**Battle of the Neighborhoods**

**Introduction:**

A family has the opportunity to move their business into either Toronto, Canada or New York City, USA, but first they want to check which place would suit them better. So, this project will study the neighborhoods in Toronto, Canada and New York City, USA and compare local business in the cities so that the family has a good life quality and room to expand. We will explore similarities and difference between neighborhoods in these two cities to see which one suits the family best.

**Data**:

The data used for this project will be acquired from the cities Wikipedia pages. The datasets consist of the postal codes, neighborhood names, latitude, and longitude information for each neighborhood. Foursquare API search will be used to collect neighborhood venue information. Details about local venues and locality will be provide insight into the qualities of a neighborhood. In addition to Foursquare, various python packages will be used to create maps and machine learning models to further provide insights into our neighborhood battle project.

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

**Methodology**:

1. HTTP requests would be made to this Foursquare API server using zip codes of the Seattle city neighborhoods to pull the location information (Latitude and Longitude).

2. Foursquare API search feature would be enabled to collect the nearby places of the neighborhoods. Due to http request limitations the number of places per neighborhood parameter would reasonably be set to 100 and the radius parameter would be set to 700.

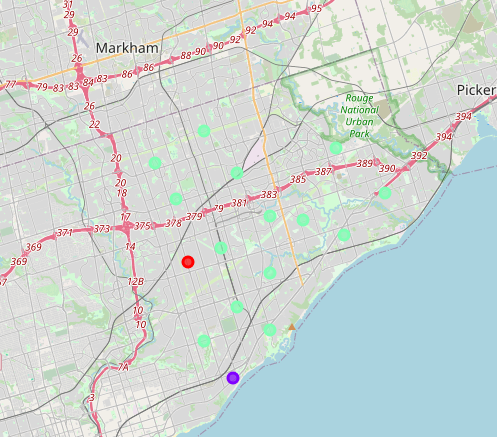
3. Folium- Python visualization library would be used to visualize the neighborhoods cluster distribution of Seattle city over an interactive leaflet map.

4. Extensive comparative analysis of two randomly picked neighborhoods world be carried out to derive the desirable insights from the outcomes using python’s scientific libraries Pandas, NumPy and Scikit-learn.

5. Unsupervised machine learning algorithm K-mean clustering would be applied to form the clusters of different categories of places residing in and around the neighborhoods. These clusters from each of those two chosen neighborhoods would be analyzed individually collectively and comparatively to derive the conclusions.

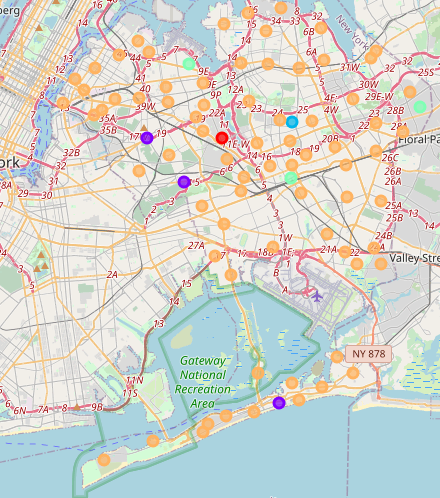
**Results:**

**Scarborough Borough in Toronto, Canada**



I use k-means to group the neighborhoods in Scarborough into 3 clusters. Cluster\_0 has 1 neighborhood and the most common venues are Accessories Stores, Auto garages, and Bakeries. Cluster\_1 has 1 neighborhood 1 neighborhood, and the most common venues are General Entertainment, Skating Rinks, and Cafes. Cluster\_2 has 166 neighborhoods, and the most common venues are Fast Food restaurants and Bakeries.

**Queens Borough in New York City**



I used k-means to group the Queens borough into 5 clusters. Cluster\_0 has 1 neighborhood and the most common venue is Gym. Cluster\_1 has 3 neighborhood and the most common venues are pizza places, Deli, and Beaches. Cluster\_2 has 1 neighborhood and the most common venue is Deli. Cluster\_3 has 3 neighborhoods and the most common venues are Donuts shops, Pizza places, and Indian restaurants. Cluster\_4 has 81 neighborhoods and the most common venues are Pharmacies, beaches. Deli, and Bakeries.

**Discussion**

Toronto has 11 boroughs and 103 neighborhoods. The geographical coordinate of Toronto, Canada are 43.7170226, -79.4197830350134.

In Scarborough borough, found 89 venues in 17 neighborhoods In Scarborough borough, the neighborhoods with the most venues are L’Amoreaux West and Steeles West. There are 80 distinct venues in 55 categories.

New York City has 5 boroughs and 306 neighborhoods. The geographical coordinate of New York City is 40.7308619, -73.9871558. Foursquare found 2103 venues in 81 neighborhoods in Queens borough. Many of the neighborhoods are homogenous and are very similar to each other. There are 1736 distinct venues in 277 categories

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 greater 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. Queens offer way more in choices for restaurants, gyms, grocery stores, and extracurricular activities for individuals and families.