

## Capstone Project - The Battle of Neighborhoods

### *Introduction*

In this study we propose to compare localities in New York City. The location data are proposed to be collected from various open source data providers like Foursquare. The location dataset are expected to provide us with various information like the neighborhood, working hours, geographical location etc. By comparing the different locations we can find out a suitable location.

This study is mainly done for finding a suitable location for a new restaurant. The location data is helpful to find out the suitable location for the restaurant. The food menu can be suitably designed by looking into the location data.

### *Downloading and Prepping Data*

The location data can be downloaded from different open source locations. Some of the location data can even be accessed from Wikipedia site. The other valuable sites from which we can access location data are FourSquare API, Google Location and YELP. We propose to access the data from Wikipedia from the location from the url: [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M).

### *Methodology*

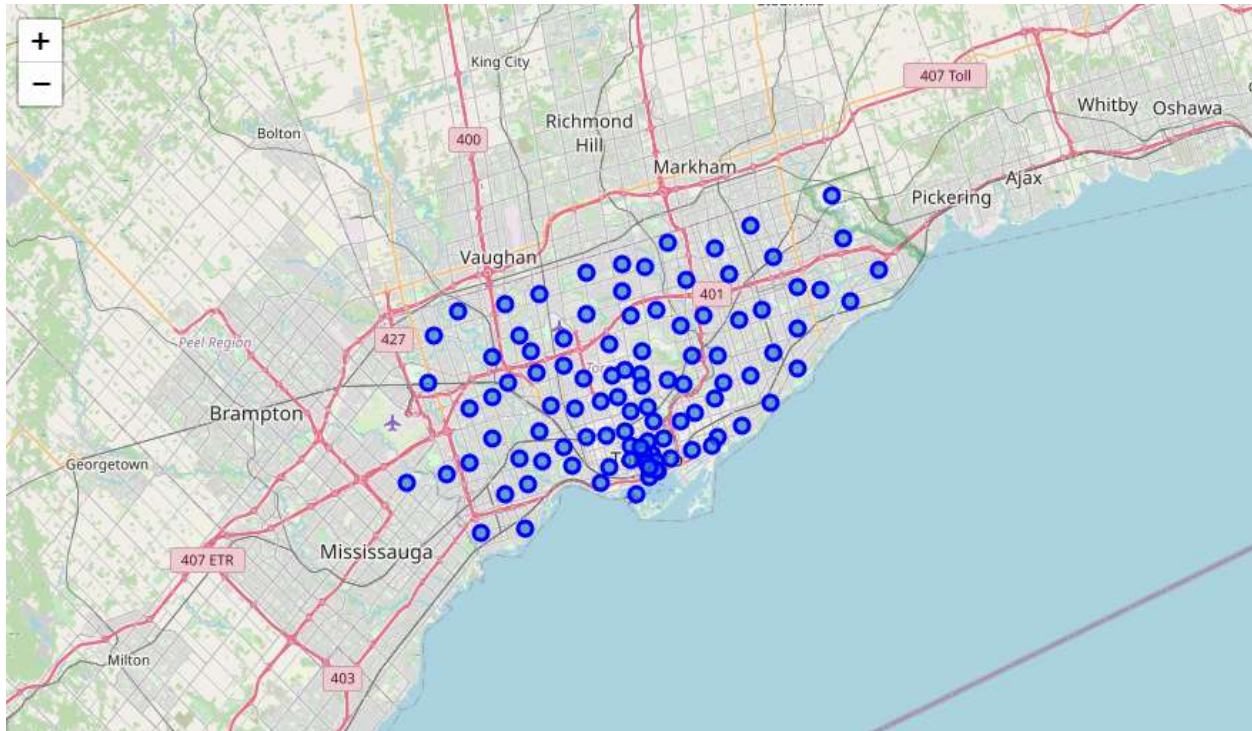
In this study we analyze the location data collected from various sources to find a suitable location for opening a restaurant. We use python as programming tool for analyzing the data. In Python we will use Folium particularly for creating visualization for map data. Moreover we have used K-Means clustering algorithm for finding out the similarity among the locations.

### *Results:*

After reading the data and applying the methods, we could find the different localities and their information. Following is a part of data extracted in data frame format.

	Borough	Neighbourhood	Latitude	Longitude
0	Scarborough	Rouge, Malvern	43.806686	-79.194353
1	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497
2	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711
3	Scarborough	Woburn	43.770992	-79.216917
4	Scarborough	Cedarbrae	43.773136	-79.239476

And following is a part of the map drawn from the data.



### *Discussion*

The study proposes some groups or clusters of the localities depending on their activities. These locations may meet the requirements of the business for opening the restaurants.

### *Conclusion*

This is a very preliminary study done with limited open source data. This study will be fruitful for achieving the business objectives. Still further analysis of the data with detailed methodology and tools may provide us with better results.