The Battle of Neighborhoods: |Finding a Better Place in Scarborough, Toronto

1. 1. Introduction:

2. The purpose of this Capstone 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, Toranto.

3. Lots of people are migrating to various states of Canada 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, like minded people, etc.

4. This Capstone Project aim 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.

5. 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 or to start a new fresh life.

2. 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
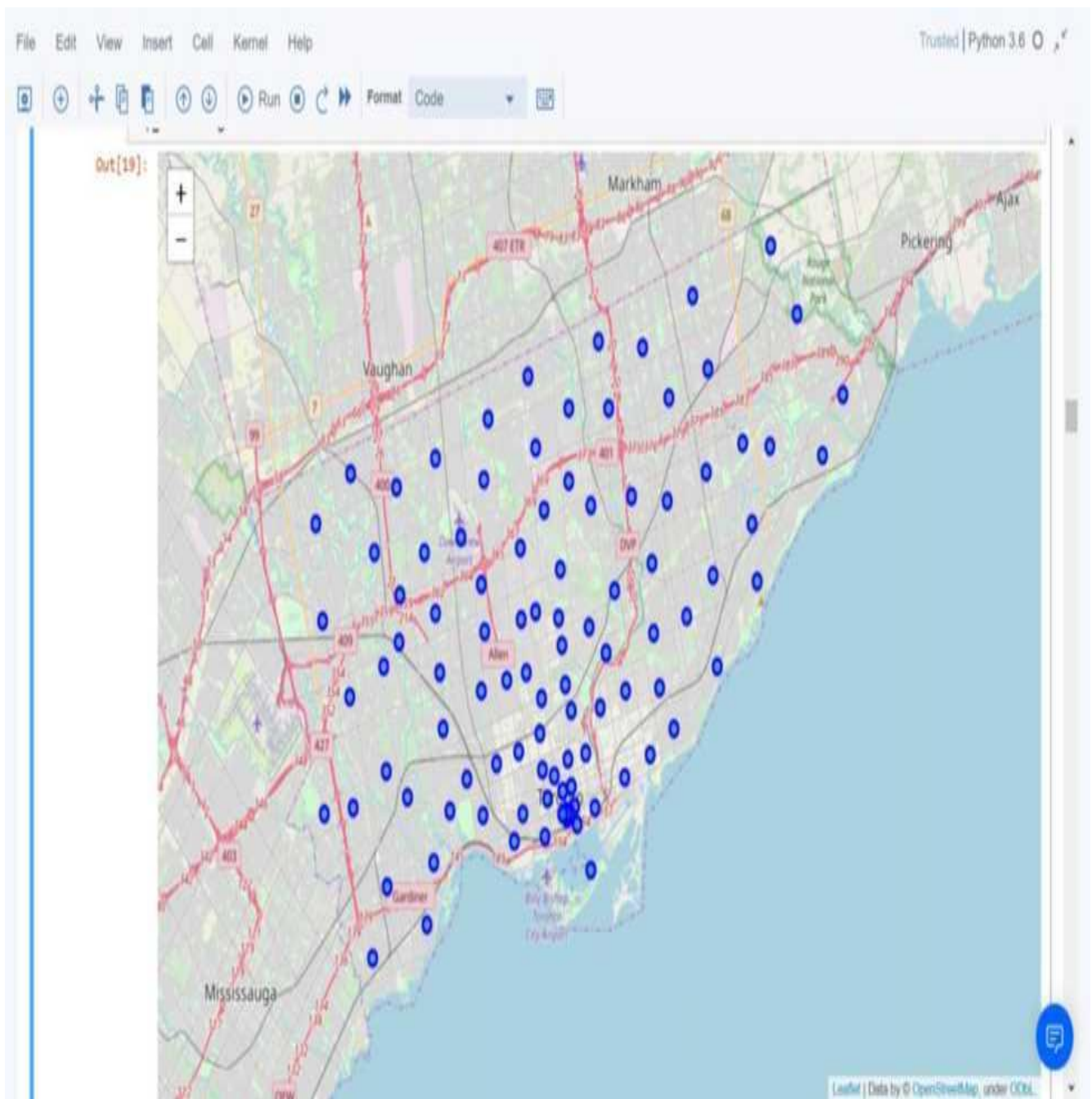
3. Methodology Section

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.

Using K-Means Clustering

```
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

neighborhoods_venues_sorted = pd.DataFrame(columns=columns)
neighborhoods_venues_sorted['Neighborhood'] = Scarborough_grouped['Neighborhood']

for ind in np.arange(Scarborough_grouped.shape[0]):
    neighborhoods_venues_sorted.iloc[ind, 1:] = return_most_common_venues(Scarborough_grouped.iloc[ind, :], num_top_venues)

neighborhoods_venues_sorted.head()
```

[129]:

	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	Shopping Mall	Bank	Clothing Store	Breakfast Spot	Skating Rink	Latin American Restaurant	Café	Supermarket	Sandwich Place	Sushi Restaurant
1	Alderwood, Long Branch	Sandwich Place	Gas Station	Pizza Place	Print Shop	Pub	Coffee Shop	Gym	Dance Studio	Dumpling Restaurant	Dive Bar
2	Bathurst Manor, Wilson Heights, Downsview North	Pizza Place	Coffee Shop	Men's Store	Park	Sandwich Place	Mediterranean Restaurant	Intersection	Sushi Restaurant	Restaurant	Fried Chicken Joint
3	Bayview Village	Park	Flower Shop	Locksmith	Trail	Yoga Studio	Doner Restaurant	Donut Shop	Dumpling Restaurant	Eastern European Restaurant	Electronics Store
4	Bedford Park, Lawrence Manor East	Pet Store	Pizza Place	Sandwich Place	Coffee Shop	Italian Restaurant	Indian Restaurant	Pharmacy	Liquor Store	Sports Club	Butcher

Approach | Most Common Venue

Most Common Venues near Neighborhood | Using Clustering

```python
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

neighborhoods_venues_sorted = pd.DataFrame(columns=columns)
neighborhoods_venues_sorted['Neighborhood'] = Scarborough_grouped['Neighborhood']

for ind in np.arange(Scarborough_grouped.shape[0]):
    neighborhoods_venues_sorted.iloc[ind, 1:] = return_most_common_venues(Scarborough_grouped.iloc[ind, :], num_top_venues)

neighborhoods_venues_sorted.head()
```

[129]:

|	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	Shopping Mall	Bank	Clothing Store	Breakfast Spot	Skating Rink	Latin American Restaurant	Café	Supermarket	Sandwich Place	Sushi Restaurant
1	Alderwood, Long Branch	Sandwich Place	Gas Station	Pizza Place	Print Shop	Pub	Coffee Shop	Gym	Dance Studio	Dumpling Restaurant	Dive Bar
2	Bathurst Manor, Wilson Heights, Downsview North	Pizza Place	Coffee Shop	Men's Store	Park	Sandwich Place	Mediterranean Restaurant	Intersection	Sushi Restaurant	Restaurant	Fried Chicken Joint
3	Bayview Village	Park	Flower Shop	Locksmith	Trail	Yoga Studio	Doner Restaurant	Donut Shop	Dumpling Restaurant	Eastern European Restaurant	Electronics Store
4	Bedford Park, Lawrence Manor East	Pet Store	Pizza Place	Sandwich Place	Coffee Shop	Italian Restaurant	Indian Restaurant	Pharmacy	Liquor Store	Sports Club	Butcher

4. Results Section

Map of Clusters in Scarborough
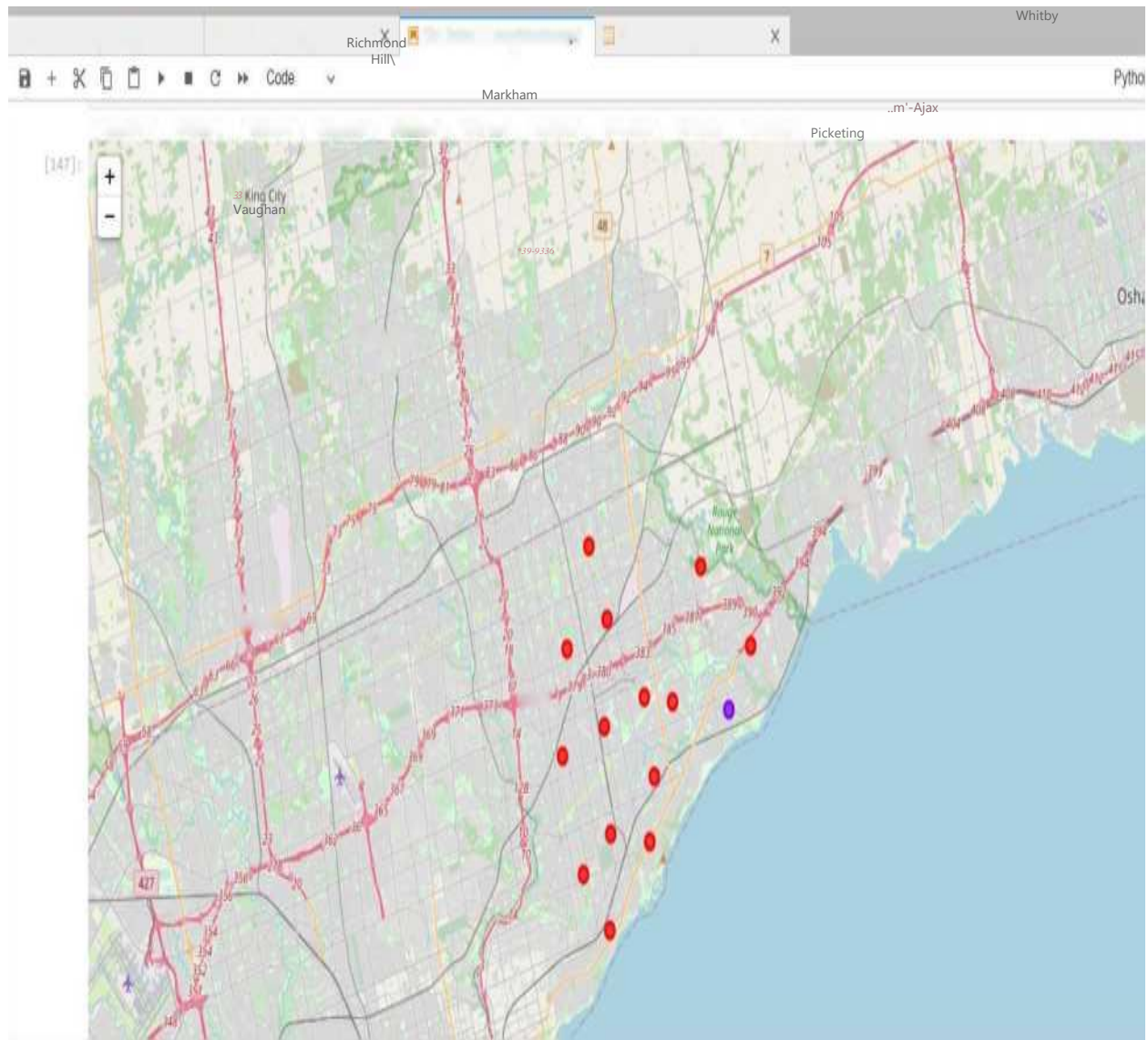
Average Housing Price by Clusters in Scarborough
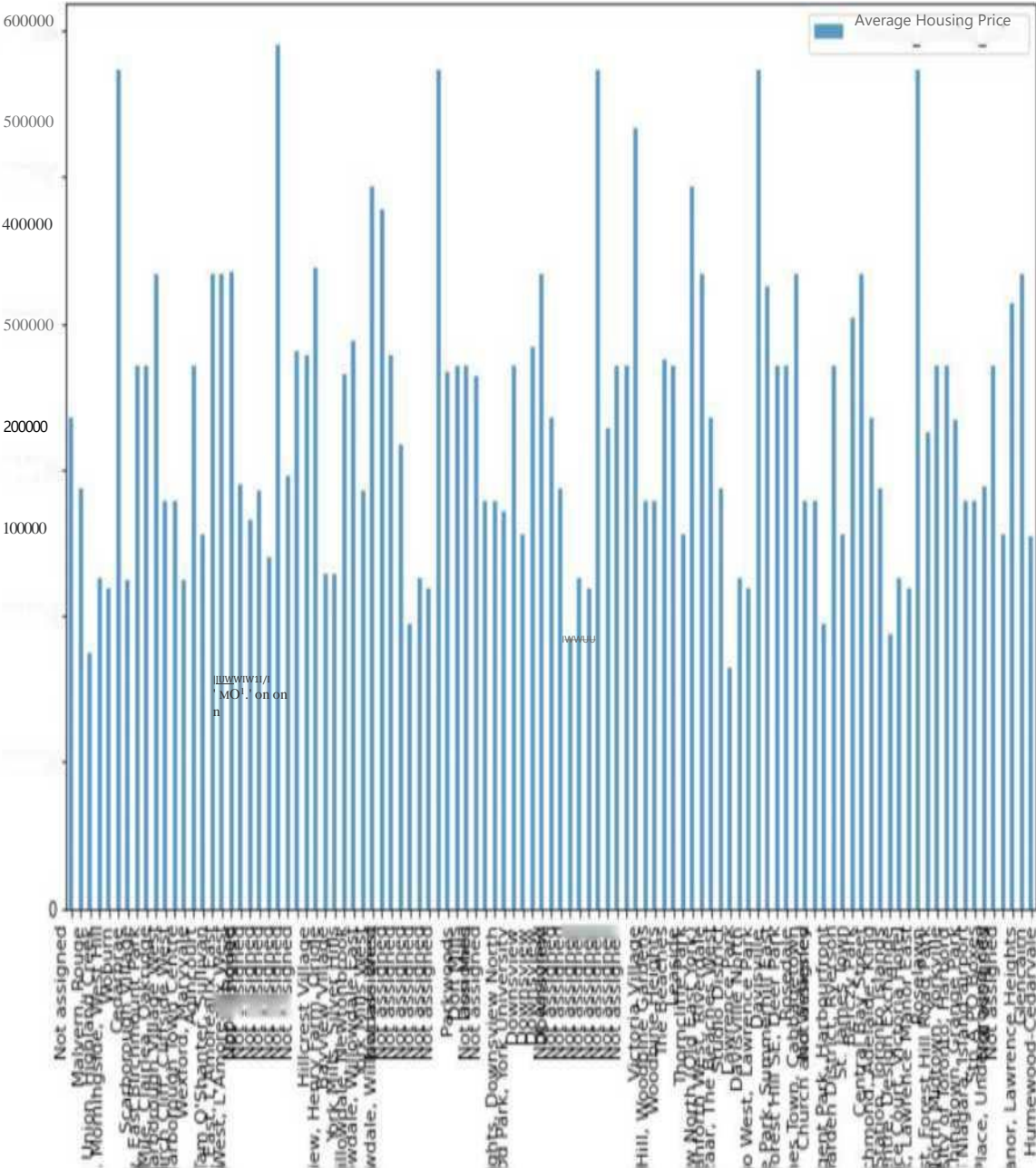
School Ratings by Clusters in Scarborough
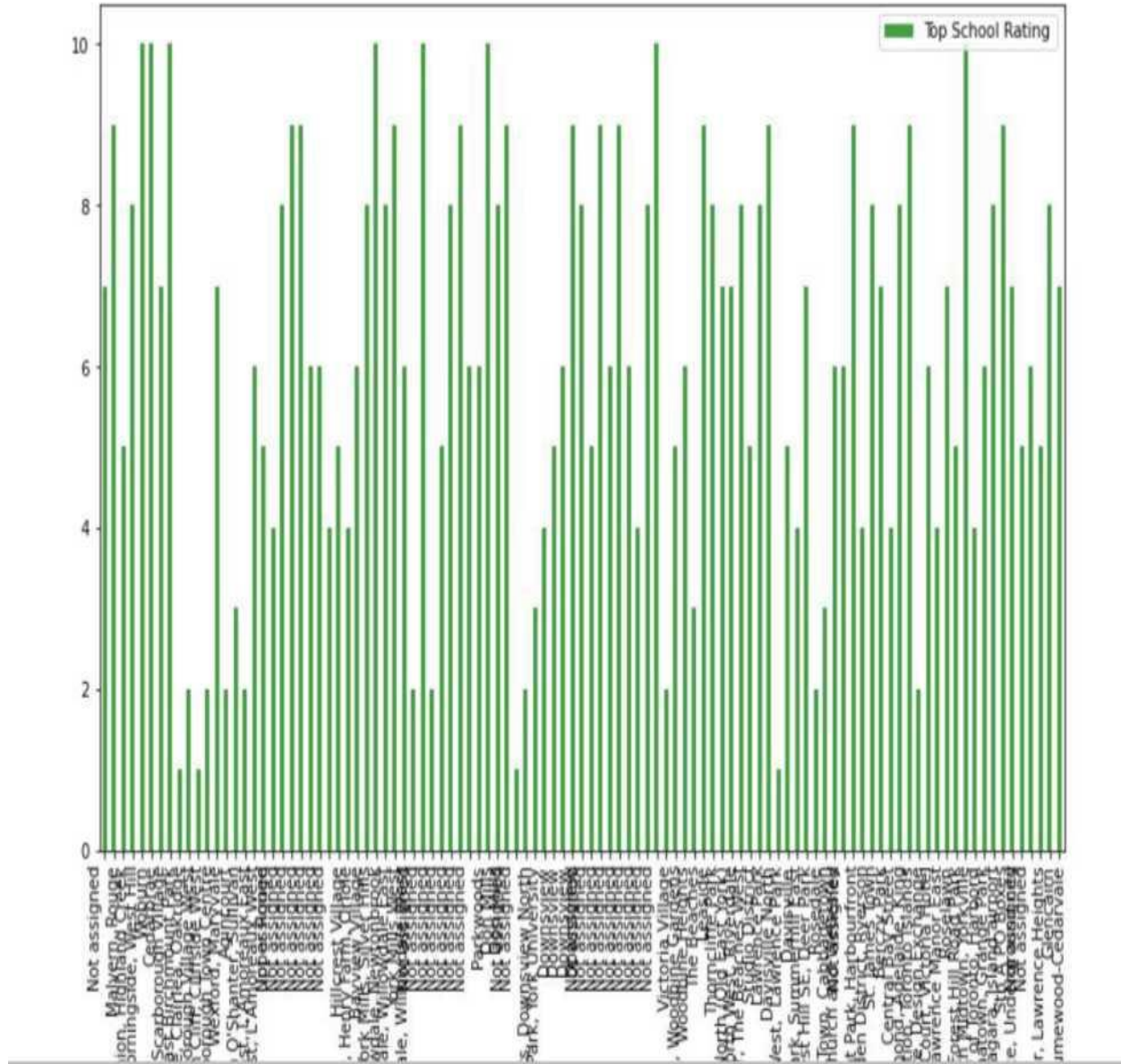
The Location:

5. Discussion Section

Problem Which Tried to Solve:

The major 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 mindedpeople. Connectivity to the airport, bus stand, city center, markets and other daily needs things nearby.

- 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

• 6. Conclusion Section

- In this Capstone project, using k-means cluster algorithm I separated the neighborhood into 10(Ten) different clusters and for 103 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.

- I feel rewarded with the efforts and believe this course with all the topics covered is well worthy of appreciation.
  This project has shown me a practical application to resolve a real situation that has impacting personal and financial impact using Data Science tools.
  The mapping with Folium is a very powerful technique to consolidate information and make the analysis and decision better with confidence.