**The Battle of Neighborhoods (week 2)**

**Introduction:**

World populations are increasingly moving from rural to urban centers, making for larger cities with greater population density than ever before. This is a global phenomenon across the spectrum of developed and developing economies. We are increasingly becoming an urban world.The data is eye-opening. The United Nations in 2009 and the International Organization for Migration in 2015 both estimated that around 3 million people are moving to cities every week. Approximately 54% of people worldwide now live in cities, up from 30% in 1950. Sources estimate this will grow to 2/3 of world population in the next 15-30 years. More than half of urban dwellers live in the 1,022 cities with greater than 500,000 inhabitants. There are currently 29 megacities with populations of over 10 million, up from 2 in 1950 and projected to grow to between 41 and 53 by 2030. Additionally, there are 468 cities with a population of over 1 million, up from 83 in 1950. A Yale research group projects that urban land coverage will expand by 463,000 square miles by 2030 to cover just under 10% of the planet’s land, equivalent to 20,000 football fields being paved over every day.

The purpose of this 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, Toronto.

Lots of people are migrating to various cities 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 etc.

The Project's aim is 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 freash 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.

**Data Section:**

Data Link: <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>

Will use Scarborough dataset which we scrapped from wikipedia on Week 3. Dataset consisting of latitude and longitude, zip codes.

**Foursquare API Data:**

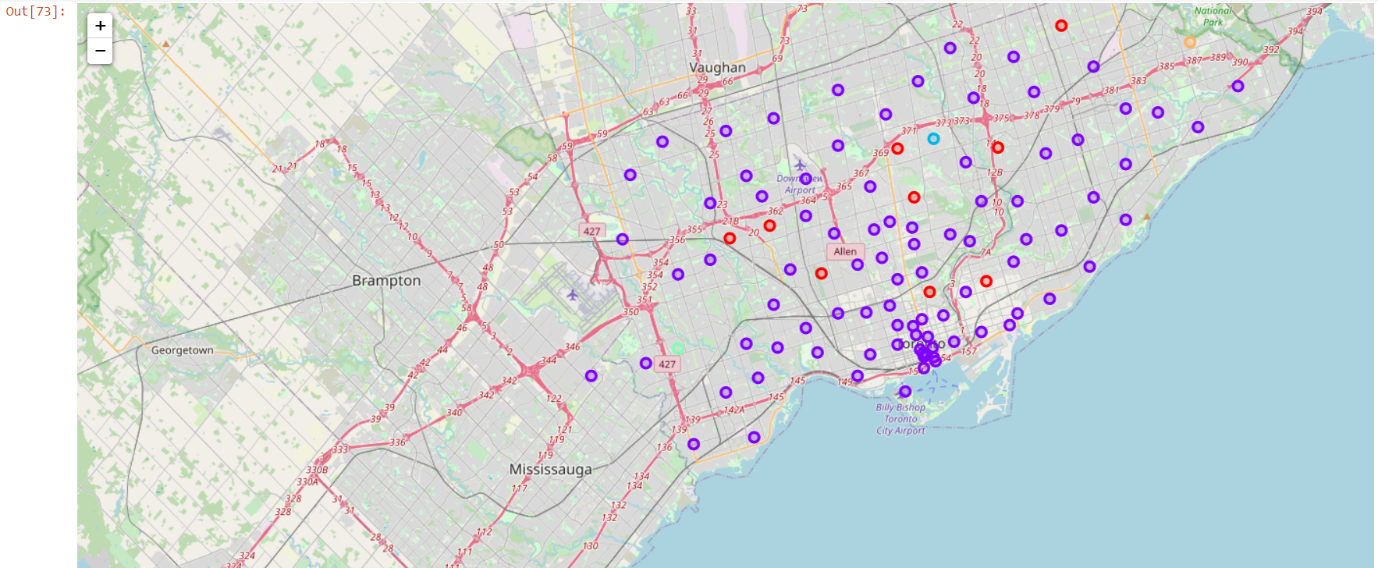
We will need data about different venues in different neighborhoods of that specific borough. In order to gain that information we will use "Foursquare" locational information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such 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

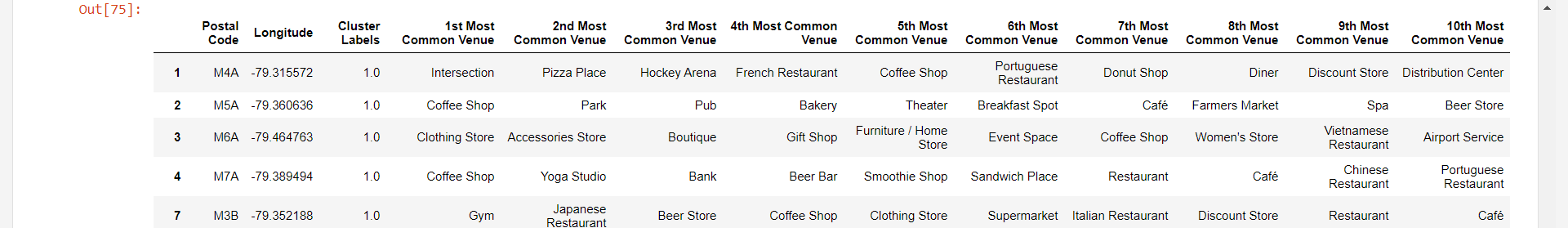
**Map of Scarborough**



**The Clustering approach:**

To compare the similarities of two cities, we should explore other neighborhoods, segment them and group them into clusters to find similar neighborhoods in a big city like Delhi and Mumbai. For that, we need to cluster data which is a form of unsupervised machine learning, i.e. k-means clustering algorithm.

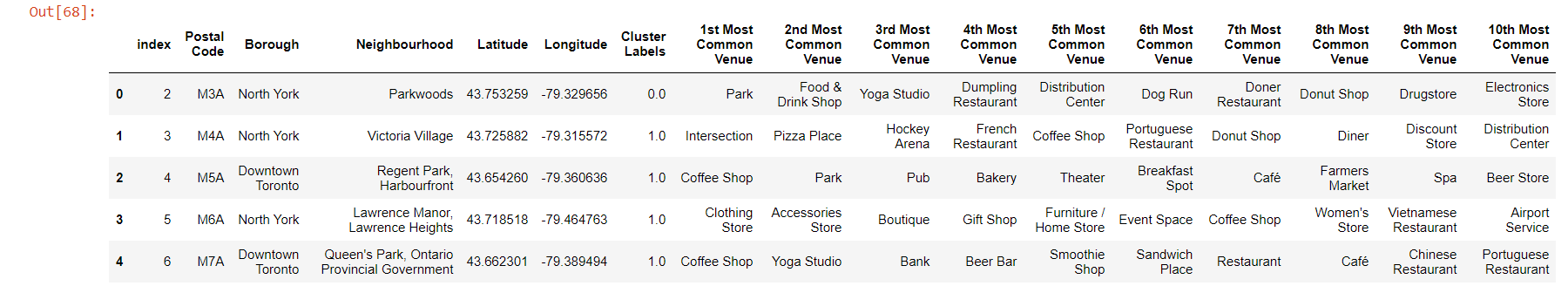
**Using K-Means Clustering Approach:**



**Most Common venues near Neighborhood:**



**Most Common venues near Neighborhood:**



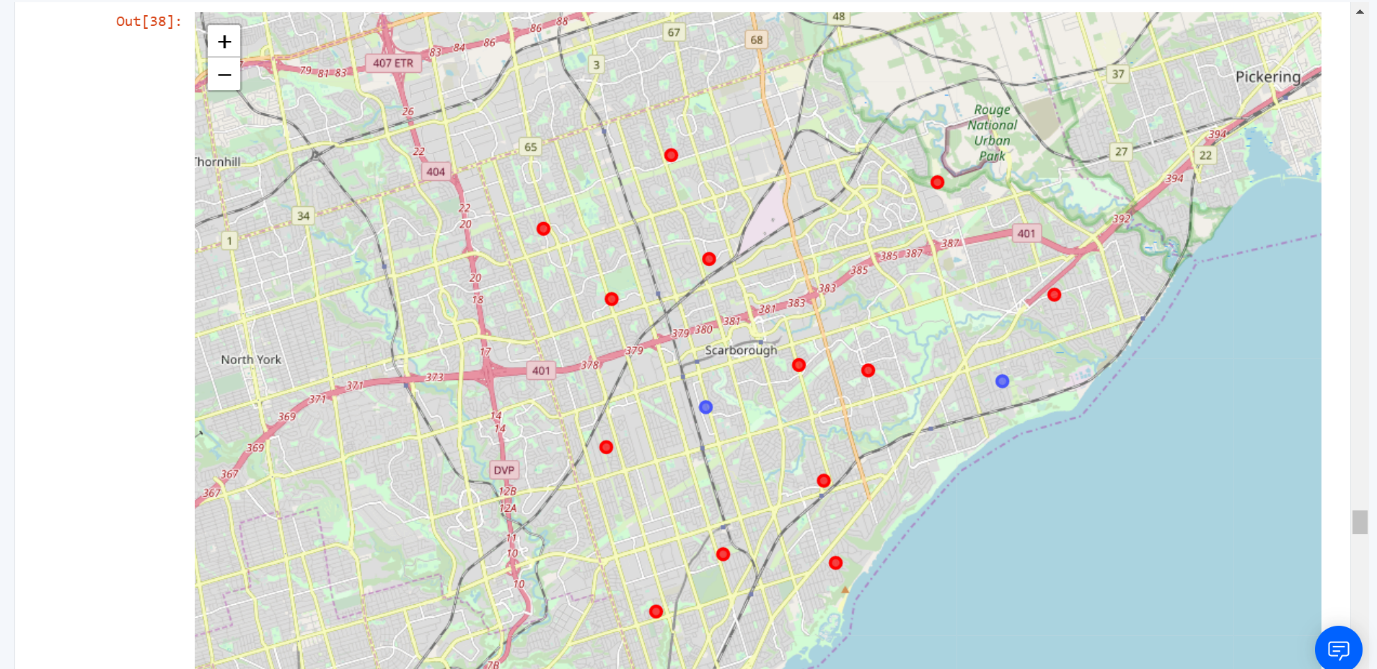
**Major Problems which are necessary to solve:**

The main purpose of this project, is to suggest a better neighborhood in a new city for the person who are shiffting there. Social presence in society in terms of like minded people. Connectivity to the airport, bus stand, city center, markets, access to the main road etc.

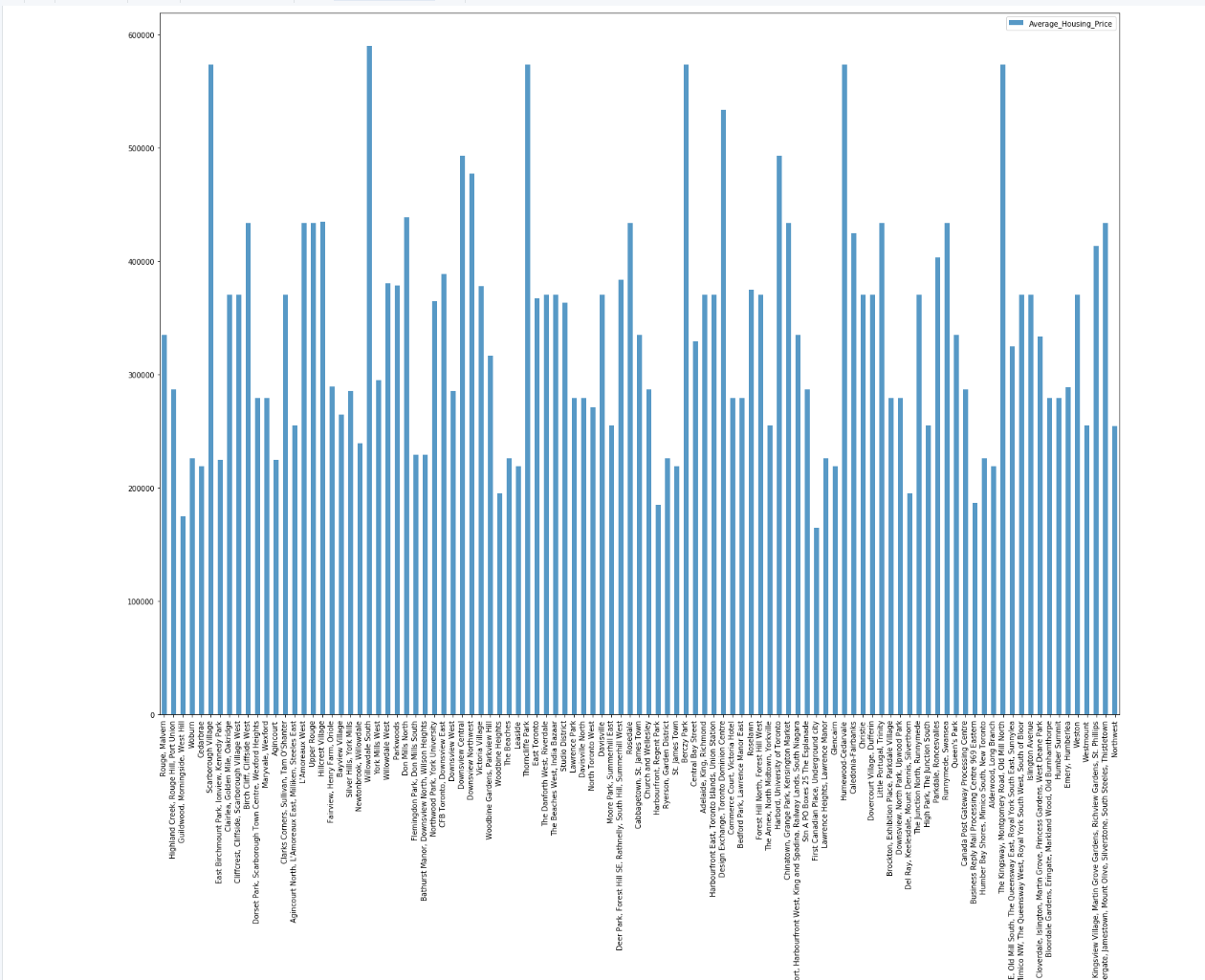
1.Sorted list of house in terms of housing prices, house size, security, reachability and traffic conjestion in a ascending or descending order 2.Sorted list of schools in terms of location, fees, transport, social activities, academic ratings etc. 3.Sorted list of Hospital, Grocery store, restraunts nearby with ratings and reviews.

**Results Section:**

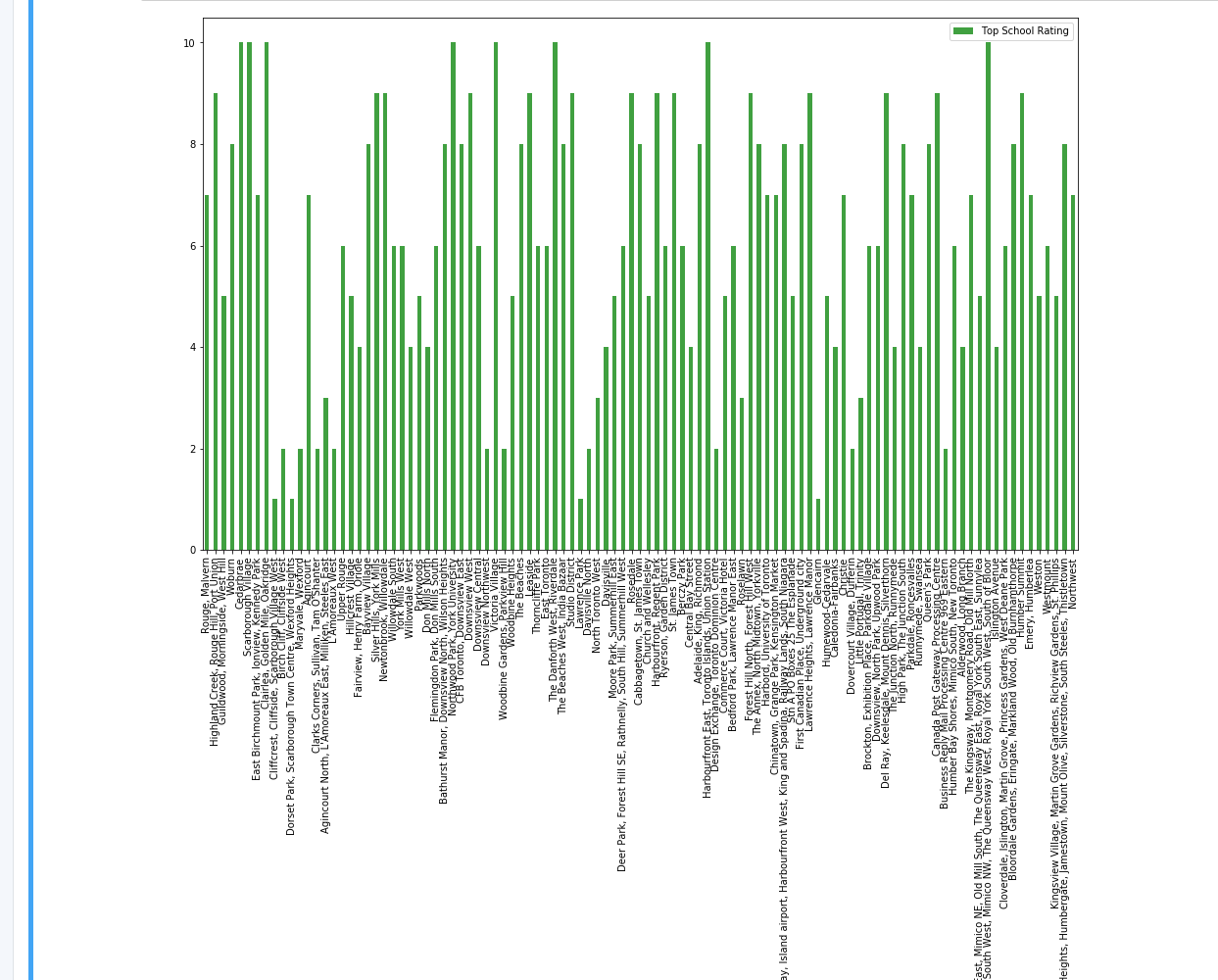
**Map of Clusters in Scarborough:**



**Average Housing Price by Clusters in Scarborough:**



**School Ratings by Clusters in Scarborough:**



**Use of Foursquare 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.

**The Workflow:**

Using credentials of Foursquare API features of near-by places of the neighborhoods would be mined. Due to http request limitations the number of places per neighborhood parameter would reasonably be set to 100 and the radius parameter would be set to 500.