary School	Doner Restaurant	Donut Shop	Dumpling Restaurant	Eastern European Restaurant
4	M1H	Scarborough	Cedarbrae	43.76944	-79.23892	0	Bakery	Bank	Caribbean Restaurant	Gas Station	Athletics & Sports	Hakka Restaurant	Thai Restaurant	Electronics Store	Doner Restaurant	Donut Shop

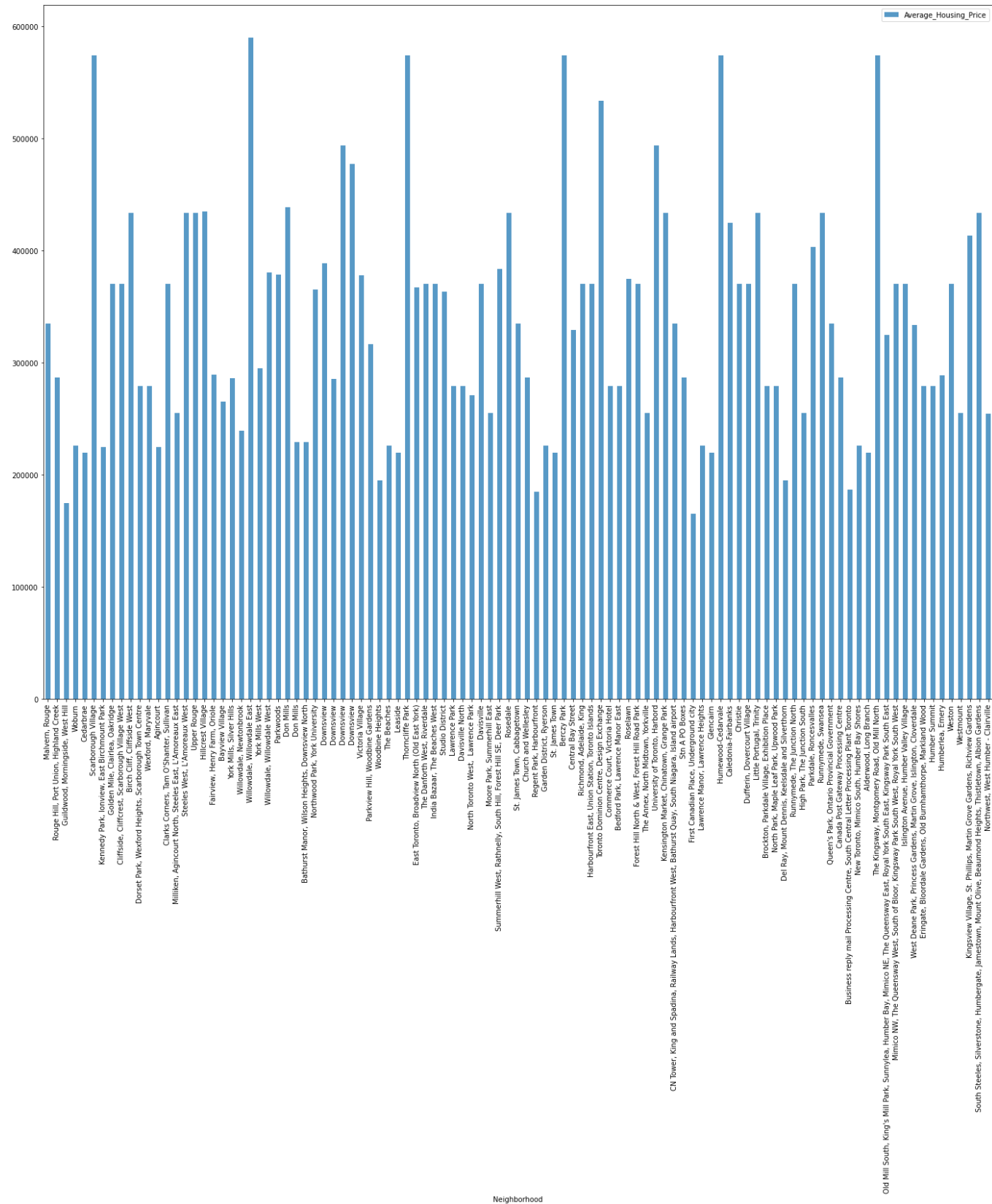
Note: The data used was obtained using the credentials of Foursquare API, all data about neighborhoods and nearby venues was mined here. Due to limitations, the API could only set number of places near neighborhood parameter to 100 and radius parameter to 500.

Results

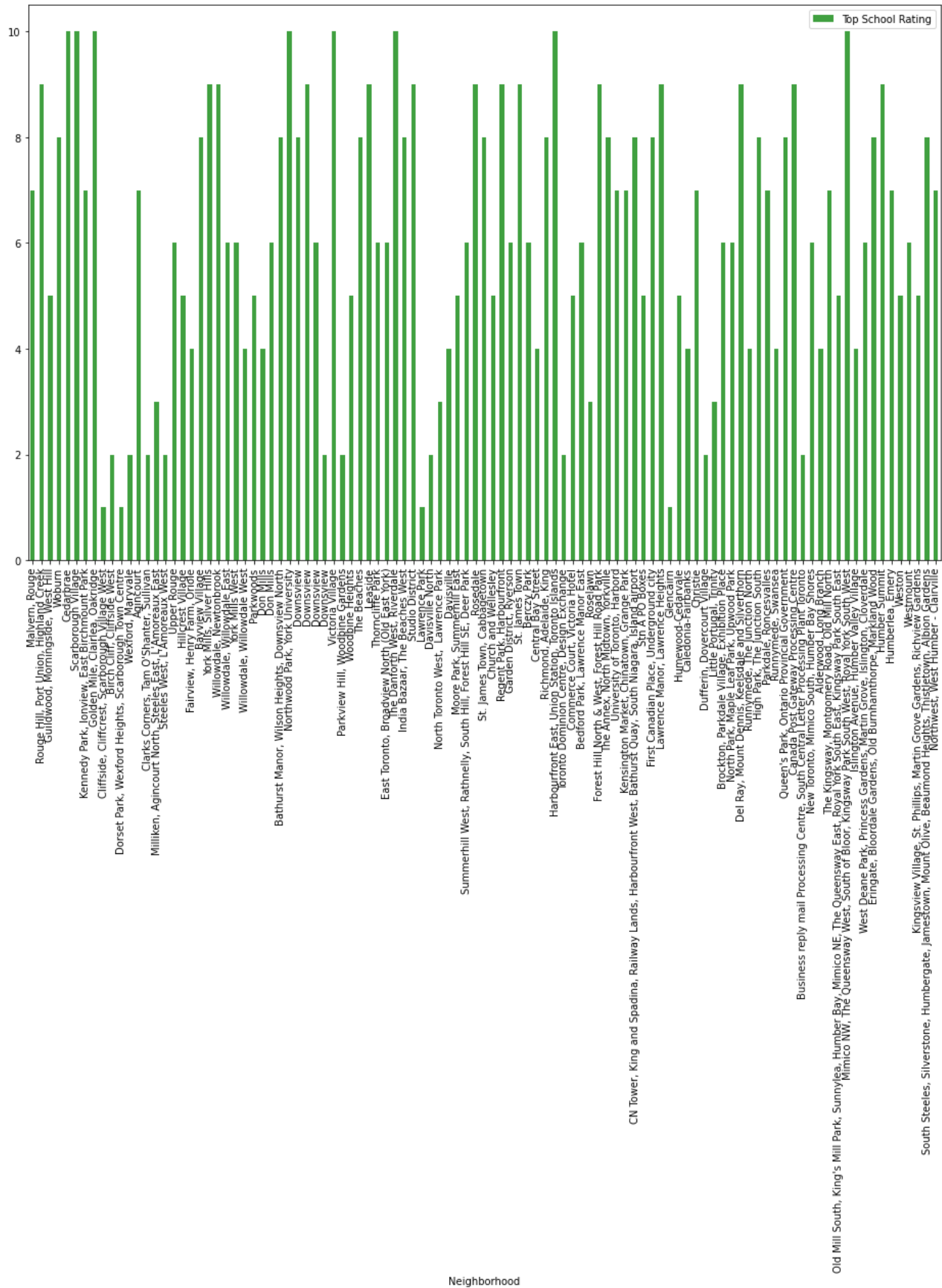
Map of clusters (3) formed using K-means in Scarborough



Average House pricing in Scarborough by clusters



School ratings in Scarborough by clusters



The Location:

Scarborough may be a popular destination for brand spanking new immigrants in Canada to reside. As a result, it is one among the foremost diverse and multicultural areas within the Greater Toronto Area, being home to varied religious groups and places of worship. Although immigration has become a hot topic over the past few years with governments seeking more restrictions on immigrants and refugees, the overall trend of immigration into Canada has been increasing.

Foursquare API:

The project has extensively used the foursquare API for location data as the database has extensive data on millions of venues worldwide. The Places API offers real-time access to Foursquare's global database of rich venue data and user content to power your location-based experiences in your app or website.

Conclusion

In battle of the neighborhoods, I used k-means cluster algorithm and separated the neighborhood into 10(Ten) different clusters for all the 103 different latitude and longitude that I could find on the Wikipedia page, forming groups/clusters on the basis of common venues present near those neighborhoods. Using charts, I have presented the average house pricing and also the school ratings in these neighborhoods to help the person moving to Scarborough in getting extra information that might affect his/her selection of neighborhood.

It was really great working on such a unique project. It helped me learn a lot in terms of how data science helps us in day to day. I have worked hard on this and I won't lie It took a lot of time but at the same time This course was definitely worth it. All the new tools that I used here for web-scraping, geospatial data, real-live maps will definitely help me in the future.