## Battle of neighborhoods

## Introduction: Background & Problem Description

Toronto, the provincial capital of Ontario, Canada, is the most populous city in Canada. Today tourism is one of the major pillars of economy for any country. People visit to the places, which are rich in heritage and developed from Foreign prospective. In current date, information is so common regarding any location that, the more the information available about a place, the place is easier to explore. Information availability is the major criteria for a tourist to choose place of travel. Food availability and quality serves a valid criteria for choosing the place. There are hundreds of locality in and around Toronto approximately.

It is impossible to run an analysis for each type of neighborhood and provide results. I have chosen the neighborhood of Toronto. The study of other types of restaurants can be conducted with the same method.

## Data Preparation

Data used in the analysis are listed below:

· Postal codes of Toronto -- [**https://en.wikipedia.org/wiki/List\_of\_postal\_codes\_of\_Canada:\_M**,](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M) I cleaned the data and reduced it to boroughs of Toronto so that I can use it to find geological locations for further venue analysis.

· Using Geopy to get geological location by address name.

Geospatial data- [**http://cocl.us/Geospatial\_data**](http://cocl.us/Geospatial_data)

Pandas were used to transform the data in the table on the Wikipedia page into the above *pandas*data frame.

* The data frame will consist of three columns: Postal Code, Borough, and Neighborhood
* Only process the cells that have an assigned borough. Ignore cells with a borough that is not assigned.
* More than one neighborhood can exist in one postal code area. For example, in the table on the Wikipedia page, you will notice that M5A is listed twice and has two neighborhoods: Harbourfront and Regent Park. These two rows will be combined into one row with the neighborhoods separated with a comma as shown in row 11 in the above table.
* If a cell has a borough but a not assigned neighborhood, then the neighborhood will be the same as the borough.
* Clean your Notebook and add Markdown cells to explain your work and any assumptions you are making.
* In the last cell of your notebook, use the .shape method to print the number of rows of your data frame.

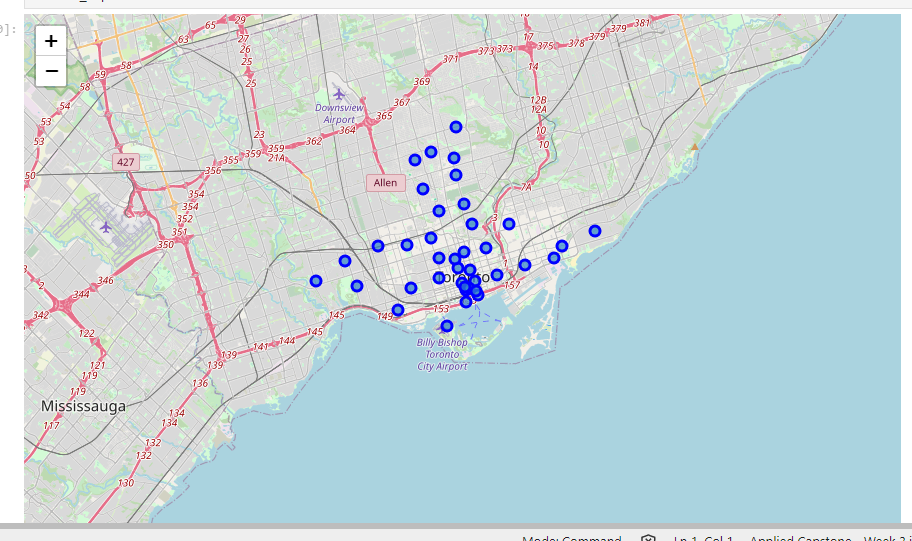
## Methodology and Analysis

After transforming the URL data of Postal code of neighborhoods of Toronto through data frames using Pandas, I used Geopy to get the geological location of each neighborhood. Then I merged both the data frames for better clarity. It returned 19 records in total.



Then neighborhood of Toronto were filtered. 

Then I used folium to visualize the distribution of the neighborhood in Toronto as below:



## Conclusion

After clustering the data of the respective neighborhoods, both cities (Boroughs) have venues which can be explored and attract the Tourists. The neighborhoods are much similar in features like Theaters, opera houses, food places, clubs, museums, parks etc. As far as concern to dissimilarity, it differs in terms of some unique places like historical places and monuments.

As far as concern to recommendations, we recommend Downtown Toronto Neighborhoods will be considered first to visit. The tourists have an easily travelling access due to Airport facility, which not only saves time but also helps to save money. This saved money can be utilized to explore more, the attracting venues