**Capstone Project : Battle of the Neighborhoods Report**

**Sevda Tekin**

**1. Introduction**

One of the best things about Toronto's multiculturalism is all the great food from diverse cultures. Among one of the best cuisines in Toronto is Thai food. This project aims to find the best locations to open a Thai Restaurant in Toronto, Canada. The people who want to open a Thai Restaurant in Toronto or the investors in food industry would be interested in this project.

**2. Data**

In order to find best locations to open a Thai Restaurant in Toronto, data on Toronto neighborhoods, boroughs,postcodes, latitude, longitude, restaurants, and restaurant ratings and tips are required.

* Data containing the neighborhoods, boroughs and postcodes of Canada will be obtained from the data source:

<https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>

* Data containing the geographical coordinates of each postal code has a link below :

<http://cocl.us/Geospatial_data>

* All data related to locations and quality of Italian restaurants will be obtained via the FourSquare API utilized via the Request library in Python.

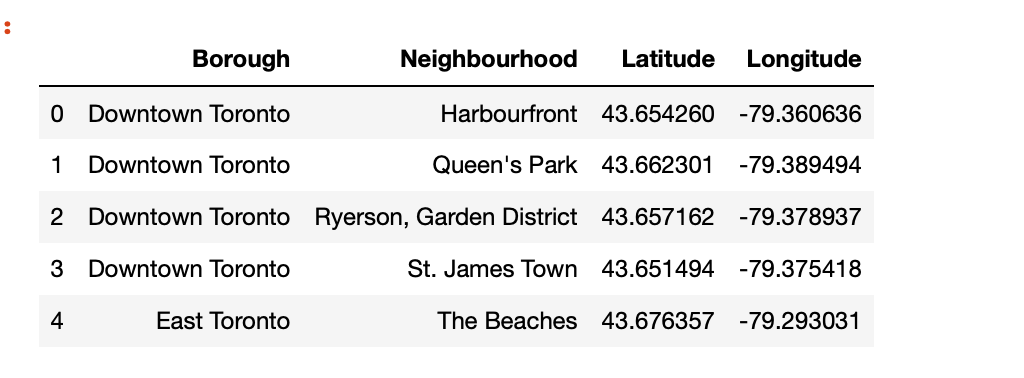
**3. Methodology**

Firstly we need to extract the datas from wikipedia page "List of Postal Codes of Canada" to a csv file and then we will build a dataframe

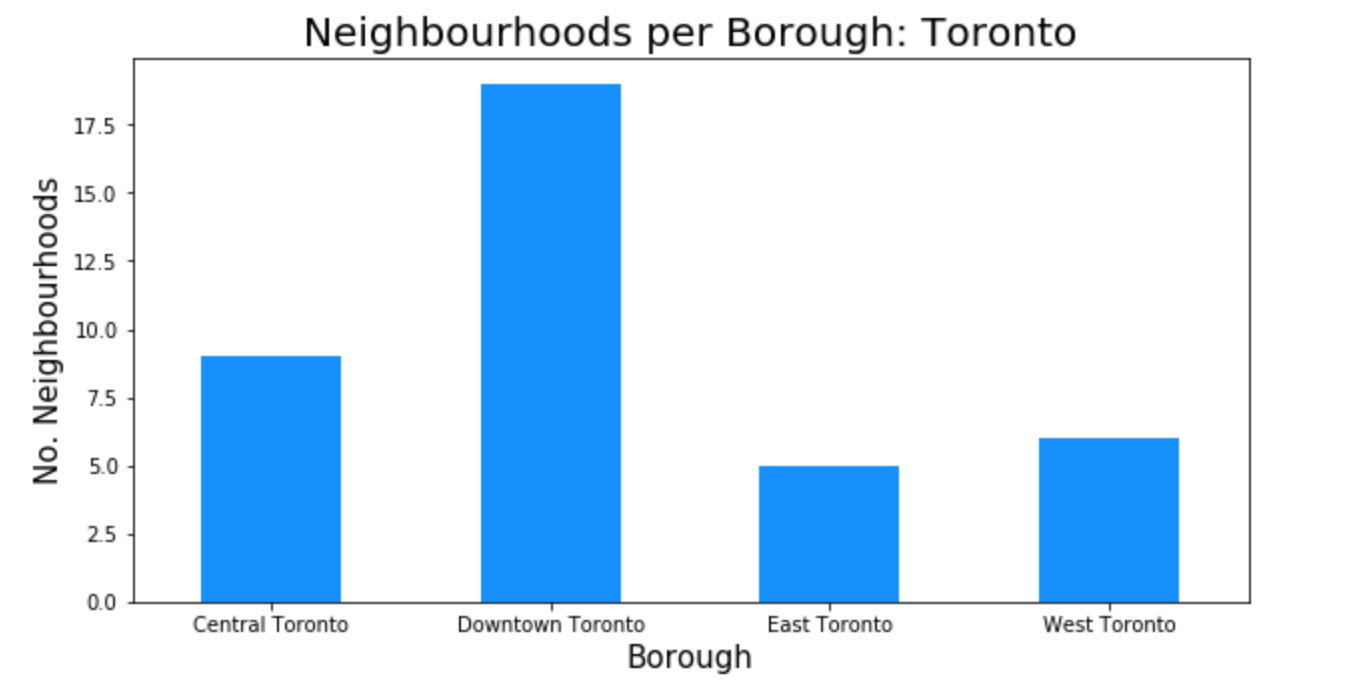
After Dataframe is generated, we need to clean data. These are the steps :

* First we remove the unused columns if any. We only need Postcode, Borough and Neigbourhood. Other columns can be removed.
* We only process the cells that have an assigned borough. Thats why, we ignore cells with a borough that is Not assigned.
* Some of the neighborhoods exist in one postal code area, these neighborhoods will be combined into one row with the neighborhoods separated with a comma.

Then we built and clean second dataframe with latitude and longtidue and merge 2 dataframe. After merging we filter the data to only use boroughs in Toronto :

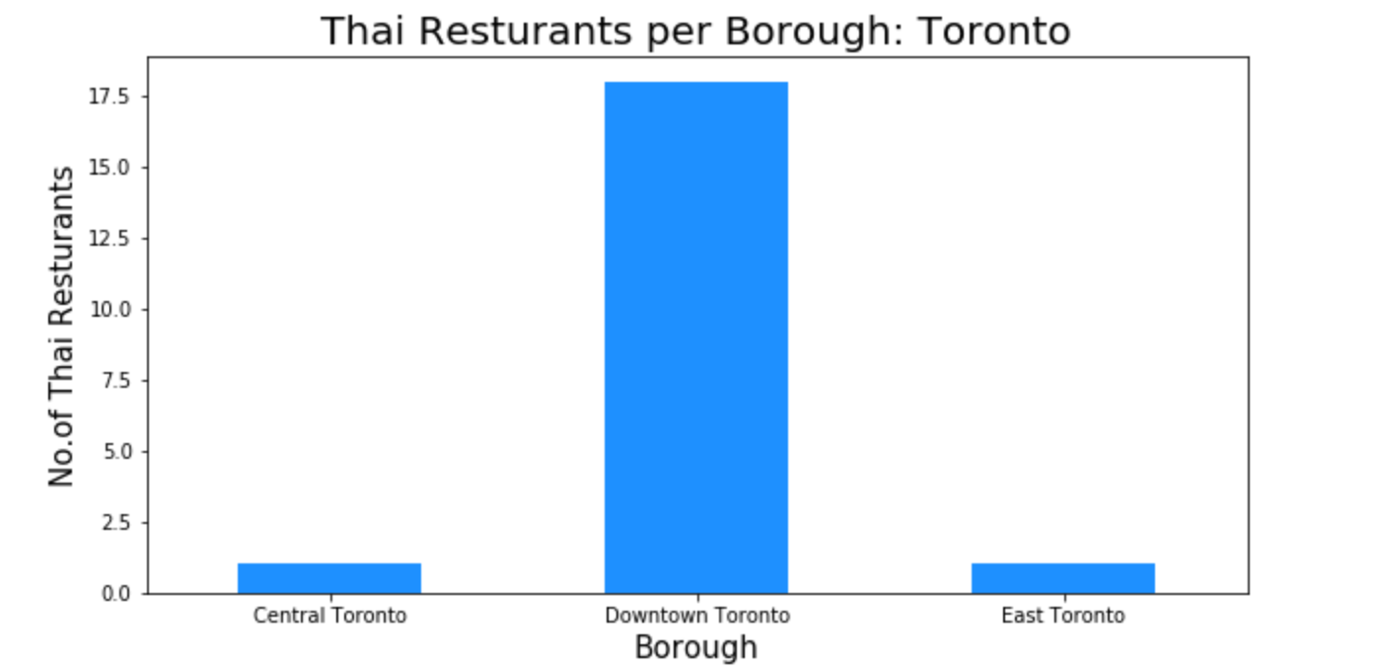


Ok, now our Toronto DataFrame is ready, we can start explore the neighborhoods in Toronto :

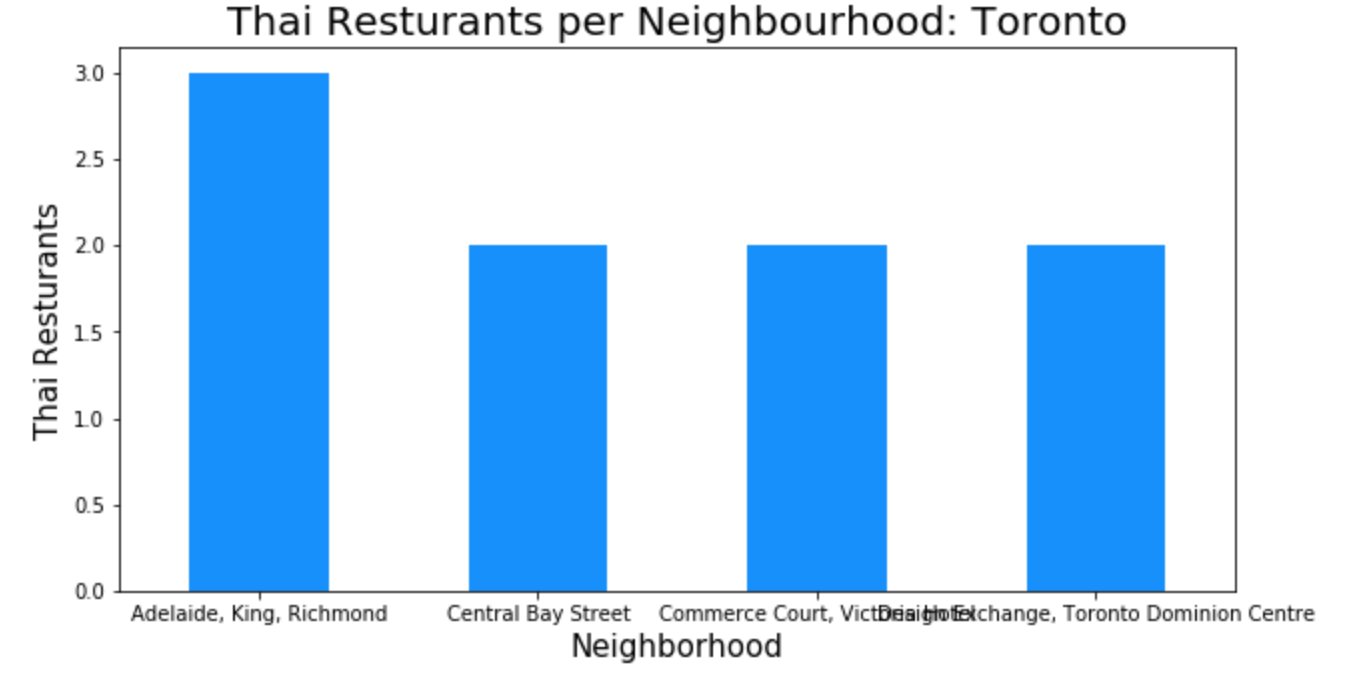


Lets analyze further and see how many Thai restaurants there are in each neighborhood and borough and graph the results.

To do this, we need to call Foursquare API to get venues and venue deteails. As a result we can see below graph:



As we continue our analysis, we see above that Downtonwn Toronto had the most number of neighborhoods and it does have the highest number of Thai restaurants. Additionally, we see there is no Thai Restaurant in Borough West Toronto. Now let's visualize Thai Restaurants for each Neighbourhood :



After some coding in notebook, we can see that only 12 neighbourhood has Thai Restaurant out of 38 neighbourhood. Adelaide, King, Richmond Neighbourhood has the maximumnumber of Thai Restaurants in Toronto. Now we are going to cluster all the neighborhoods by using K-means Clustering Model.

First we create a function to get the top 100 venues for each neighbourhood in Toronto within a radius of 500 metres. Then we will cluster them regarding the venues' similarity.

Now, lets see top 10 venues frequency for each neighborhood



Now, we will run \*k\*-means to cluster the neighborhood into 5 clusters.

**Results :**

Let's create a new dataframe that includes the cluster as well as the top 10 venues for each neighborhood.



**Discussion**

As we can see above 34 neighborhoods are in 'Cluster 1' out of 38. So maybe we can check the Thai restaurant's rating and then we can decide where we open the Thai Restaurant.

The neighborhood of Adelaide, King, Richmond has the highest number of Italian restaurants in all of Toronto and is actually located in the borough of Downtown Toronto.

**Conclusion**

So, as a result most of the neighbourhood are in cluster 1 and most of the borough in cluster 1 is Downtown Toronto. So, I would suggest to open a Thai restaurant in Downtown Toronto.