The battle of the neibourhoods : Report

Introduction :

This project aims to make it easier for people to make decisions when in comes to comparing and selecting neibourhoods to live in.

About 300,000 immigrants come to Canada each year, and about 150,000 immigrants become citizens each year. Prior to migrating or moving out people do a lot of research about housing prices and good schools etc. The purpose of this project is to make it easier for these people to find a suitable neibourhood according to their preferences and expectaions.

In this project we will make an analysis of different components for people migrating or moving to Scarborough. These components are real estate prices, schools quality, crime statistics, weather conditions as well as proximity and variety of leisure activities.

Problem to solve :

The goal of this project is being able to offer a comparative analysis between neibourhoods to suggest the more suitable according to one's preferences.

Location :

Some Scarborough neighbourhoods are popular destinations for immigrants who add their own culture to Scarborough's.

Fousquare API :

This project would use Four-square API as its prime data gathering source as it has a database of millions of places, especially their places API which provides the ability to perform location search, location sharing and details about a business.

Data Description :

We will retrieve our data from : link : "https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M"

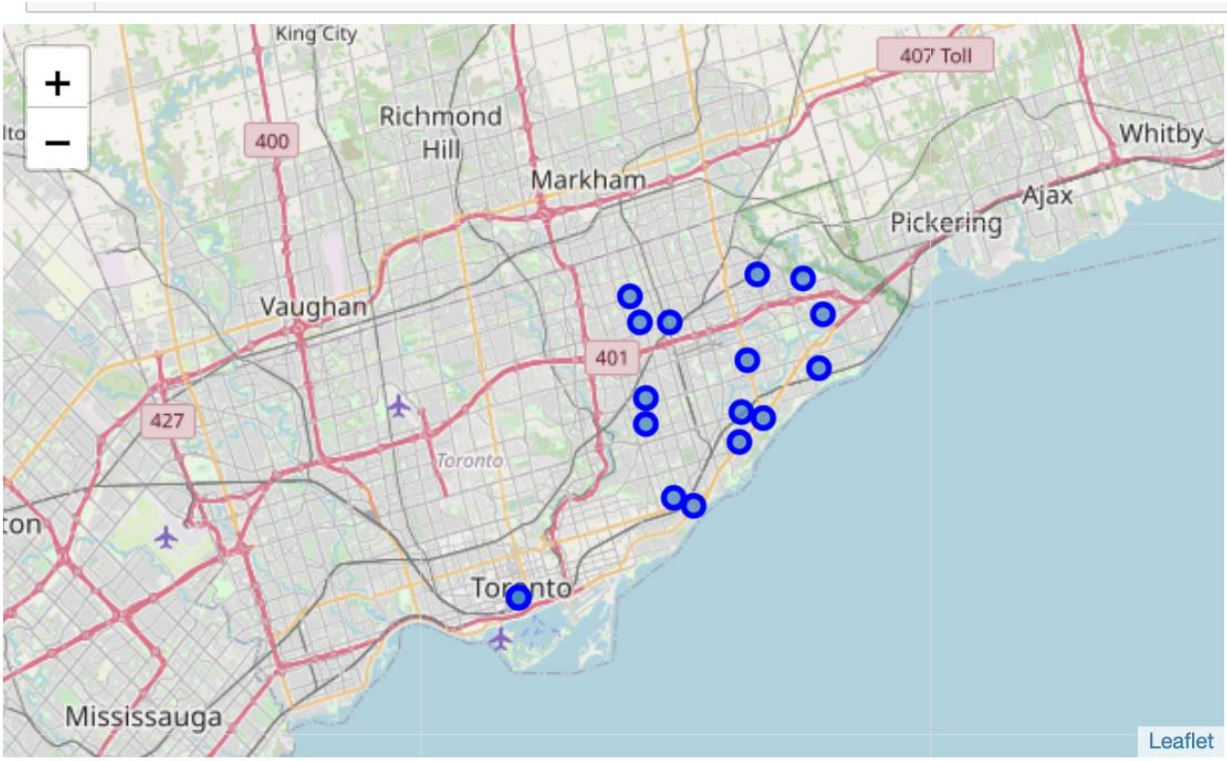
We will make use of the Scarborough dataset

Foursquare API :
We will use Foursquare locational information in order to get the information about the different venues such as venue names, location and menus as well as the menus.

The data retrieved from Foursquare API contains informations about venues within a specified distance (radius of 100 meters) such as : Neiborhood, Neiborhood Latitude, Neighborhood Longitude, Venue, Name of the venue e.g. the name of a store or restaurant, Venue Latitude, Venue Longitude, Venue Category.

Map :

Out[25]:



Methodology :

To carry out this project, we decided to explore neibourhoods segment the and group them into clusters to find similarities between the neibourhoods. In order to do that we used k-means clustering.

K-means clustering :

Out[71]:

Neighborhood	Latitude	Longitude	Cluster Label	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 Mc Comm Ven
M9AEtobicoke(Islington Avenue)	43.64869	-79.38544	0	0	Coffee Shop	Hotel	Café	Restaurant	Beer Bar	Furniture / Home Store	Bookstore	Italian Restaurant	Gy
M9BEtobicoke(West Deane Park / Princess Garden...	43.81023	-79.22038	0	0	Fast Food Restaurant	Pharmacy	Pizza Place	Park	Supermarket	Grocery Store	Bubble Tea Shop	Sandwich Place	Convenien Sto
M9CEtobicoke(Eringate / Bloordale Gardens / Ol...	43.78948	-79.17614	0	0	Convenience Store	Home Service	Wings Joint	Fast Food Restaurant	Coffee Shop	College Stadium	Comedy Club	Concert Hall	Cosmeti Sh

Most common venues in vicinity :

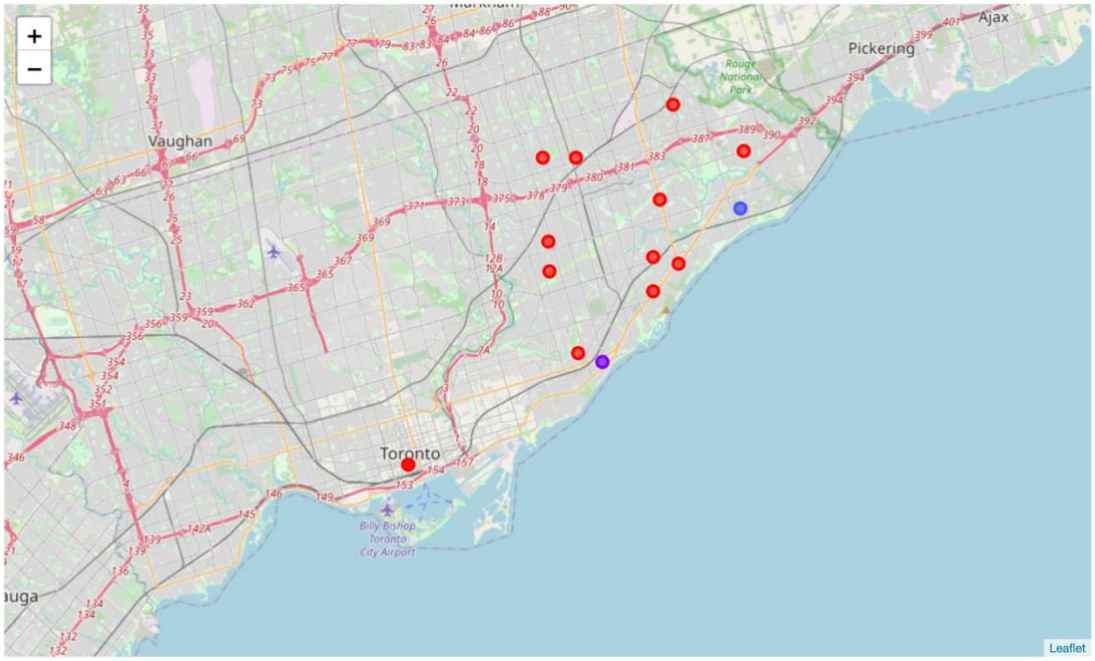
Out[71]:

rhood	Latitude	Longitude	Cluster Label	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
ington venue)	43.64869	-79.38544	0	0	Coffee Shop	Hotel	Café	Restaurant	Beer Bar	Furniture / Home Store	Bookstore	Italian Restaurant	Gym	Japanese Restaurant
e(West incess rden...	43.81023	-79.22038	0	0	Fast Food Restaurant	Pharmacy	Pizza Place	Park	Supermarket	Grocery Store	Bubble Tea Shop	Sandwich Place	Convenience Store	Skating Rink
ingate dens / Ol...	43.78948	-79.17614	0	0	Convenience Store	Home Service	Wings Joint	Fast Food Restaurant	Coffee Shop	College Stadium	Comedy Club	Concert Hall	Cosmetics Shop	Deli / Bodega

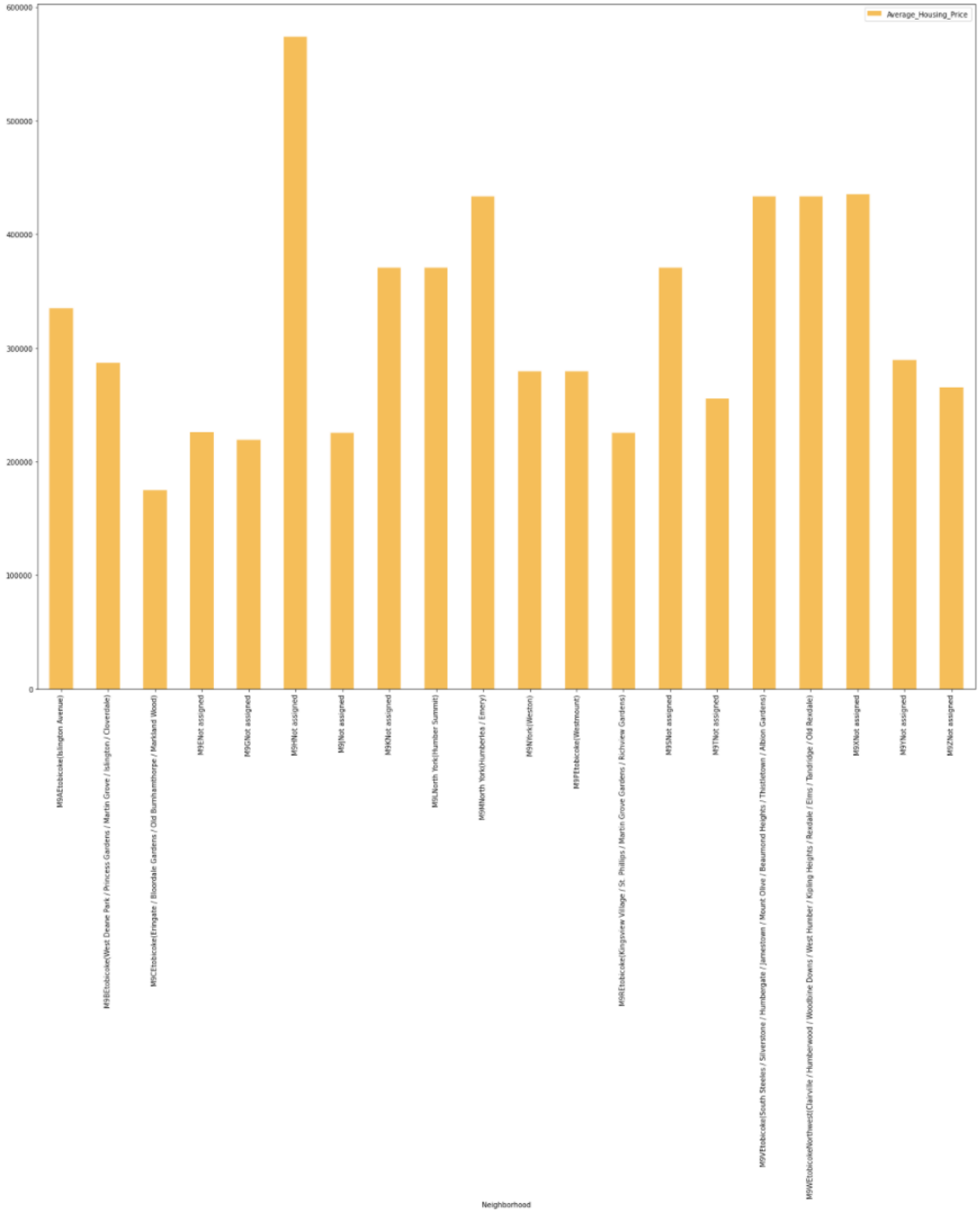
Results Section :

Map of clusters :

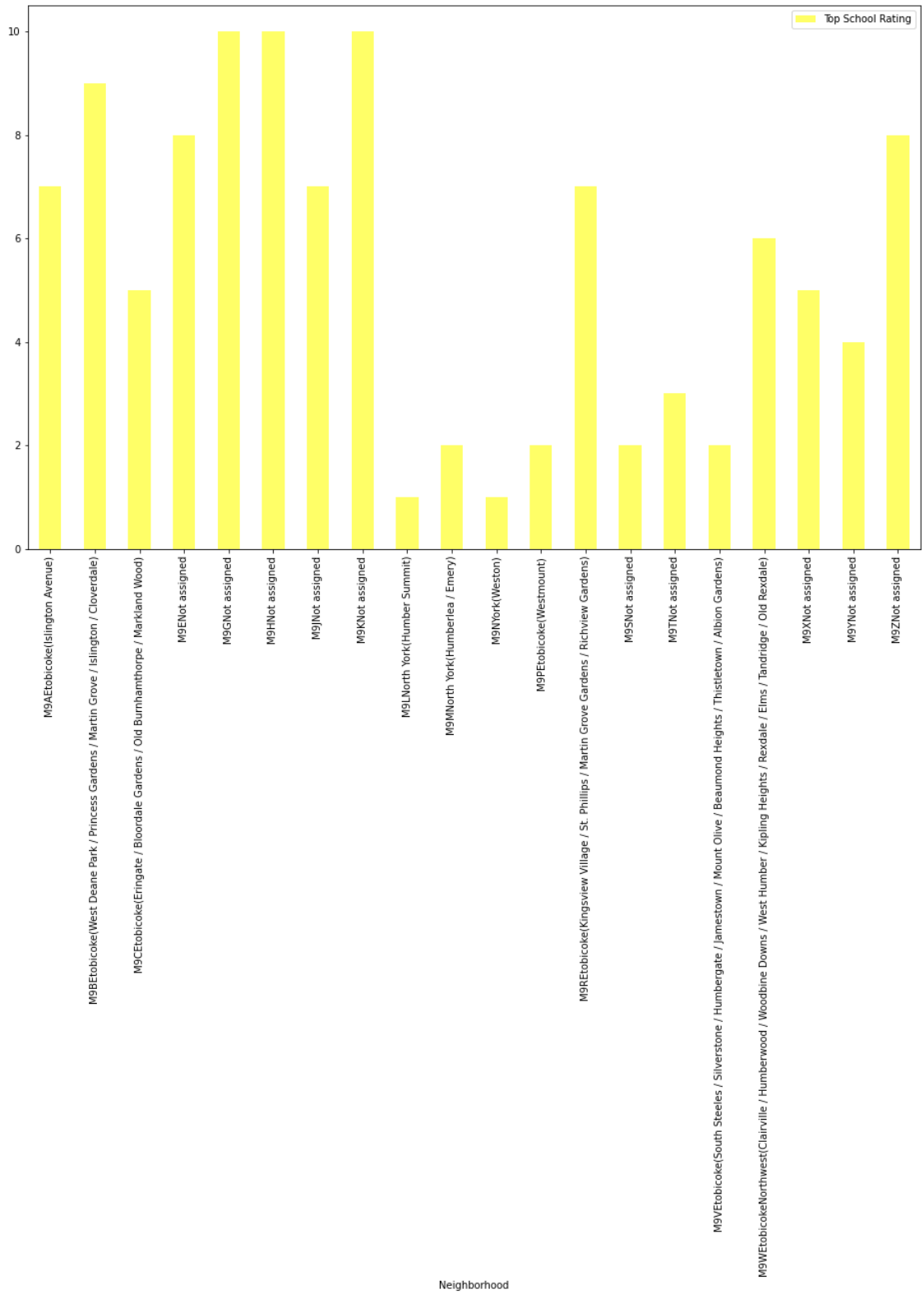
Out[73]:



Average housing price by clusters :



School Rating by clusters :



Discussion section :

The major purpose of this project, is to suggest a better neighborhood in a new city for the person who are moving there. We were able to list the most common venues near

the neighbourhoods, give an average housing price by clusters as well as a school rating by clusters.

Conclusion :

The purpose of this project was to explore Scarborough and see how attractive its neibourhood could be to potential migrants or simply people that are planning on moving there. Using k-means cluster algorithm we separated the neibourhoods into different clusters for 20 different latitude and longitude from our dataset.
This project was an opportunity for me to put my newly acquired skills in Data Science into practice and apply them to a real case. I came out of this experience motivated and determined to apply this knowledge to solve other problems.