Final Report – Capstone Project
The Battle of Neighborhoods Finding a Better Place in North York, Toronto
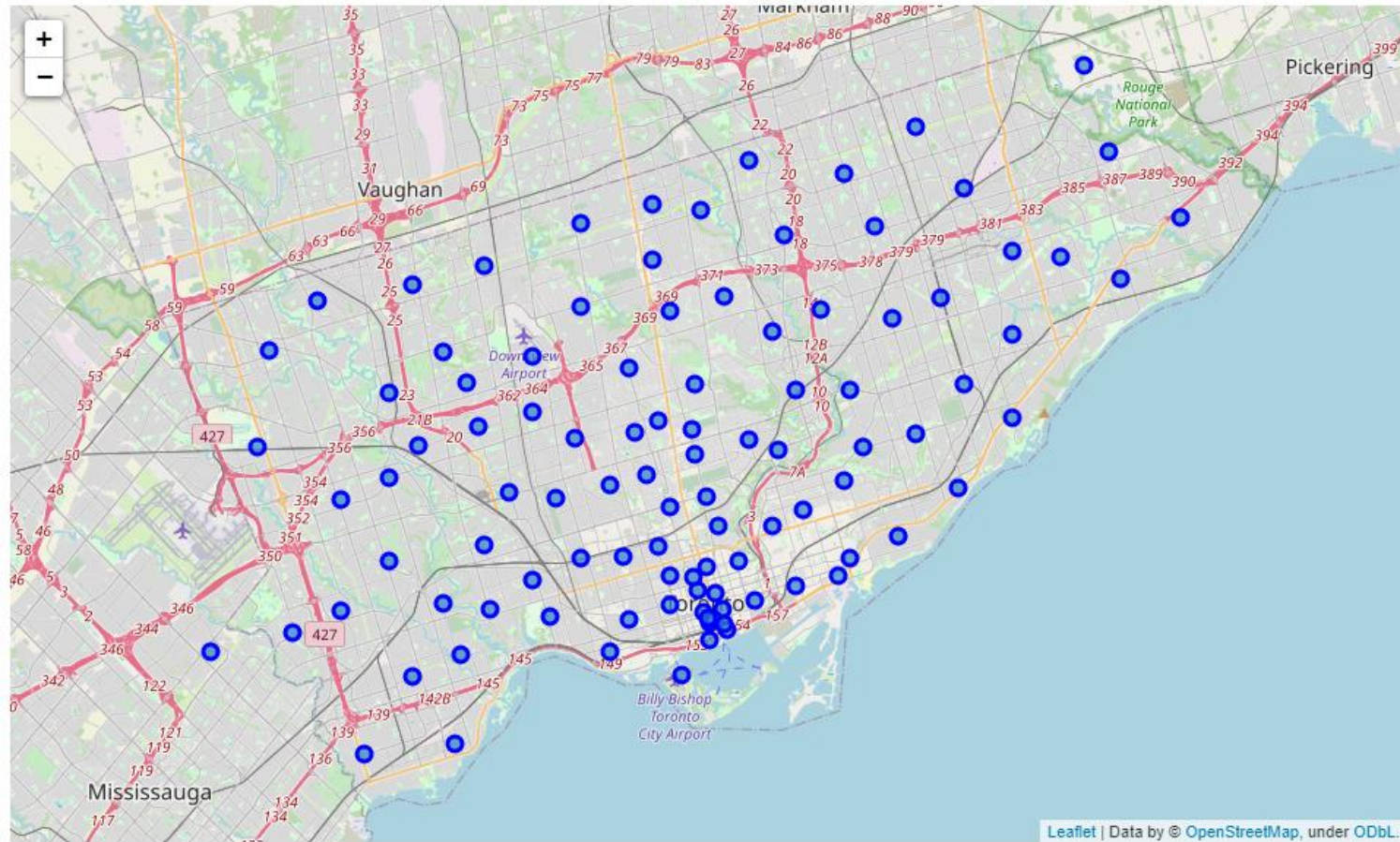
Introduction

- The goal 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 North York, Toronto.

- This project aim to create an analysis of features for a people migrating to North York to search a best neighborhood as a comparative analysis between neighborhoods for example median housing price, 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.

Data

- Data source: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

- Using North York dataset which we scrapped from Wikipedia on Week 3. Dataset consisting of latitude and longitude, and 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. 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

Data

Map of North York, Toronto

Methodology

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.

Methodology

Using K-Means Clustering Approach

Results and Discussion

The Location:

North York is an eclectic, multicultural district home to the hands-on Ontario Science Centre and the Aga Khan Museum, with exhibits on Islamic culture in a striking modern building. In the area's north, Black Creek Pioneer Village is an 1800s living museum. Sprawling Downsview Park includes a lake, event spaces, and a flea and farmers' market, while Edwards Gardens has a greenhouse, fountains, and botanic gardens.

North York 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.
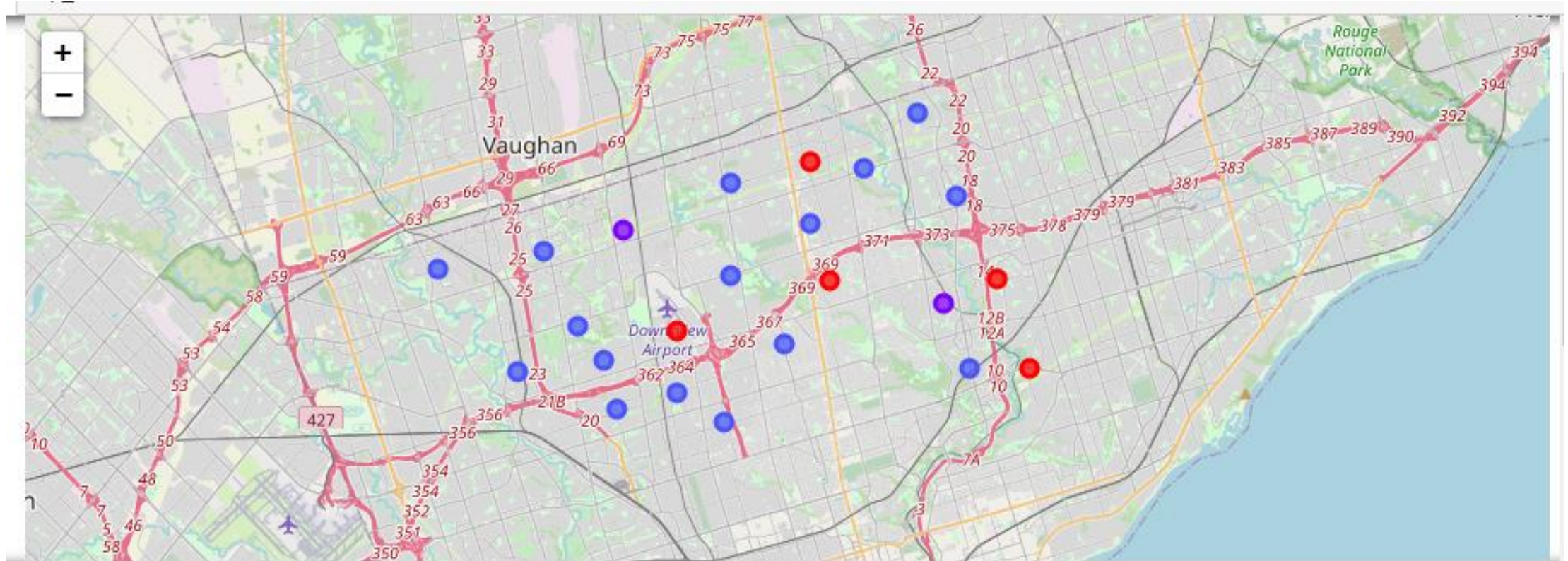
Foursquare API:

This project has used 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 major purpose of this project is to suggest a better neighborhood in a new city for the person who are shifting there. Connectivity to the airport, bus stand, city center, markets and other daily needs things nearby.

1. Sorted list of houses in terms of housing prices in ascending or descending order
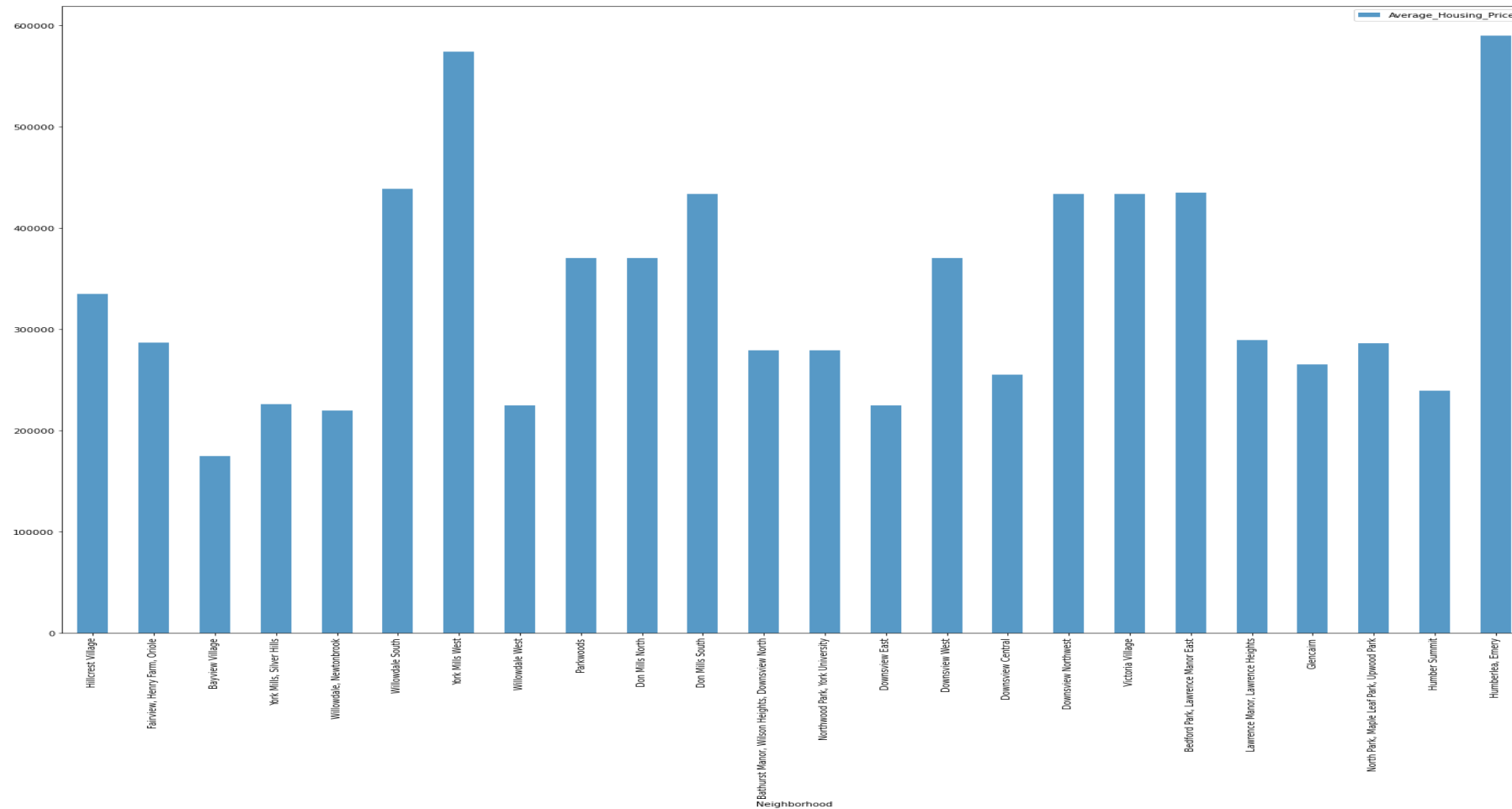2. Sorted list of schools in terms of location, fees, rating and reviews

Results and Discussion

Map of Clusters in North York

Results and Discussion

Average Housing Price by Clusters in North York

Conclusions

In this project, using k-means cluster algorithm I separated the neighborhood into 10 different clusters and for 24 different coordinates from dataset, which have very-similar neighborhoods around them. Using the charts above results presented to a particular neighborhood based on average house prices and school rating have been made.