The Battle of Neighbourhoods Project

Finding the perfect neighbourhood in Toronto Area

Introduction:

- This project is aimed at creating a detailed analysis report of the neighborhoods in toronto to help people with moving to toronto. The features will include median housing prices and better schools according to ratings, crime rates of that particular area, road connectivity, weather conditions, good management for emergency, water resources and recreational facilities.

Problems that are assesed in this project:

The major purpose of this project, is to suggest a better neighborhood in a the city of Toronto, social presence in society in terms of like minded people, connectivity to the airport, bus stand, city center, markets, and proximity of other daily needs things.

- Sorted list of house in terms of housing prices in a ascending or descending order
- Sorted list of schools in terms of location, fees, rating and reviews

Libraries and resources:

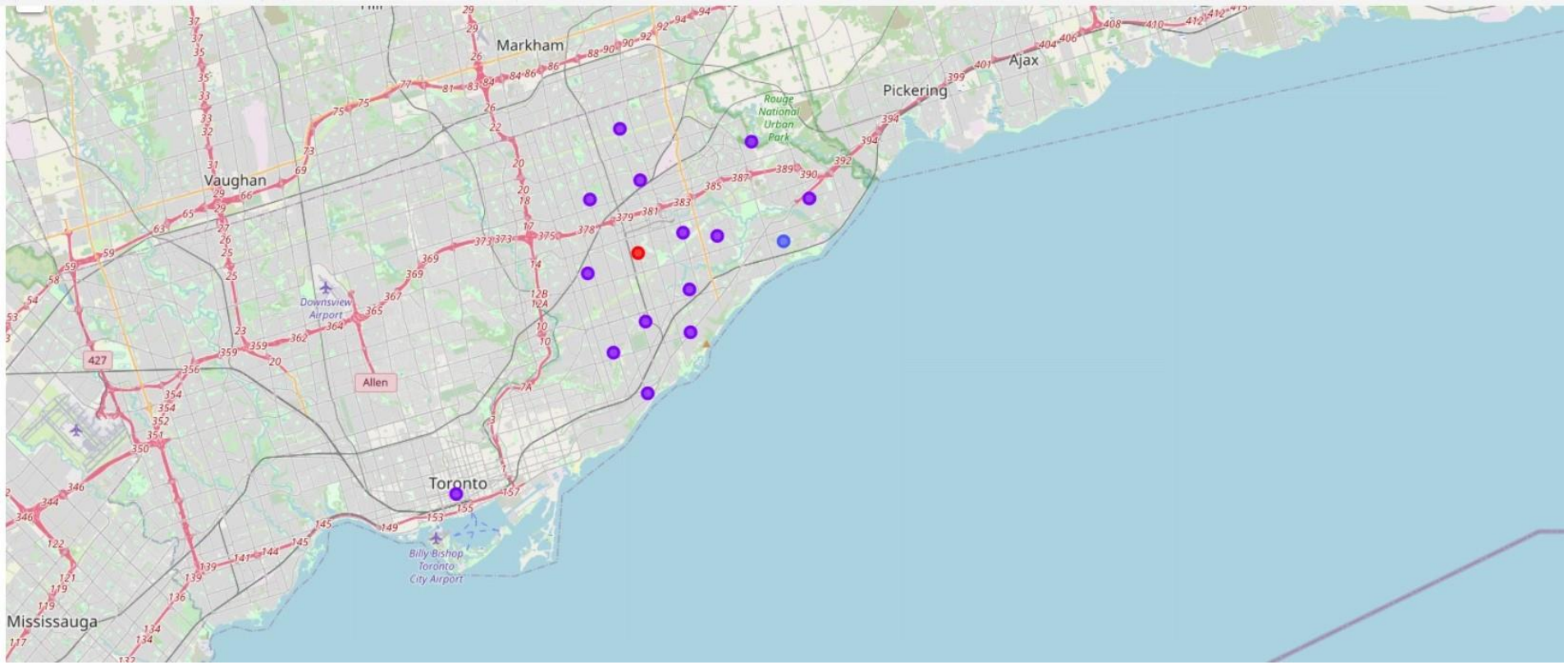
- Pandas: For creating and manipulating dataframes.
- Folium: Python visualization library would be used to visualize the neighborhoods cluster distribution of using interactive leaflet map.
- Scikit Learn: For importing k-means clustering.
- JSON: Library to handle JSON files.
- XML: To separate data from presentation and XML stores data in plain text format.
- Geocoder: To retrieve Location Data.
- Beautiful Soup and Requests: To scrap data and library to handle http requests.
- Matplotlib: Python Plotting Module.

Foursquare API:

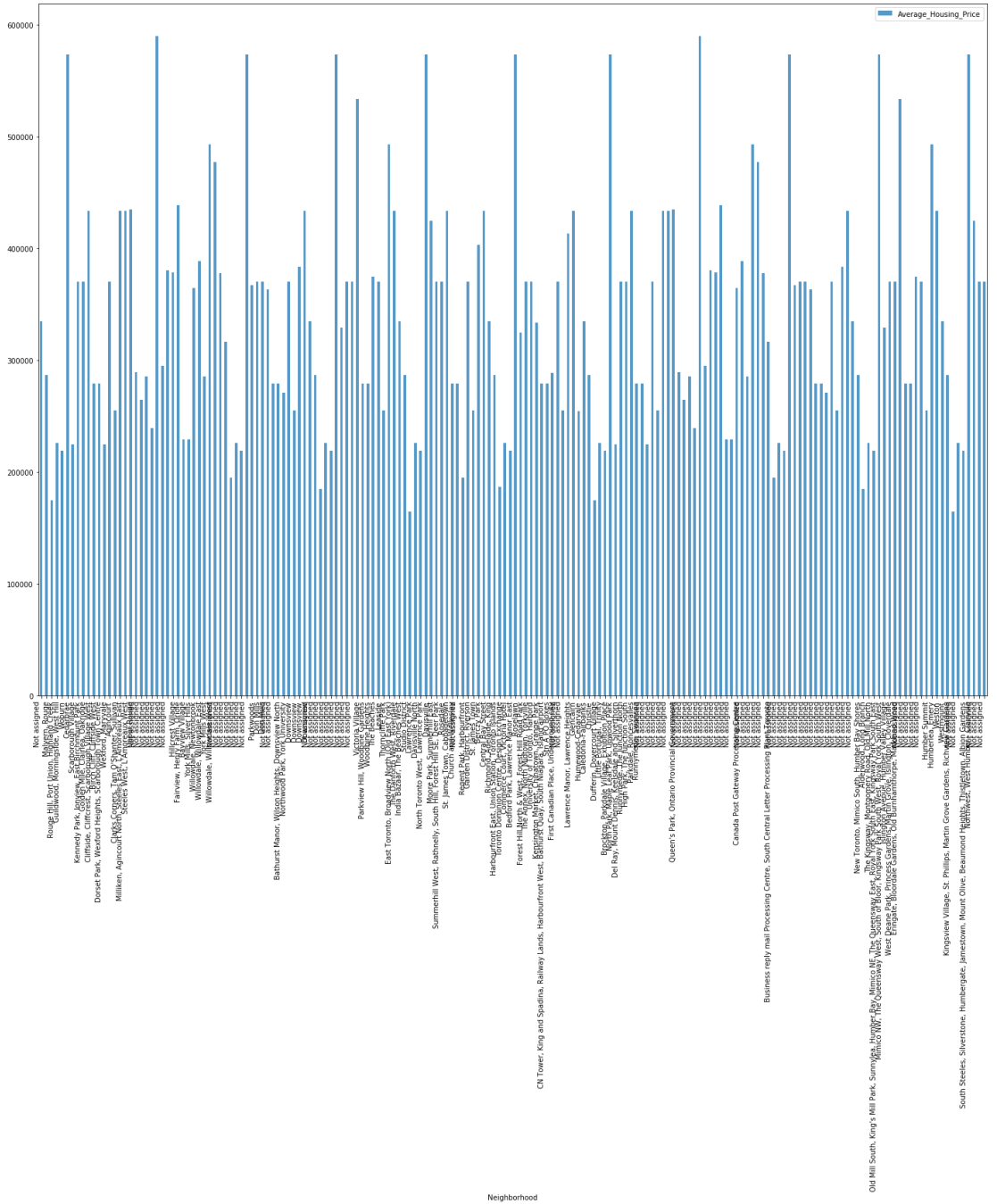
- After finding the list of neighborhoods, we then connect to the Foursquare API to gather information about venues inside each and every neighborhood and the radius is limited to 100 mts.
- The following data is retrieved from Foursquare API:
- 1. Neighborhood
- 2. Neighborhood Latitude
- 3. Neighborhood Longitude
- 4. Venue and Name
- 5. Venue Latitude
- 6. Venue Longitude
- 7. Venue Category

Results

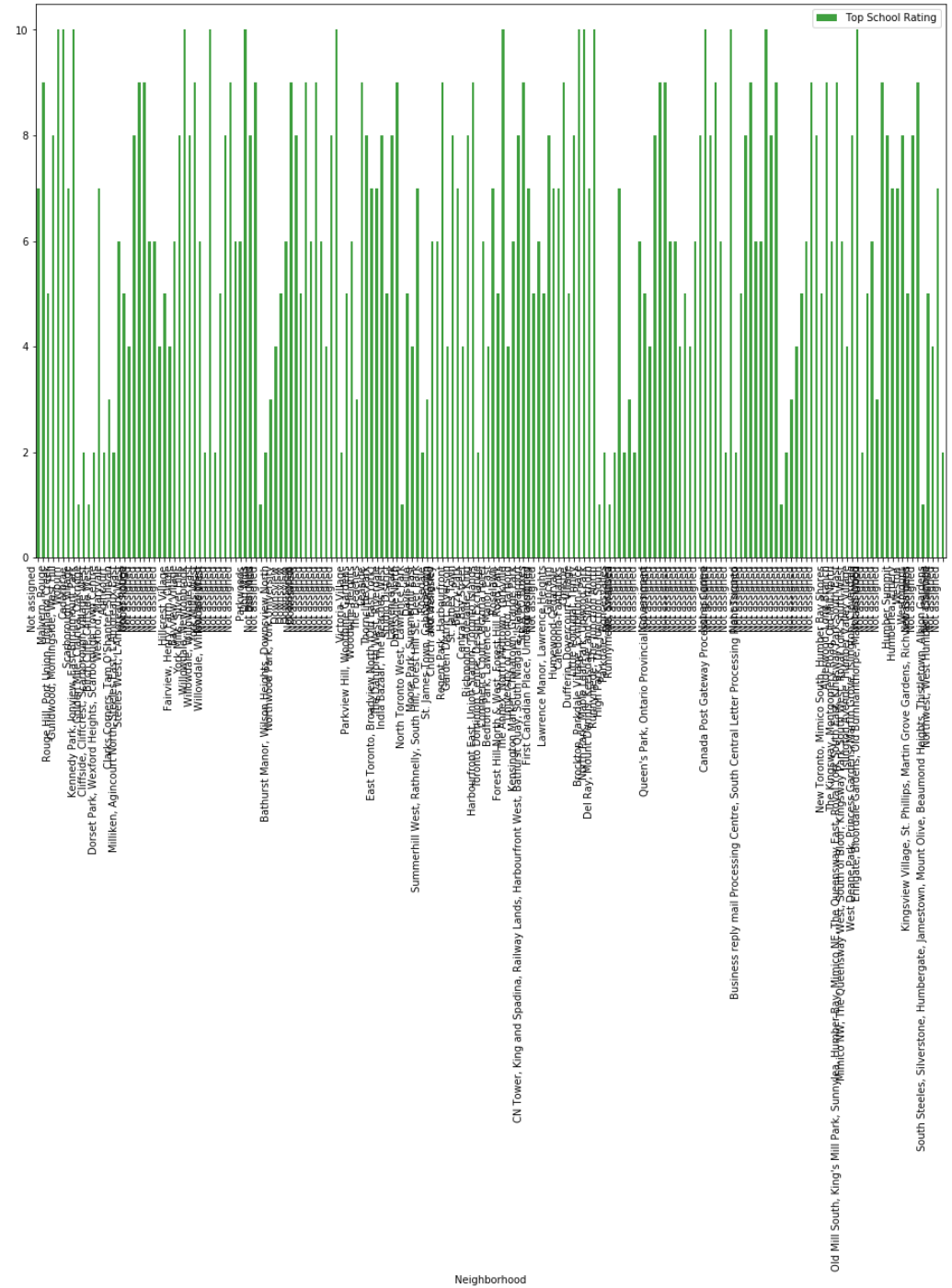
- Clustered map of Toronto area



Average housing prices



Average School ratings



Conclusion

- In this project, using k-means cluster algorithm I separated the neighborhood into 10(Ten) different clusters and for 180 different lattitude and logitude 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.