**Capstone Project- Clustering the Neighborhoods of New York and Toronto**

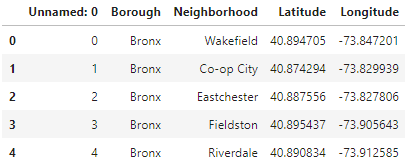
**1. Introduction**

New York City and Toronto are the financial capitals of their respective countries. In this project, I am going to explore their neighborhoods and find out the top 10 common venues for each neighborhood. I will make the comparison between both neighborhoods by clustering analysis.

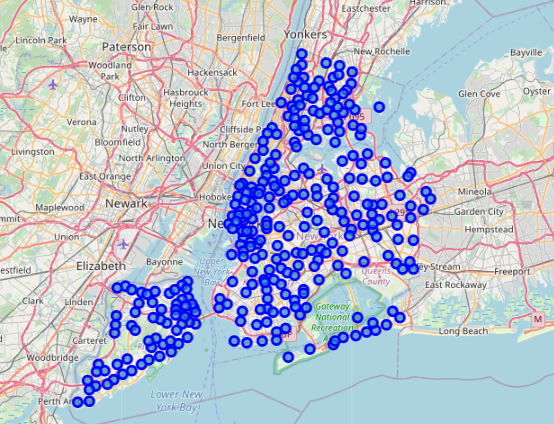
### 2. Data used

For this project, the following data is required:

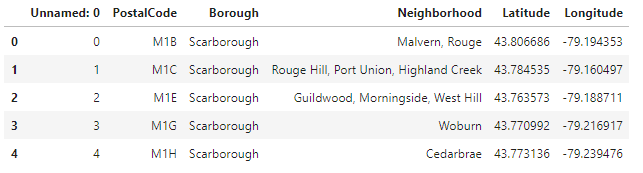
* A list of neighborhoods of New York City



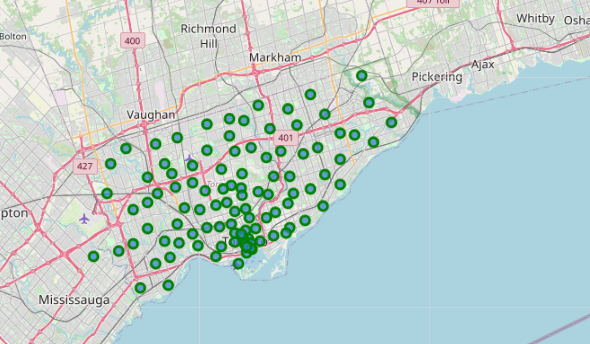
(306 Rows, 5 Columns)



* A list of neighborhoods of Toronto



(103 Rows, 6 Columns)



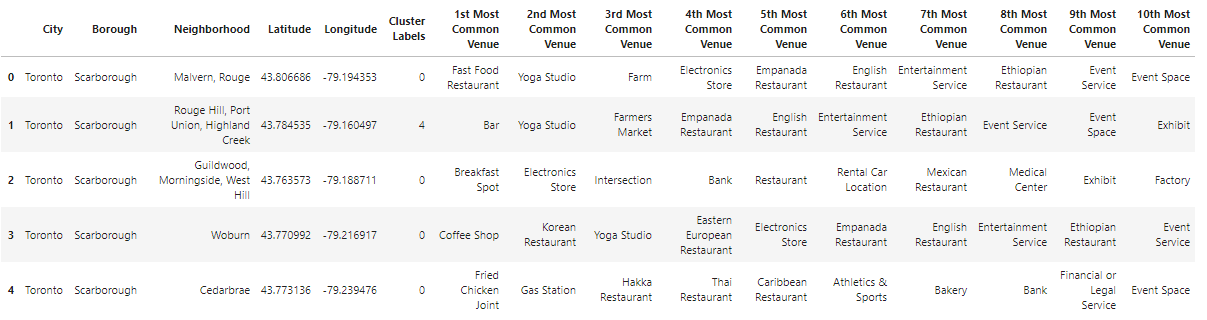
To acquire the top 10 venues and categories of each neighborhood, New York table and Toronto are merged, and Foursquare API is used.



(396 Rows, 12 Columns)

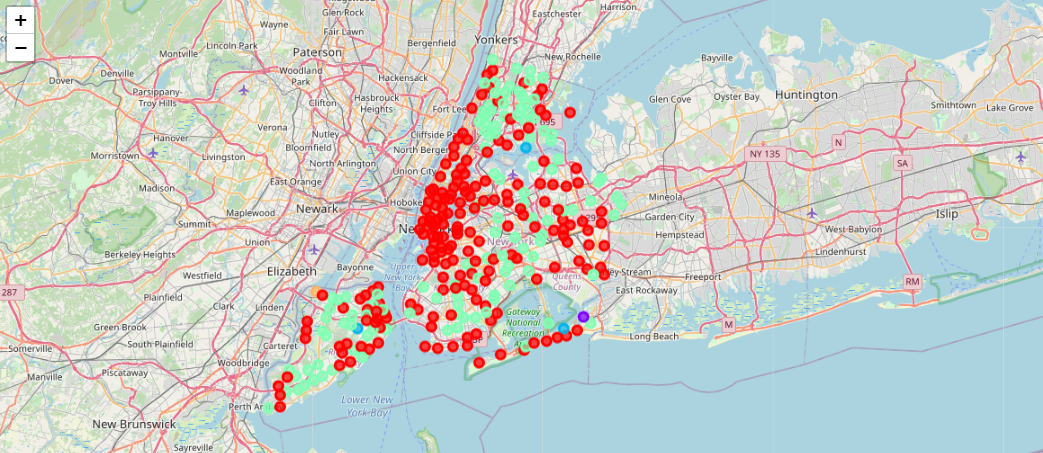
### 3. Methodology

After identifying top 10 most common venue for each neighborhood of New York and Toronto, “k-mean” clustering is applying for analysis based on top 10 most common venue. K-Cluster is assumed as 5. It means that the neighborhoods are categorized into 5 clusters: “0”, “1”, “2”, “3” and “4”.

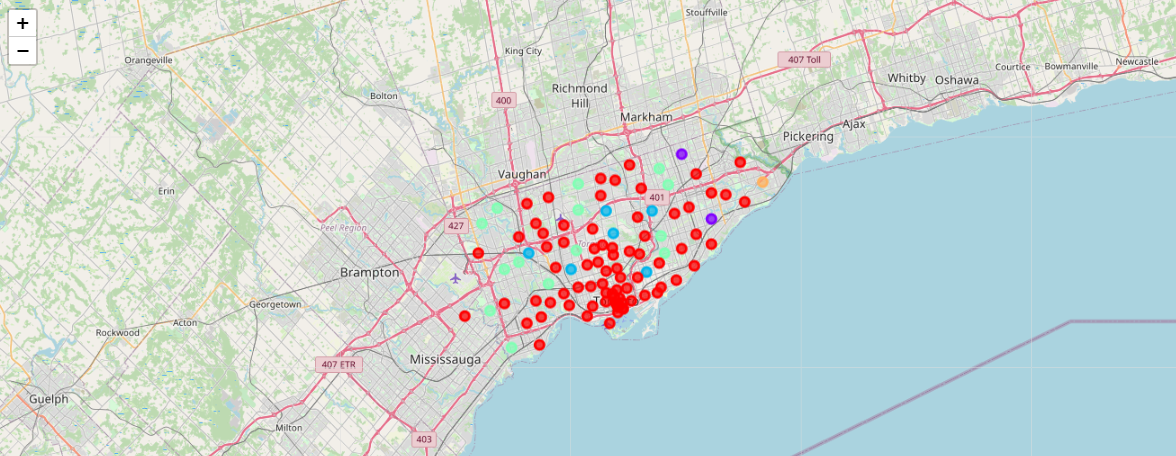


Below graphs are showing the clustered neighborhood of New York and Toronto.

New York:

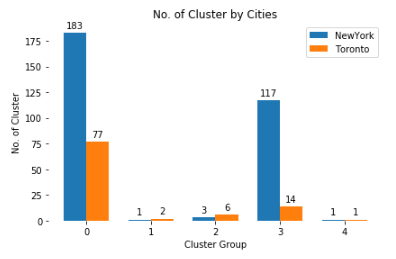


Toronto:



### 4. Analysis and Result

The below graph shows the number of neighborhoods of New York and Toronto for each cluster.



The following tables show the coverage of the most common venues by each cluster:

|  |  |
| --- | --- |
| Cluster 0 | Cluster 1 |
|  |  |

|  |  |
| --- | --- |
| Cluster 2 | Cluster 3 |
|  |  |

Cluster 4



Based on the result above, regardless of New York and Toronto, most common venues of neighborhoods are categorized into Cluster 0 and Cluster 3.

* In cluster 0, New York has 183 venues and Toronto has 77 venues; In cluster 3, New York has 117 venues and Toronto has 14 venues. In other clusters, New York and Toronto only has less than 10 venues.
* In Cluster 0, Coffee Shop are the most common venues which cover 12.25%; In Cluster 3, Pizza Place is the most common venues which cover 15.94%.

### 5. Conclusion

### In this project, I have categorized neighborhoods of New York City and Toronto into 5 clusters. The results show the most common venues for each cluster. Such analysis can help people to identify what venues the people in the financial capital: New York City and Toronto are most likely to go.