Recommending an optimal place for opening a restaurant to owners

Project Report

# Introduction:

A lot of fast food businessmen who own a chain of restaurants or a franchise of a famous chain or even a food truck will need to ask the question that where should they expand their business to or where should they open up a new franchise or in case of a food truck, around where should they be so that they get maximum number of customers to increase their sales and generate revenue.

Using data science on geospatial data provided to us by foursquare API, we can analyze the neighborhood in Toronto and give those owners the answers to their questions. We can tell them the appropriate place where they can hit the maximum number of customers and generate profit.

# Data:

We can apply web scraping for getting different neighborhoods in Toronto, CA. After we have different neighborhoods, we can use Foursquare API to get all the venues like parks, cafes, offices in those neighborhoods and what are the trending venues in those neighborhoods. We can also use other sources to check the eating trends of people in Toronto, CA but that’s just if we ever need it.

After we have the data, we can analyze as to where are the maximum number of theme parks or fun places to go but have lesser number of fast food restaurants? We can give recommendations to restaurant owners to put their restaurants in crowdy places because crowd is normally looking to eat something as well.

# Methodology:

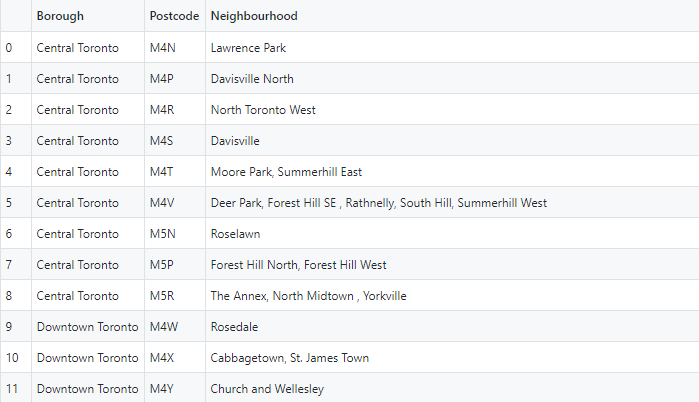
## Data Collection:

Firstly, the data of neighborhoods and boroughs along with their postal codes was scraped through Wikipedia site[[1]](#footnote-1) using Pandas and Beautiful Soup (Python Libraries). The Geospatial coordinates were provided through a CSV file which contained Toronto’s neighborhood postal codes and their longitudes and latitudes.

## Data Pre-Processing:

The data was very inconsistent so some null values had to be dropped and some neighborhoods which had null values with same boroughs were corrected so that data could be made inconsistent.

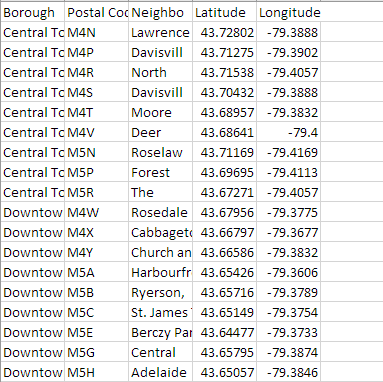
After cleaning the data, it looked like this:



The Geospatial Data looked like this:



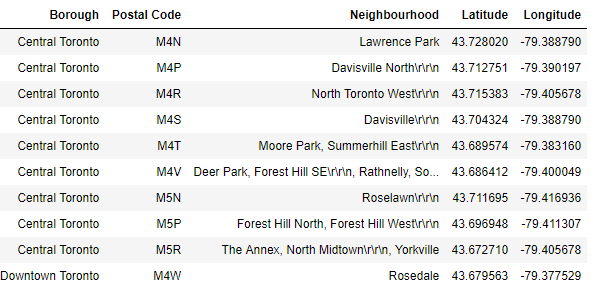
After the data was cleaned, the data had to be transformed so that we could have coordinates and neighborhoods together in one dataset. So through pandas, a new dataset was formed which looked like this:



Now the data was ready to be clustered so that we could analyze different neighborhoods.

## Data Analysis

Firstly, the data for Toronto was extracted from the main dataset like:



Now we had different boroughs like Central Toronto, West Toronto, North Toronto, Downtown Toronto etc.

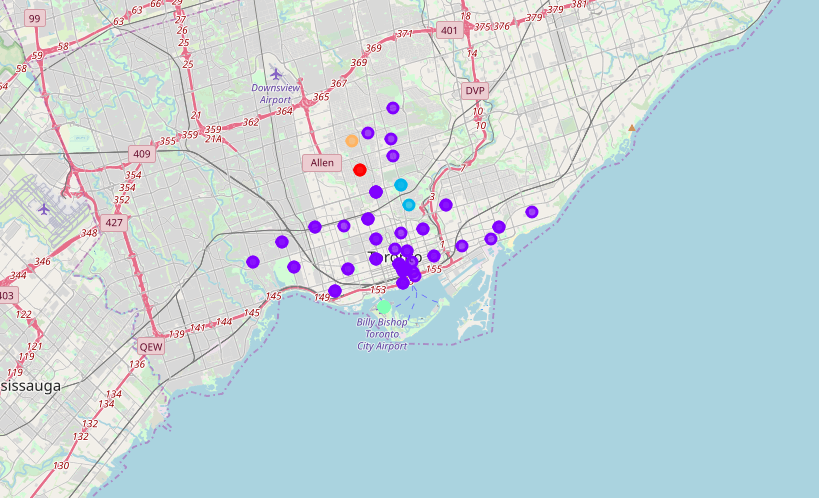
And they all belonged to Toronto because we were trying to find optimal location in Toronto for the restaurant owners.

After using Foursquare API, venues around the neighborhoods were imported and after some data manipulation and transportation, the final dataset was formed for clustering which had neighborhoods, boroughs and its 10 most common venues. It looked like this:



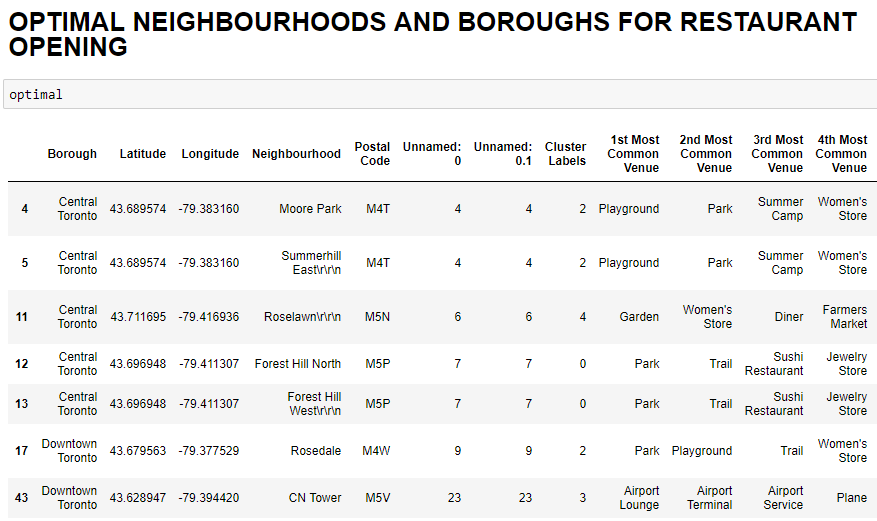
After the dataset was ready, we put it up for K-means clustering and tried total numbers of clusters to be 5 because it was found optimum after number of tries.

Then we applied cluster labels to the data so that every neighborhood is labelled with its cluster and then the clusters were put on a map using Folium library



After the clusters were formed, each cluster was observed carefully and those clusters which had maximum venues, but lower number of eating places were put together as a single dataset and those 10-12 neighborhoods were put on recommendation list as the optimal location for opening a new restaurant. The image will be put under the next heading.

# Results:



The result showed that almost every neighborhood in Downtown Toronto and some in Central Toronto will prove to be good for any restaurant owner who wants to open a fast food restaurant nearby or café to improve his overall business.

# Discussions:

Following suggestions could be made based on the neighborhoods that were found to be optimal for a restaurant business:

* In **Moore Park** and **Summerhill East** in **Central Toronto,** any restaurant (fast food, sushi etc.) will prove to be beneficial as the number of restaurants or even cafes in these two neighborhoods are relatively low.
* In **Roselawn** and **Forest Hill North and West** in **Central Toronto,** fast food restaurant will prove to be better than any other type of food business because of lack of any famous fast food restaurant
* In **Downtown Toronto,** there are few to zero restaurants so a good quality restaurant will prove to be profitable as there will be no competition in those neighborhoods

# Conclusion:

This overall data analysis will prove to be beneficial for anyone who is trying to open a new restaurant in Toronto somewhere. This will provide enough insights for restaurants to make smarter decision and avoid any loss in business while investing in a new place.

# For Future:

The same data with neighborhood could be used for other business who are trying to invest their money in food business, transport business or even want to develop parks or boutiques.

1. <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M> [↑](#footnote-ref-1)