Capstone
The Battle of Neighbourhoods
Analysing Median House Prices and School Ratings for Scarborough Canada
for Immigrants

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.

Project Description:

Due to various reasons like Jobs, Business many people migrating to various states of Canada and it requires search of a good housing prices as well as good rating schools for their children.

This projects aim is to create an analysis of features for a neighbourhood as a comparative analysis between neighbourhoods. The features include median house price and school ratings, crime rates, weather conditions, recreational facilities. This would help people to get awareness of the places before moving to a new country, state, city or place for their work or to start a new life

Also this project helps people explore different possibilities and take a better decision on choosing the best neighbourhood out of many neighbourhoods in **Scarborough** city based on the distribution of various facilities in and around that neighbourhood.

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.

Work Flow:

Using credentials of Foursquare API features of near-by places of the neighbourhoods would be mined. Due to http request limitations the number of places per neighbourhood parameter would reasonably be set to 100 and the radius parameter would be set to 700.

1. Data acquisition and cleansing
2. Data preparation
3. Feature selection
4. Clustering

Data acquisition and cleansing

Data acquisition is a 2 step process
1. Getting the Postal Codes for Neighbourhoods in Toronto
2. Getting venues within these Neighbourhoods

Clustering Approach:

To compare the similarities of two cities, we decided to explore neighbourhoods, segment them, and group them into clusters to find similar neighbourhoods 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.

Neighbourhoods Most Common Venues

	Neighborhood	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	Agincourt	Coffee Shop	Skating Rink	Convenience Store	Motorcycle Shop	Clothing Store	Pool Hall	Lounge	Sandwich Place	Latin American Restaurant	Shanghai Restaurant
1	Birch Cliff, Cliffside West	Park	College Stadium	Skating Rink	Diner	Café	General Entertainment	Thai Restaurant	Hakka Restaurant	Fish Market	Department Store
2	Cedarbrae	Coffee Shop	Indian Restaurant	Athletics & Sports	Caribbean Restaurant	Hakka Restaurant	Gas Station	Fried Chicken Joint	Flower Shop	Pizza Place	Playground
3	Clarks Corners, Tam O'Shanter, Sullivan	Intersection	Pharmacy	Pizza Place	Seafood Restaurant	Italian Restaurant	Gas Station	Fried Chicken Joint	Fast Food Restaurant	Convenience Store	Noodle House
4	Cliffside, Cliffcrest, Scarborough Village West	Wings Joint	Hardware Store	Chinese Restaurant	Restaurant	Fried Chicken Joint	Diner	Discount Store	Electronics Store	Fast Food Restaurant	Fish Market

Average Housing Price
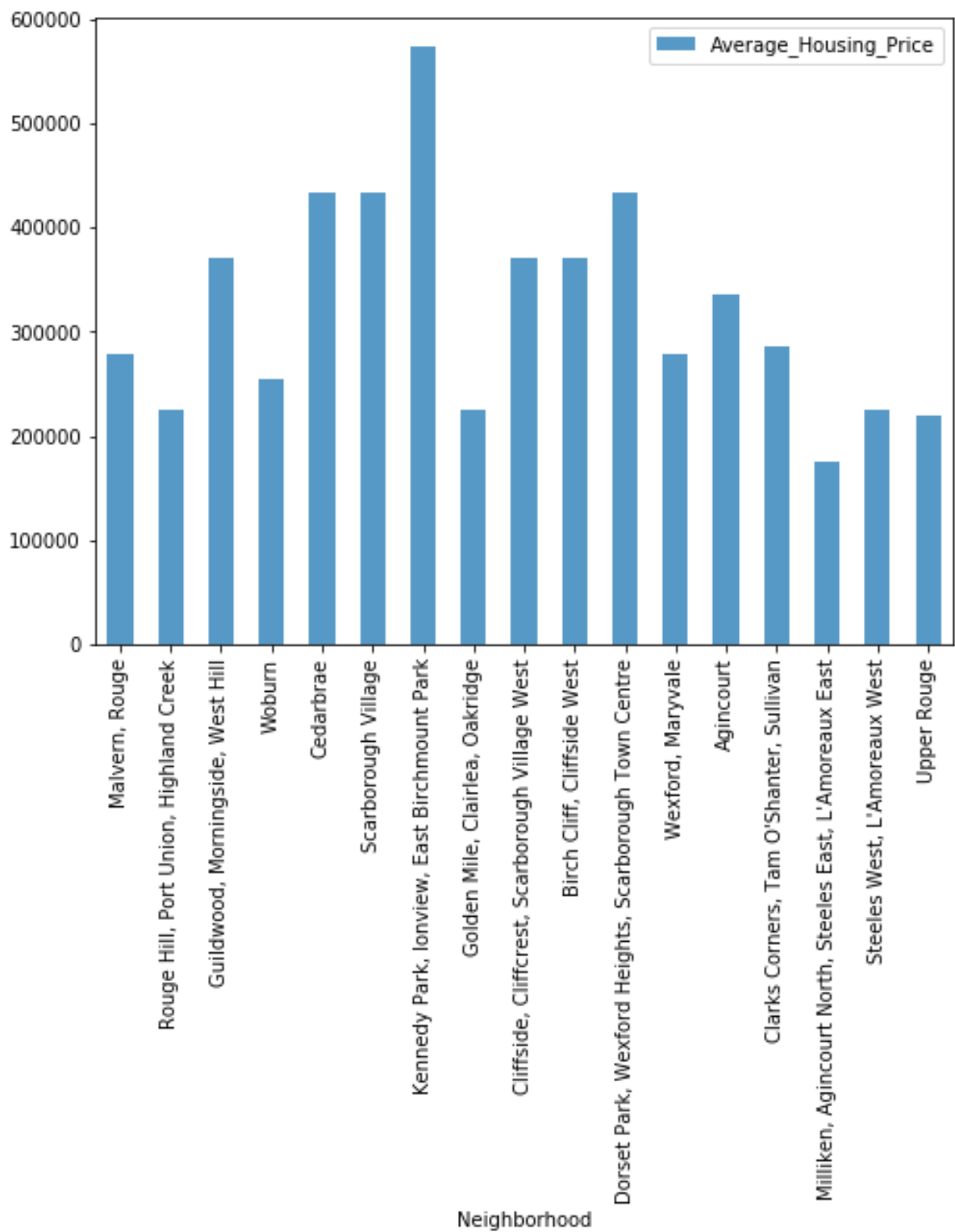
Neighbour School Ratings
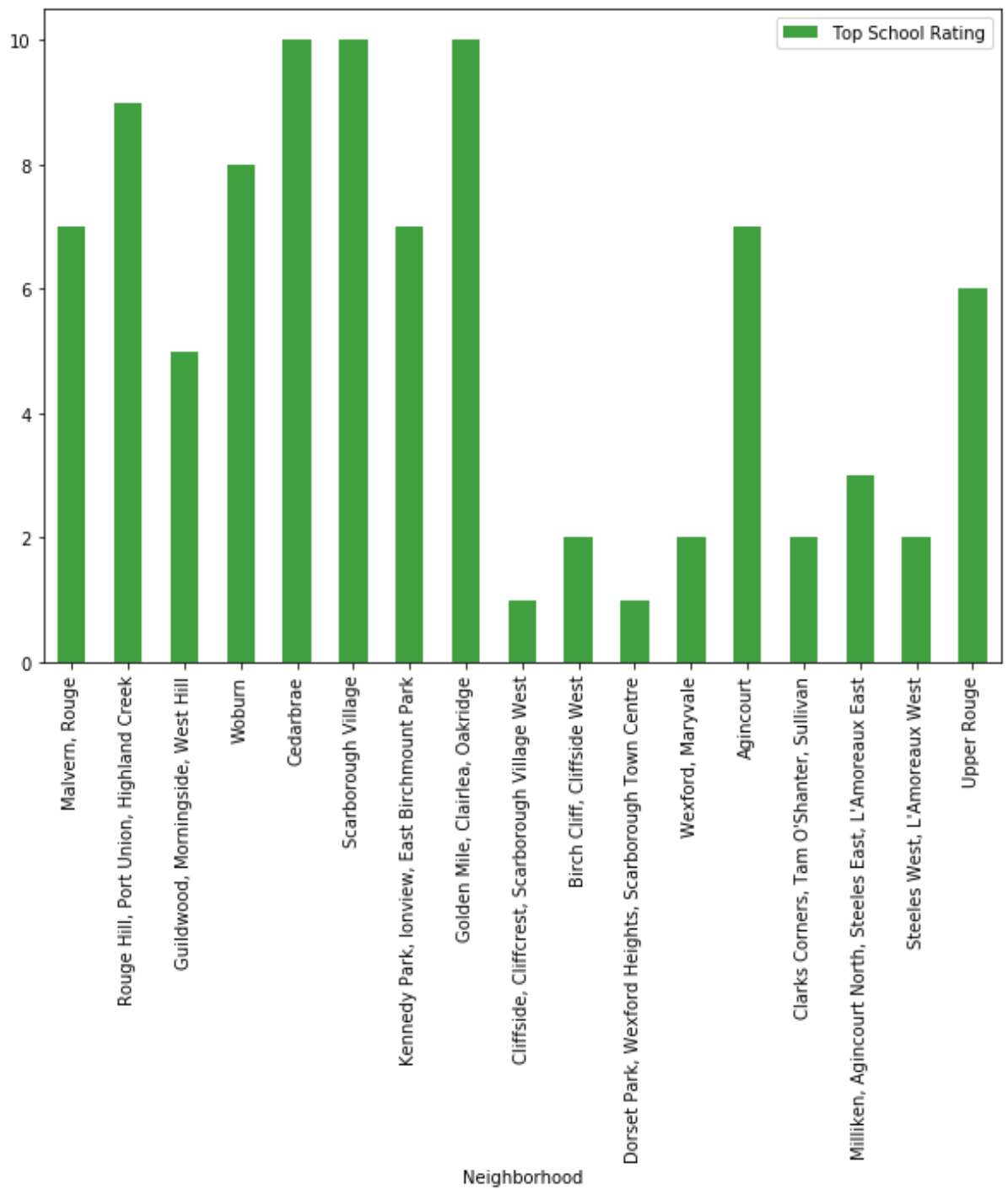
Conclusion:

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