



THE BATTLE OF NEIGHBORHOODS

LOOK FOR A NICE RESIDENT IN
DOWNTOWN TORONTO

INTRODUCTION

Business Problem:

- As we know a lot of people have been migrating to Canada for better study and career opportunities
- They are not local residents of Canada and are unaware of various locations, places in Canada.
- Specifically when it comes to looking for a nice and better place to live.

Target Audience:

- Immigrants to Canada who are not familiar with Canada and are looking for a better place to reside.
- This project is aimed at solving this by providing a better way to explore and find various neighborhoods based on various factors like access to nearby supermarket, grocery stores, malls, average housing price, school ratings.
- It will provide all results and information via an interactive map and bar charts making it easier for end user.



DATA SECTION

For this project we will be focusing only on Downtown Toronto area which is a popular destination for new immigrants moving to Canada.

Data Source: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

Libraries Used:

- Pandas: Standard library for all required dataframes operations.
- Scikit Learn: to import K-Means clustering.
- Matplotlib: to plot bar charts.
- Folium: To visualize cluster distribution of neighborhoods using interactive leaflet map.
- Beautiful Soup: To fetch and handle http requests operations.
- JSON: For JSON files operations.
- XML: To fetch data from XML and store it in dataframe.
- Geocoder: To fetch location coordinates.



This data contains Postal codes, Borough and Neighborhood

Postalcode	Borough	Neighborhood
M1B	Scarborough	Malvern, Rouge
M1C	Scarborough	Rouge Hill, Port Union, Highland Creek
M1E	Scarborough	Guildwood, Morningside, West Hill
M1G	Scarborough	Woburn
M1H	Scarborough	Cedarbrae

METHODOLOGY SECTION

Following methods and approaches are used to get required results

Foursquare API:

- Used it to fetch information related to nearby venues located inside each and every neighborhood.
- Explore endpoint has been used to get nearby venues names and coordinates.

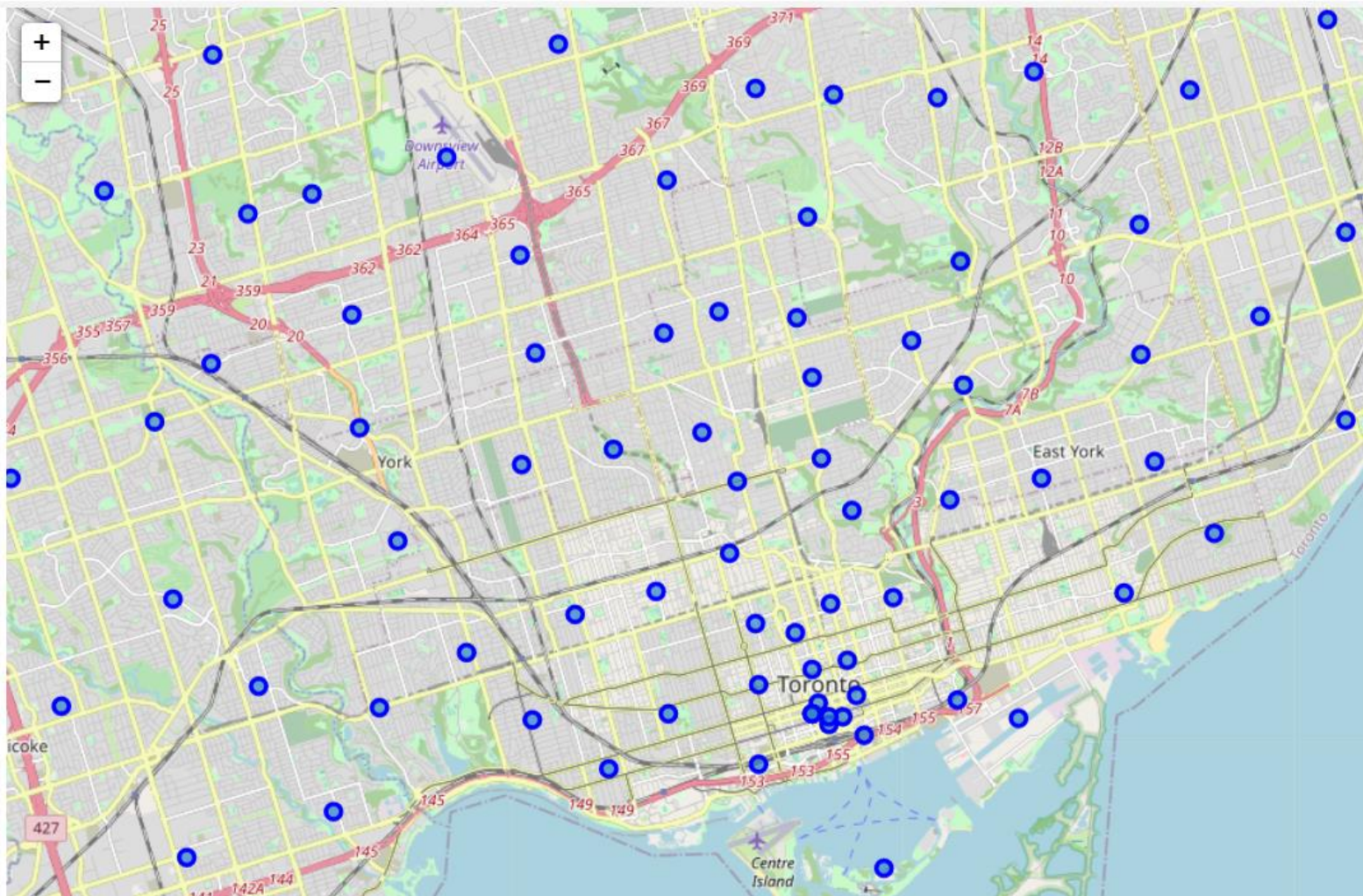
```
url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
    CLIENT_ID,
    CLIENT_SECRET,
    VERSION,
    lat,
    lng,
    radius,
    LIMIT)
```

- Used one hot encoding technique and clustering to segment neighborhoods and group them into common clusters
- K-Means clustering to split nearby neighborhood into 3 clusters and find most common venues near neighborhood

	Postalcode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
0	M1B	Scarborough	Malvern, Rouge	43.808626	-79.189913	2	Park	Trail	Women's Store	Eastern European Restaurant
1	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek	43.785779	-79.157368	0	Bar	History Museum	Park	Fish & Chips Shop

© 2010 Blackwell Publishing Ltd *Journal of Internal Medicine* 267: 221–229

Neighborhoods Segmented into clusters over Downtown Toronto area

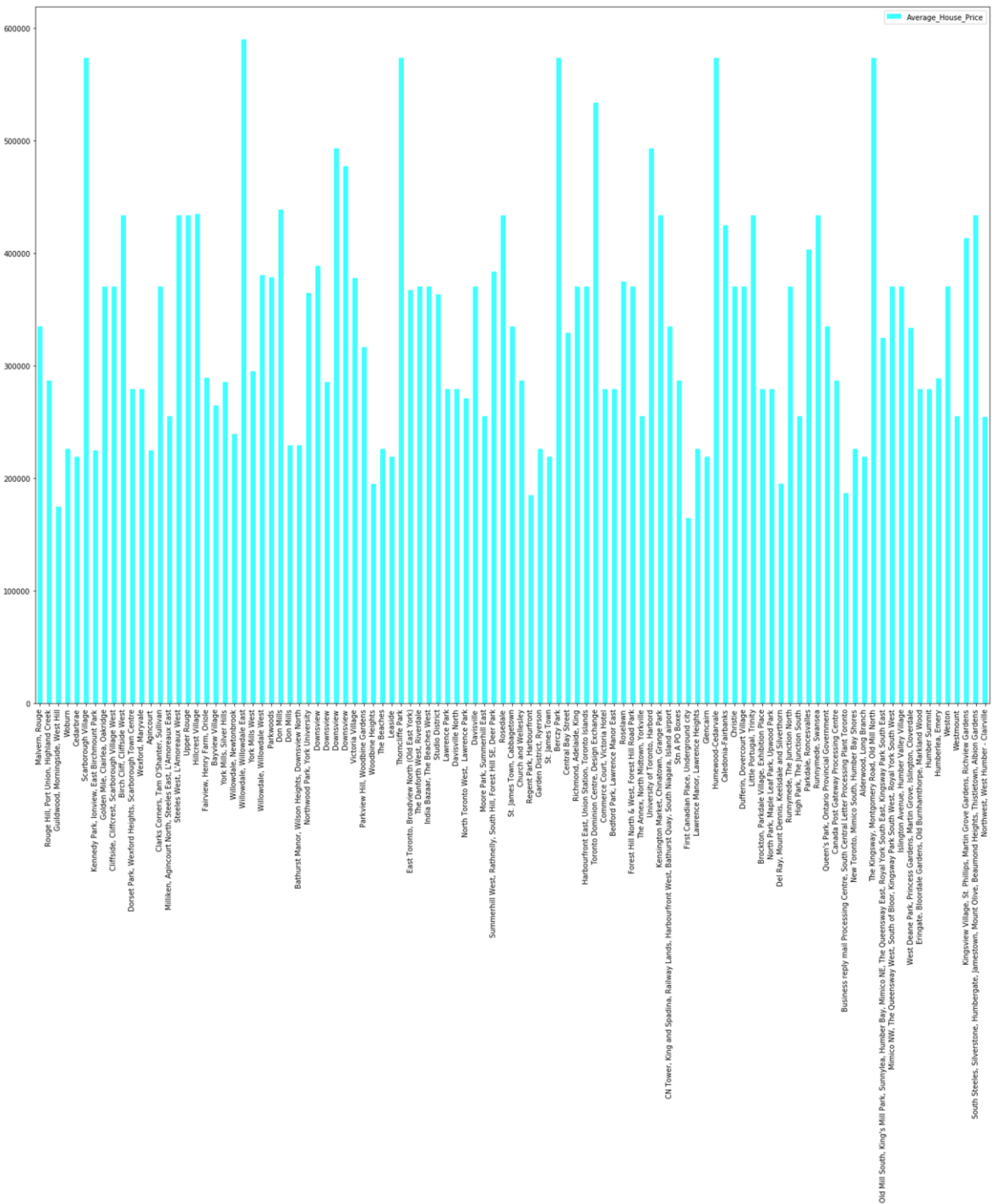


School Ratings by Clusters in Downtown Toronto Area



RESULTS SECTION

Average Housing Price by Clusters in Downtown Toronto Area



DISCUSSION AND CONCLUSION

Discussion Section - Observations:

- House prices should be the major contributing factor in decision making as opposed to School ratings.
- Most often schools with high rates are located in most expensive places which could also be one of deciding factors in House prices.
- So, neighborhoods with schools will probably have more demand and more price.

Conclusion Section:

- It would be best to first shortlisted neighborhoods based on specific price ranges say, 270000 to 320000.
- Then one by one eliminate those neighborhoods with School ratings below 5.
- Then ranking remaining neighborhoods first by lowest house price and then by highest School rating and based on it pick your desired neighborhood.





THANK YOU