# Capstone Project :- The Battle of Neighborhoods

## Business Problem:

Students who want to pursue Masters in Countries like USA, CANADA, Australia , UK etc or the parents who want to migrate to countries like above from India considers search of good rating Universities / Schools and also the good Housing Prices. So to Address this problem we predict the solution of to create an analysis of features of neighborhood cities in a state of Canada between Neighborhoods. The features includes School ratings , median Housing price, weather conditions , crime rate etc. This will help students to figure out which country to choose , which universities to plan , what will the approx budget before going to that place where they are gonna start their new life.

The aim of the Project is to help students/ parents explore different possibilities and take better decision on choosing the best neighborhood out of many neighborhoods in Scarborough city based on the various features in and around that neighborhood.

## Selection Criteria of Neighborhood:

In this project we define good neighborhood by the good Housing price and the good university / schools ratings. We can consider more parameter to like low crime rates in that area , good weather conditions.

1.  Compare School ratings
2.  Compare Median of House Prices

## The Location : Scarborough:

Scarborough is a place which has good schools and universities so students are interested going there . even its is great destination for immigrants in Canada. 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.

## Use of FourSquare API:

The project will Use FourSquare API as source of the data of places as it has millions of places. They even provide ability to perform location search , location sharing and business details.

We will Use 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.

# K-Means Clustering:

We will 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 Toronto. To be able to do that, we need to cluster data which is a form of unsupervised machine learning: k-means clustering algorithm

## Work Flow:
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. Steps taken were:
1. Data acquisition and cleansing
2. Data preparation
3. Feature selection
4. Clustering

## Data acquisition and cleansing:
Data acquisition was a 2-step process:
1. Obtaining the postcodes for neighborhoods in Toronto
2. Obtaining venues within these neighborhoods

## Clustering Approach:
1. 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

## Neighborhood Most Common Venus:

| | 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 Mo Comm Venu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | M1B | Scarborough | Rouge, Malvern | 43.811525 | -79.195517 | 0 | Zoo Exhibit | Paintball Field | Fast Food Restaurant | Construction & Landscaping | Business Service | Automotive Shop | Departme Sto |
| 1 | M1C | Scarborough | Highland Creek, Rouge Hill, Port Union | 43.785730 | -79.158750 | 2 | Bar | Construction & Landscaping | Department Store | Caribbean Restaurant | Chinese Restaurant | Coffee Shop | Colleg Stadiu |
| 2 | M1E | Scarborough | Guildwood, Morningside, West Hill | 43.765690 | -79.175256 | 0 | Park | Athletics & Sports | Gym | Gym / Fitness Center | Zoo Exhibit | Deli / Bodega | Caribbe Restaura |
| 3 | M1G | Scarborough | Woburn | 43.768359 | -79.217590 | 0 | Coffee Shop | Park | Fast Food Restaurant | Business Service | Convenience Store | Cantonese Restaurant | Caribbe Restaura |
| 4 | M1H | Scarborough | Cedarbrae | 43.769688 | -79.239440 | 0 | Caribbean Restaurant | Lounge | Athletics & Sports | Flower Shop | Thai Restaurant | Bakery | Bar |

# Neighborhood Median Housing Prices:
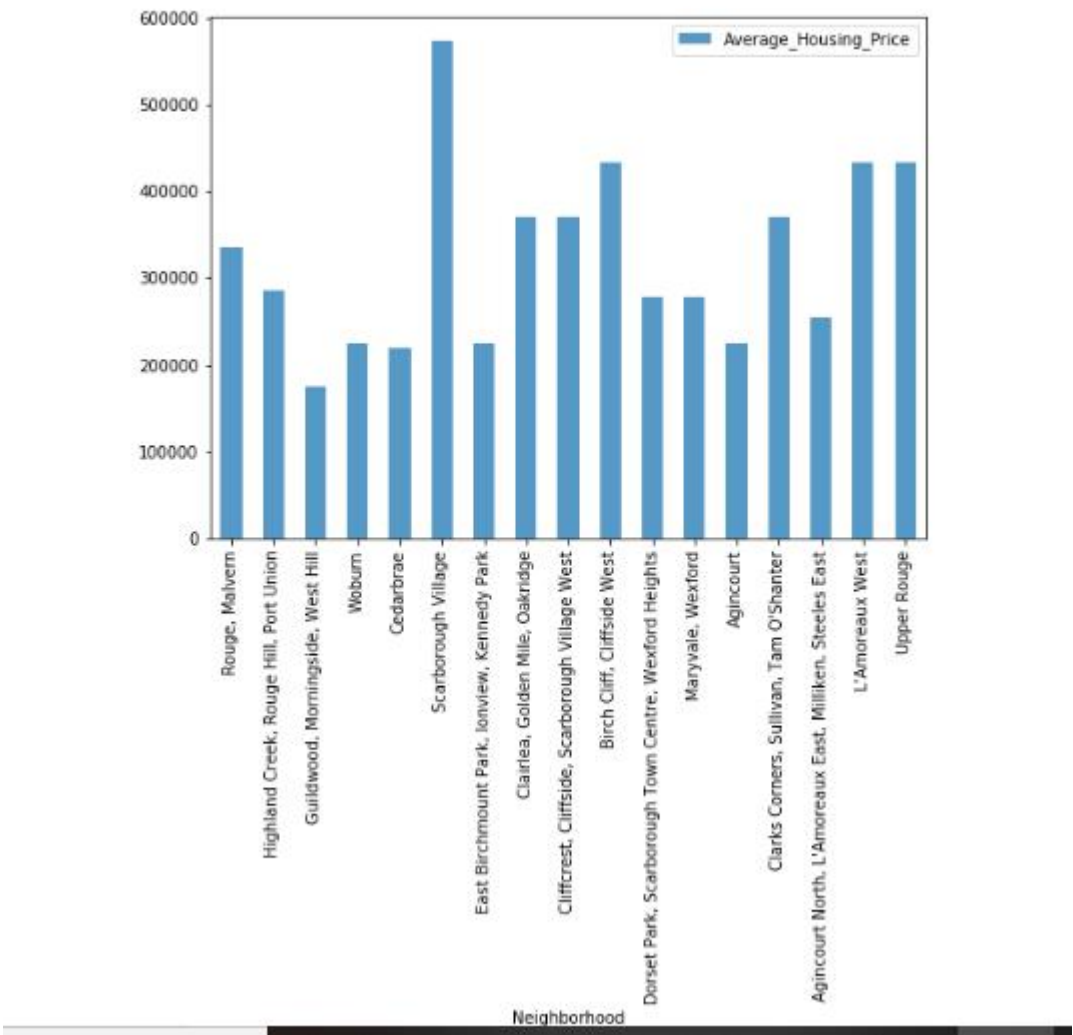
## Neighborhood School Ratings:
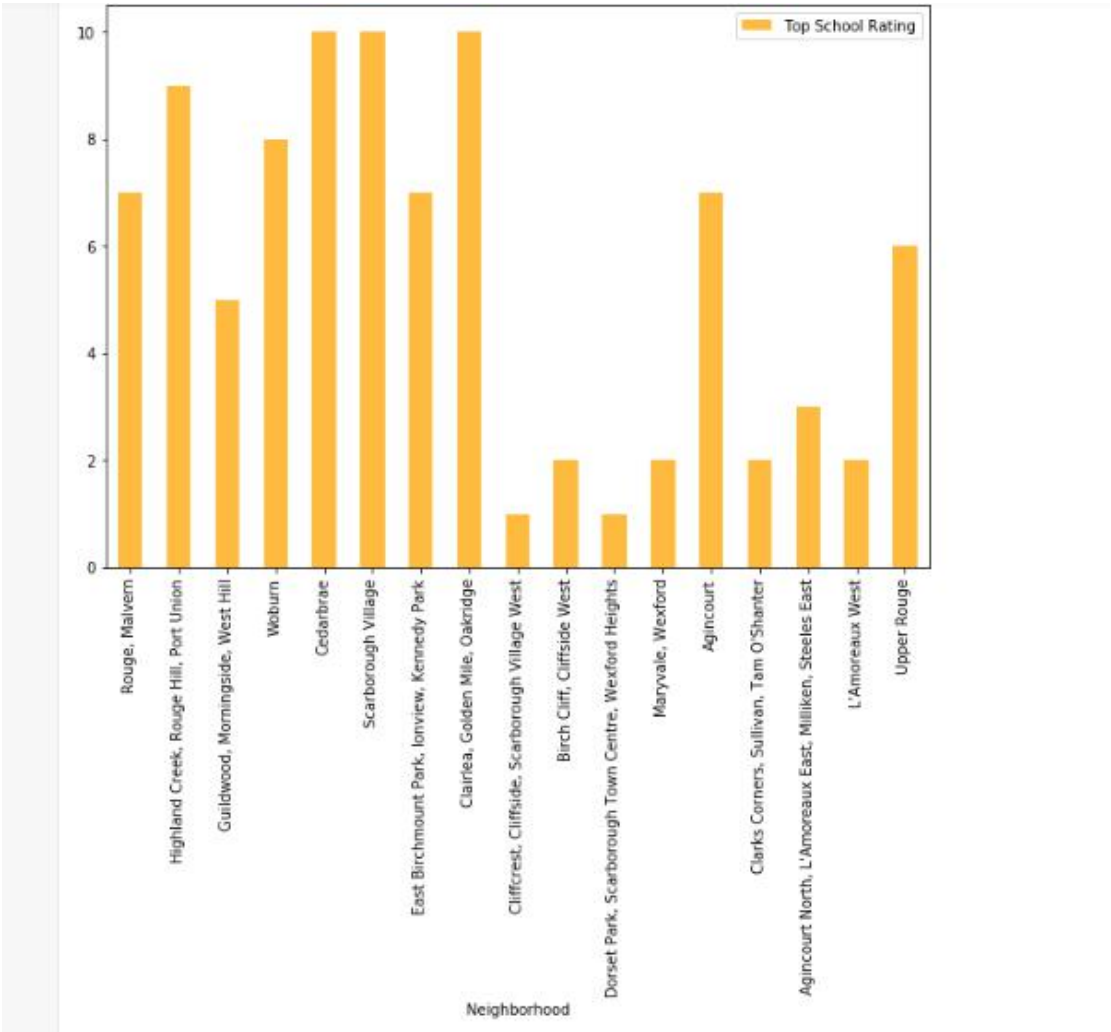
# Conclusion:

In this project, through a k-means cluster algorithm we separate the neighborhood into 03 clusters, which have similar neighborhoods around them. Using the charts above decision leading to a particular neighborhood based on average house prices and school rating can be made