CAPSTONE PROJECT

Kunal Raval

Introduction

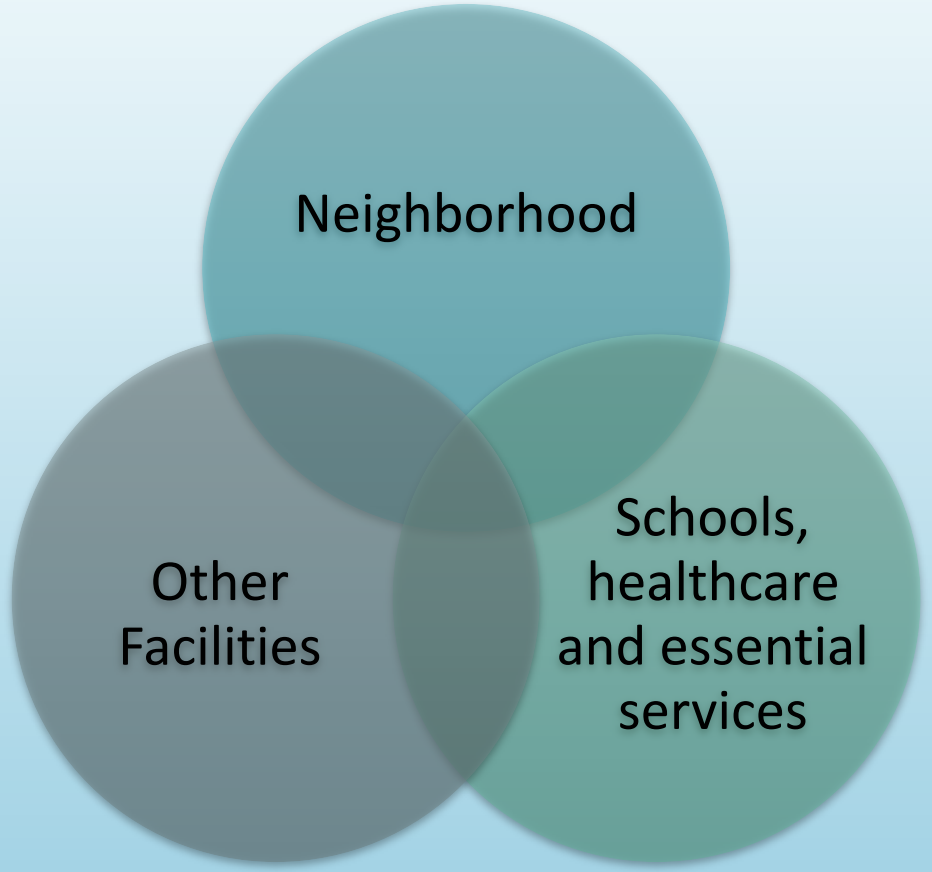
- Main goal of this project is to help the citizens to help and provide accurate and precise detail about the facilities around their neighborhood.
- Thousands of people and families are migrating to Canada and Greater Toronto Area is one of the most popular place in Ontario. So, Scarborough is one of the city which welcomes these immigrants in significant amount.
- It will help people making smart and efficient decision on selecting great neighborhood out of numbers of other neighborhoods in Scarborough, Toronto.

Business Problem:

- The challenge is to find a suitable apartment for rent in Scarborough, Toronto.
- That complies with the demands on location, price and venues.

Major factors:

- These are the main three factors based on which a new - comer can decide in which area they should buy or rent the property.



- **Transform the data into a pandas data frame:** The next task is essentially transforming this data of nested Python dictionaries into a pandas data frame. So let's start by creating an empty data frame.
- **Use Geopy library to get the latitude and longitude values of Scarborough City.** In order to define an instance of the geocoder, we need to define a user agent.
- **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.
- **Beautiful Soup and Requests:** To scrap and library to handle http requests.
- **Matplotlib:** Python Plotting Module.

Approach to solve this problem:

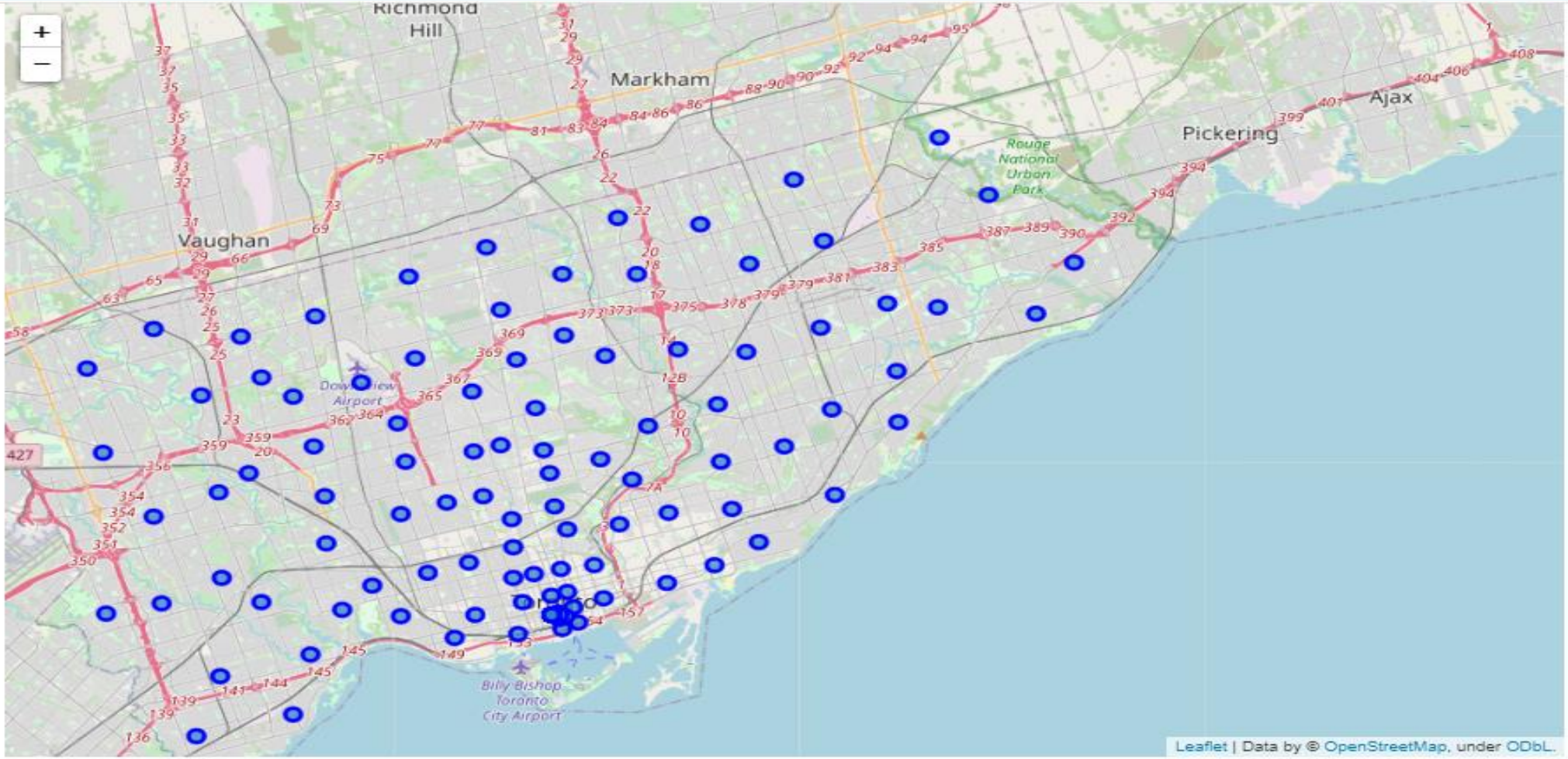
Steps to find this results:

- Install important libraries
- Fetch data from the wiki file
- We need to clean and filter out the data
- once we have filtered data then visualize and diplay it on the map
- get nearby Venues and locations & get categories of those locations
- resolve data duplications.
- Do "one hot encoding"
- Most Common venues near neighborhood
- Perform K-Means Clustering
- Visulize Data and show the hotspots on the map.

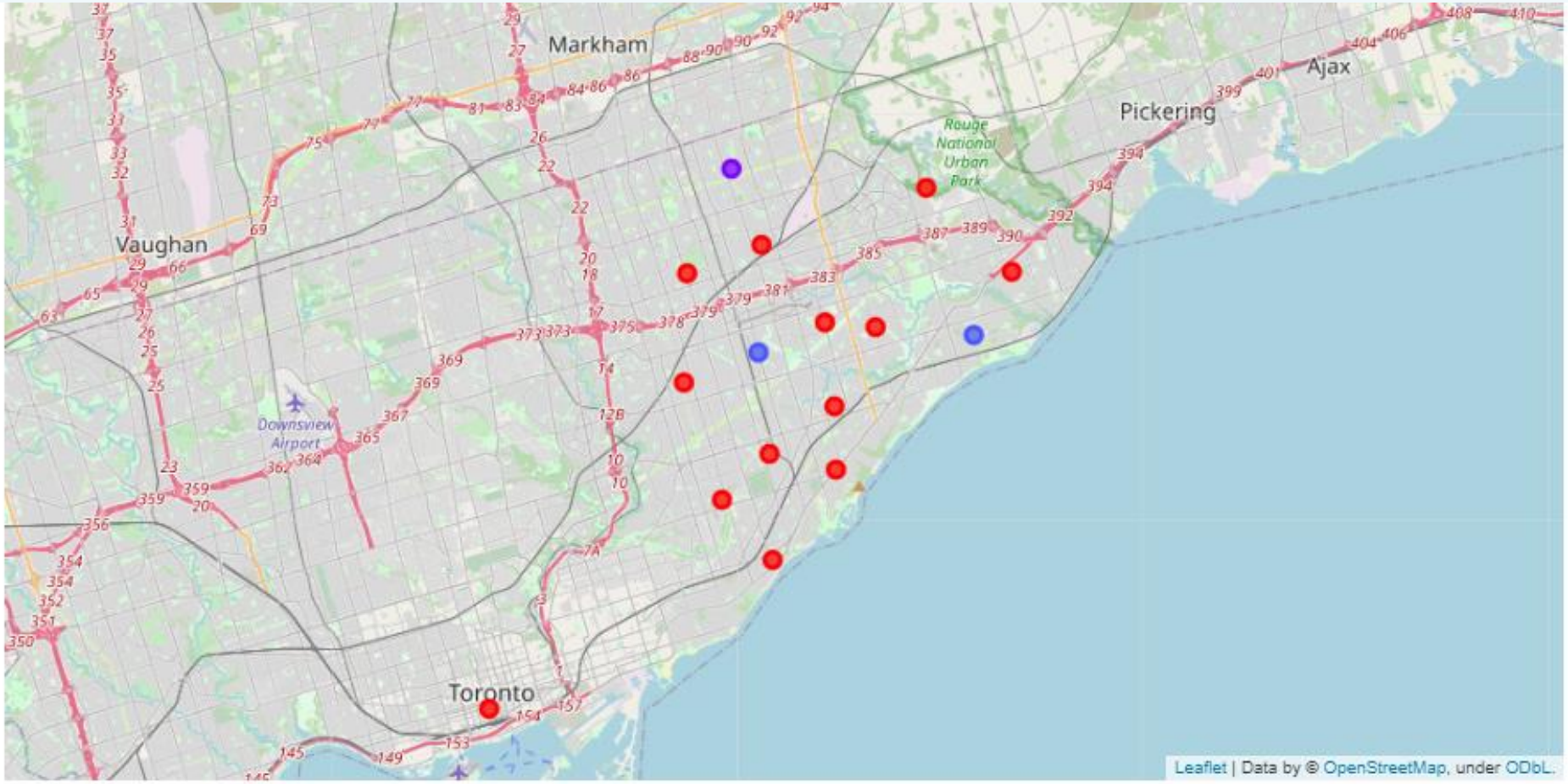
	Postalcode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue
0	M1A\n	Not assigned\n	Not assigned\n	43.64869	-79.38544	0	Coffee Shop	Hotel	Café	Restaurant	Gym	Theater	Movie Theater
1	M1B\n	Scarborough\n	Malvern, Rouge	43.81153	-79.19552	0	Zoo Exhibit	Construction & Landscaping	Financial or Legal Service	Fast Food Restaurant	Falafel Restaurant	Donut Shop	Dumpling Restaurant
2	M1C\n	Scarborough\n	Rouge Hill, Port Union, Highland Creek	43.78564	-79.15871	0	Moving Target	Home Service	Fish & Chips Shop	Bar	Farm	Dumpling Restaurant	Eastern European Restaurant
3	M1E\n	Scarborough\n	Guildwood, Morningside, West Hill	43.76575	-79.17520	2	Park	Athletics & Sports	Gym / Fitness Center	Yoga Studio	Event Space	Donut Shop	Dumpling Restaurant
4	M1G\n	Scarborough\n	Woburn	43.76820	-79.21761	0	Park	Chinese Restaurant	Fast Food Restaurant	Coffee Shop	Yoga Studio	Event Space	Donut Shop

Results section:

Out[18]:



Results section:



This is how we can cleanse the data, create data frames which contains all the details raleated to the neighbourhood and based on that we used FourSquare API and K-clustering approach which finally provides the visulized and detailed view on map.

Results section:

- **Foursquare is a location data provider with information about all manner of venues and events within a particular area. 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.**

Key Benefits of using Four Square API

Conclusion:

- **As a result, people are migrating to big cities to start a business or work or to have a better lifestyle. For this reason, people can achieve better outcomes through their access to the platforms where such information is provided.**
- **As an immigrant this details are so useful and they can have better idea where they should plan to find a property (in this case, Scarborough, Toronto).We can implement the same approach to any major cities with collaboration of FourSquare API which provides millions and millions of records and this API is growing in size so rapidly and contain so much information which can be used to get the more precise result.**