1. Introduction:

The purpose of this Capstone Project is to help people in exploring better facilities around their neighborhood. It will help people making smart and efficient decision on selecting great neighborhood out of numbers of other neighborhoods in Scarborough, Toranto.
Lots of people are migrating to various states of Canada and needed lots of research for good housing prices and reputated schools for their children. This project is for those people who are looking for better neighborhoods. For ease of accessing to Cafe, School, Super market, medical shops, grocery shops, mall, theatre, hospital, like minded people, etc.
This Capstone Project aim to create an analysis of features for a people migrating to Scarborough to search a best neighborhood as a comparative analysis between neighborhoods. The features include median housing price and better school according to ratings, crime rates of that particular area, road connectivity, weather conditions, good management for emergency, water resources both fresh and waste water and excrement conveyed in sewers and recreational facilities.
It will help people to get awareness of the area and neighborhood before moving to a new city, state, country or place for their work or to start a new fresh life.

2. Data Section

Data Link: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
Will use Scarborough dataset will be used to scrap the data from wikipedia.
Dataset consists of latitude and longitude, zip codes.

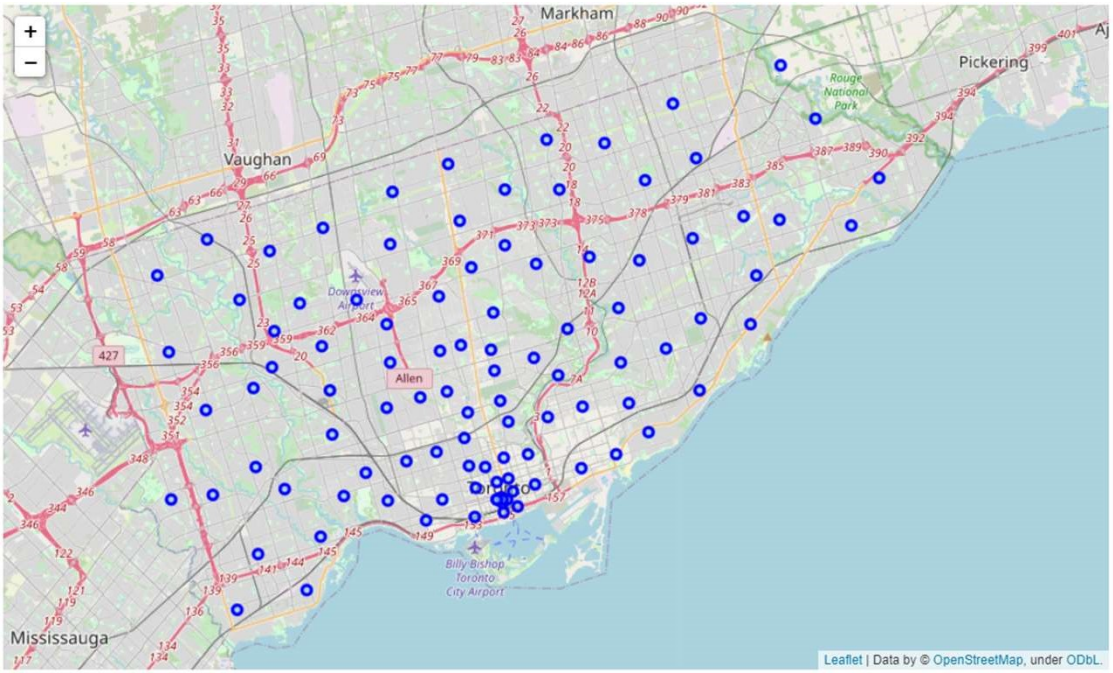
Foursquare API Data:
We need data about different venues in different neighborhoods specifically for borough.
In order to gain that information we will use “Foursquare” locational information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API.
After finding the list of neighborhoods, we then connect to the Foursquare API to gather information about venues inside each and every neighborhood. For each neighborhood, we have chosen the radius to be 100 meter.
The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

1. Neighborhood
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Venue
5. Name of the venue e.g. the name of a store or restaurant
6. Venue Latitude
7. Venue Longitude
8. Venue Category

3. Methodology Section

Clustering Approach:

To compare the similarities of two cities, we decided to explore neighborhoods, segment them, and group them into clusters to find similar neighborhoods in a big city like New York and Toronto. To be able to do that, we need to cluster data which is a form of unsupervised machine learning: k-means clustering algorithm.



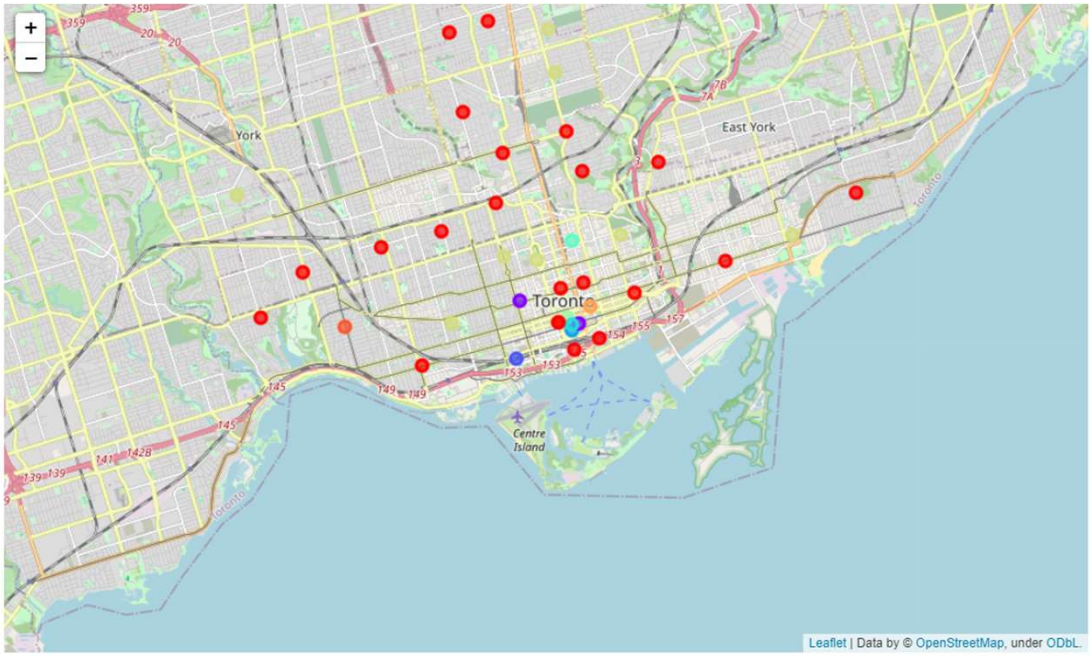
Map of Scarborough

	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 M Comm Ve
0	M4E	East Toronto	The Beaches	43.67709	-79.29547	9	Trail	Health Food Store	Pub	Wine Shop	Discount Store	Event Space	Ethiopian Restaurant	Electror St
1	M4K	East Toronto	The Danforth West, Riverdale	43.68375	-79.35512	0	Park	Bus Line	Grocery Store	Business Service	Intersection	Coffee Shop	Discount Store	Ice Cre Sl
2	M4L	East Toronto	India Bazaar, The Beaches West	43.66797	-79.31467	0	Fast Food Restaurant	Pizza Place	Park	Pet Store	Italian Restaurant	Sandwich Place	Pub	Ice Cre Sl
3	M4M	East Toronto	Studio District	43.66213	-79.33497	0	Italian Restaurant	Brewery	Diner	American Restaurant	Coffee Shop	Arts & Crafts Store	Gay Bar	Su Restau
4	M4N	Central Toronto	Lawrence Park	43.72843	-79.38713	6	Swim School	Bus Line	Wine Shop	Dog Run	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electror St

Clustering Data Analysis

4. Results Section

Map of Clusters in Scarborough



The Location:

Scarborough is a popular destination for new immigrants in Canada to reside. As a result, it is one of the most diverse and multicultural areas in the Greater Toronto Area, being home to various religious groups and places of worship. Although immigration has become a hot topic over the past few years with more governments seeking more restrictions on immigrants and refugees, the general trend of immigration into Canada has been one of on the rise.

5. Discussion

The major purpose of this project, is to suggest a better neighborhood in a new city for the person who are moving to a new town. Social presence in society in terms of like minded people. Connectivity to the airport, bus stand, city center, markets and other daily needs things nearby.

6. Conclusion

In this Capstone project, using k-means cluster algorithm I separated the neighborhood into 10 clusters and for 103 different latitude and longitude from dataset, which have very-similar neighborhoods around them.

I feel rewarded with the efforts and believe this course with all the topics covered is well worthy of appreciation.

This project has shown me a practical application to resolve a real situation that has impacting personal and financial impact using Data Science tools.

The mapping with Folium is a very powerful technique to consolidate information and make the analysis and decision better with confidence.