# The Battle of Neighborhoods

## Introduction

This project is designed to make a comparison between 2 mega cities New York and Toronto. This project may benefit large company or specific individuals in their decision to make a choice between 2 cities. Companies nowadays are trying to attract best employees with highest skill sets by carefully picking their location not just by having direct economic benefits like taxes and wage rate but also, they are looking at quality life and environment that will attract best employees. In the same time, very often job seekers are willing to relocate to the place where they would have comfortable and good living. We will be focusing on Scarborough Neighborhood of Toronto and Queens of NY City.



## Data

The data that is going to be used in this project is going to come from the following data sources:

Toronto Neighborhoods - https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M.
Toronto Latitude and Longitude - http://cocl.us/Geospatial_data
New York City neighborhoods - https://geo.nyu.edu/catalog/nyu_2451_34572

New York City Latitude and Longitude = Python Geolibrar
Foursquare API search feature will be used to collect neighborhood venue information.

The following libraries will be used to preprocess and visualize data:
Pandas
JSON
Requests.json
Matplotlib
Sklearn
Folium

## Methodology:

### Workflow:
1. Using dataset to retrieve Zip Codes/Postal Codes for both Toronto and NY
2. Preprocess data. For Toronto we are going to ignore cells with a borough that is Not assigned. Combining neighborhoods that exists in one postal code and change unassigned Neighborhood to its Borough's name. For NY we are going to get data from  https://geo.nyu.edu/catalog/nyu_2451_34572 and convert json format data into pandas and assign geo locations for both Toronto and NY.
3. Foursquare API will give us most popular venues for each location
4. By using Folium Python visualization library output maps for both cities.
5. By using k-mean algorithms we are going to cluster each areas and determine top 5 venues.

### Results Scarborough Borough in Toronto, Canada

By using k-means to group the neighborhoods in Scarborough into 3 clusters. Cluster_0 has 15 neighborhoods and the most common venues are skating rinks, international cuisine restaurants and



breakfast spots.  Cluster 1 has 1 neighborhood 1 neighborhood, and the most common venues are pizza place and noodle house. Cluster 2 has 1 neighborhood, and the most common venues are Chinese restaurants and discount stores.
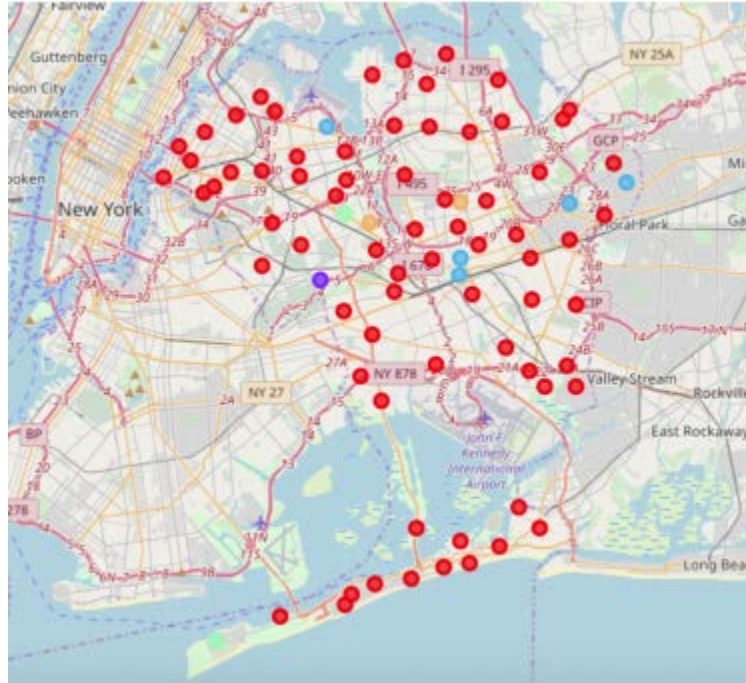
**Queens Borough in New York City**

By using k-means to group the Queens borough into 5 clusters. Cluster_0 has 81 neighborhoods and consist of many international cuisine restaurants and grocery stores. The most common venues are pizza places, deli, and Chinese restaurants.

Cluster_1 has 1 neighborhood and the most common venue is a dance studio.

Cluster_2 has 5 neighborhoods and the most common venue are donut shops and international cuisine restaurants.

Cluster_3 has 2 neighborhoods and the most common venues are the beach and a bakery.

Cluster_4 has 2 neighborhoods and the most common venues are gyms and donut shops.



## Discussion

Toronto has 11 boroughs and 103 neighborhoods. The geographical coordinate of Toronto, Canada are 43.7170226, -79.4197830350134. In Scarborough borough, found 85 venues in 17 neighborhoods In Scarborough borough, the neighborhoods with the most venues are L'Amoreaux West and Steeles West.

There are 79 distinct venues in 50 categories.

New York City has 5 boroughs and 306 neighborhoods. The geographical coordinate of New York City are 40.7308619, -73.9871558. Foursquare found 2108 venues in 81 neighborhoods in Queens borough.

Many of the neighborhoods are homogenous and are very similar to each other. Both Scarborough and Queens borough consist of neighborhood cluster that contain majority of the neighborhoods, and the remaining cluster had 1-5 neighborhoods. Queens borough had a significantly more number of neighborhoods and venues than Scarborough.

## Conclusion

In conclusion, based on the quantity of venues and variety of venues, I would choose Queens over Scarborough as a choice to establish presence for a company that wants to attract more employees.

Queens offer way more in choices for restaurants, gyms, grocery stores, and extracurricular activities for individuals and families of the company's workers.