

|  |
| --- |
| The Battle of Neighborhoods 2019 |
|  |
| April 20  Manik’s ASSIGNMENT  Authored by: Manik Garg |

# Capstone Project – The Battle of Neighborhoods

|  |
| --- |
| **Introduction**Background There are many cities with many differences. This report will find the similarity and dissimilarity of Toronto and New York. The audience of this report are tourists, investors and who interested in different cultural diversity study.  All these cities are the financial capital of their respected countries. Though these cities are occupied with many things still they have scope for the new local business-like coffee shop, gym, health club etc. Everyone want to sift to a place where all basic amenities are available. We will be predicted that Toronto and New York have close similarity and also dissimilarity between these cities with the help of coffee shops in neighborhood, but see what our exploring, segmenting and clustering of their neighborhoods have to say. Data Description  * New York data is from https://cocl.us/new\_york\_dataset * Toronto data is from the week 3 analysis. City based data is from <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M> * The Find Latitude and Longitude site (<https://www.findlatitudeandlongitude.com/>) has used to get the boundary of these cities and make sure we are not out of metropolitan area and used http://cocl.us/Geospatial\_data * Foursquare API service has used to get the required data for the report. * In total, 288 rows in the raw dataset. Cleaned data has 103 rows. |

### B. Methodology

As a database, I used GitHub repository in my study. My master data which has the main components PostalCode *,Borough,* Neighborhood



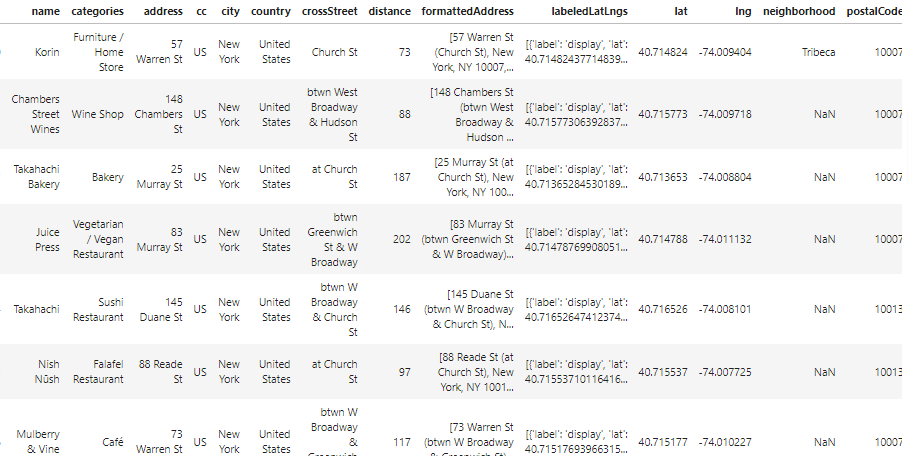
I used python folium library to visualize geographic details of New York and Toronto.We have built a dataframe of the postal code of each neighborhood along with the borough name and neighborhood name, in order to utilize the Foursquare location data, we need to get the latitude and the longitude coordinates of each neighborhood.

We will use the Geocoder Python package instead: https://geocoder.readthedocs.io/index.html.

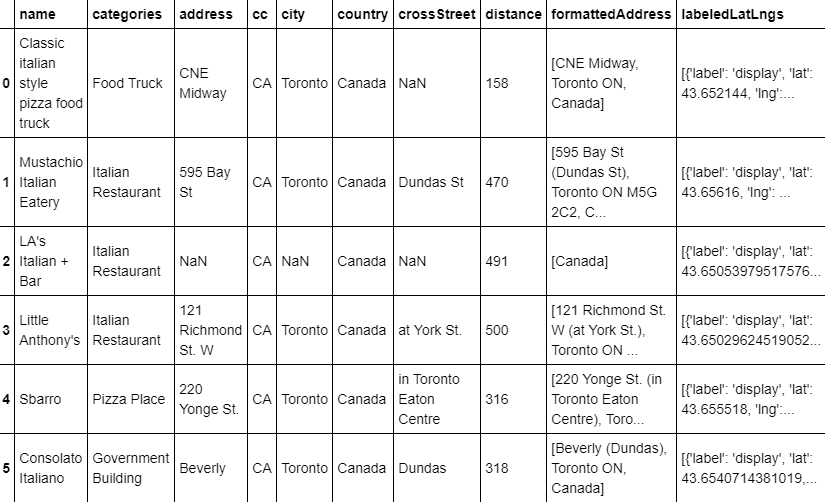
The result doesn’t mean that inquiry run all the possible results in boroughs. Actually, it depends on given Latitude and Longitude informations and here is we just run single Latitude and Longitude pair for each borough. We can increase the possibilities with Neighbourhood information with more Latitude and Longitude information.

Below are screenshots of some important neighbourhood.

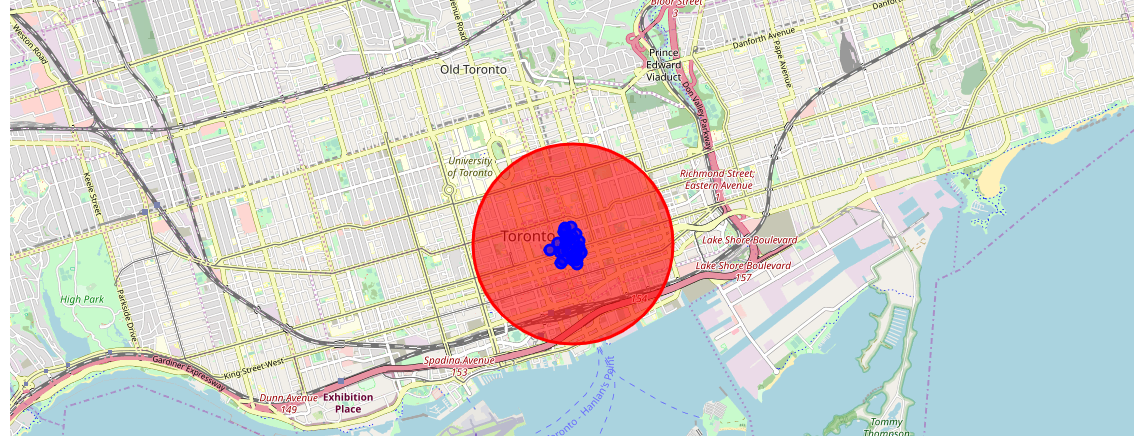
New York Data for Coffee Shop Neighbourhood.



Toronto Data for Coffee Shop Neighbourhood.



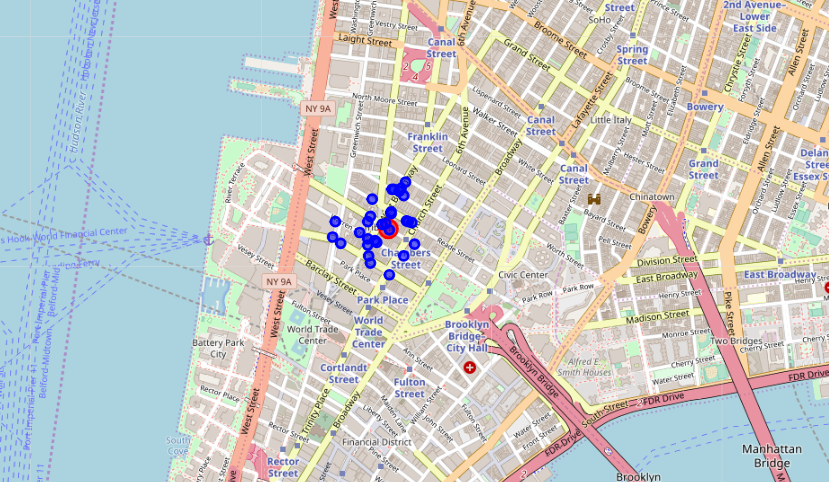
Toronto Government building



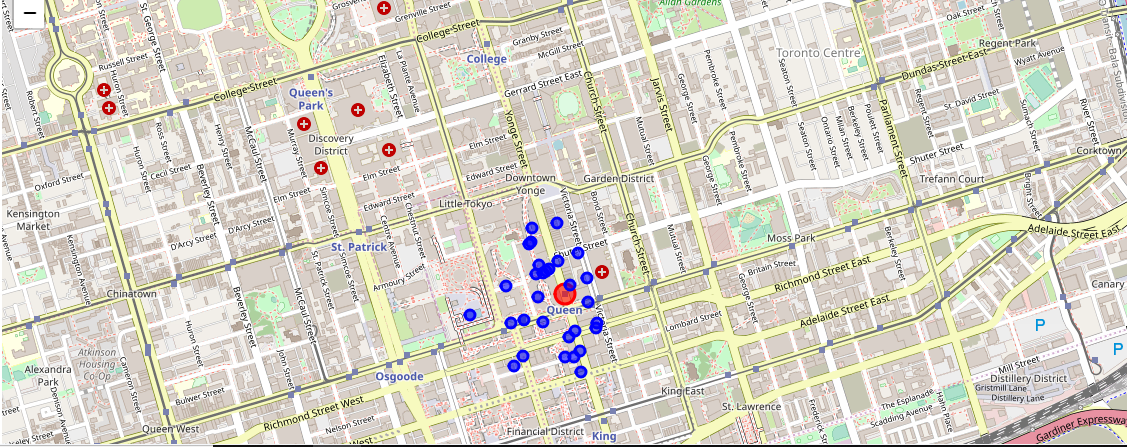
### Results

In final section, I created choropleth map which also has the below information for each borough:

New York Map for Coffee Shop



Toronto Map



### D. Discussion

As I mentioned before, New York and Toronto are big cities with a high population density in a narrow area. The total number of measurements and population densities of the 12 unique districts of Toronto in total can vary. As there is such a complexity, very different approaches can be tried in clustering and classification studies. Moreover, it is obvious that not every classification method can yield the same high-quality results for this metropole.

I used the Means algorithm as part of this clustering study. When I tested the Elbow method, I set the optimum k value to 3. However, only 12 district coordinates were used. For more detailed and accurate guidance, the data set can be expanded, and the details of the neighbourhood or street can also be drilled.

Below is the classification snapshot of US city.

I also performed data analysis through this information by adding the coordinates of districts and home sales price averages as static data on GitHub. In future studies, these data can also be accessed dynamically from specific platforms or packages.





I ended the study by visualizing the data and clustering information on the New York and Toronto map. We could easily easy the dis-similarity between New York and Toronto coffee shop density.

New York has much more coffee shop than Toronto.

In future studies, study for the neighbourhoods for any type of food joint can be carried out.

### F. Conclusion

As a result, people are turning to big cities to start a business or work. For this reason, people can achieve better outcomes through their access to the platforms where such information is provided.

Anyone who wish to open a food joint or mainly coffee shop near building like government building can use the analysis.