The Battle of Neighborhoods

By Judith Duinisveld

Toronto

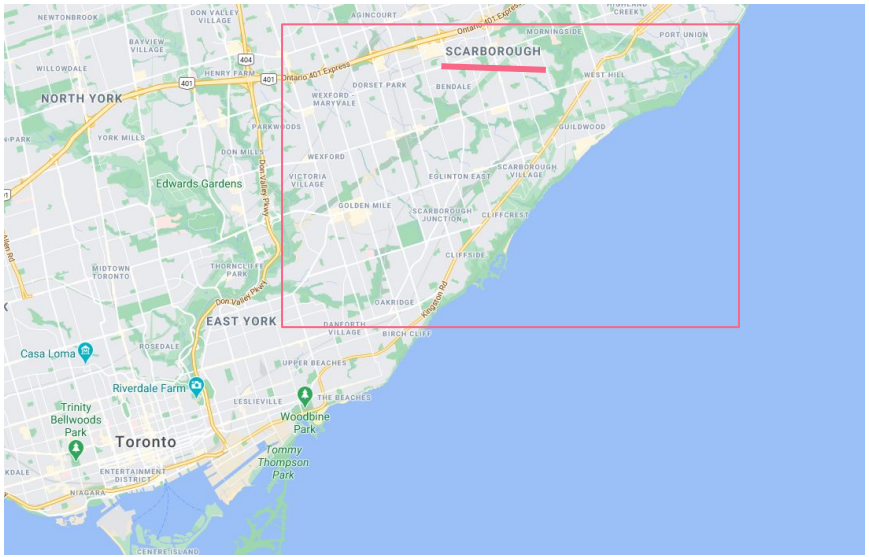
New York

Introduction

- This project **will analyze neighborhoods** between Scarborough borough Toronto (Canada) & Queens borough New York City (USA)
- A **Fortune 500 company** is looking to move its HQ to either Toronto or New York City
- In order to do so the company **wants insight** into the neighborhoods and local businesses in the cities for **optimal living standards** and **quality of life** for their employees.
- This project describes the **exploration of the (dis)similarities** between certain neighborhoods in the two cities, so we can determine which neighborhood best fit the culture of the employees.

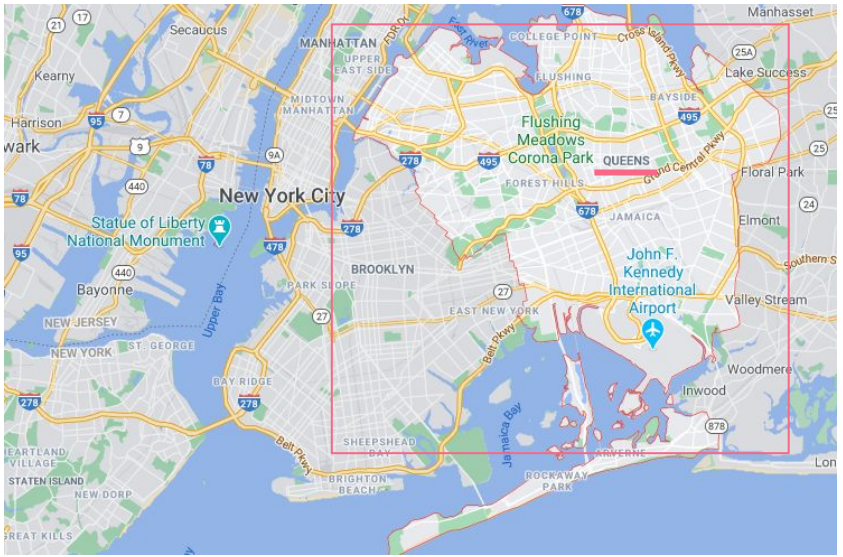
Toronto

Scarborough borough



New York City

Queens borough



Problem statement

- ➔ Understand the similarities and differences of neighborhoods between Scarborough borough in Toronto and Queens borough in New York City
- ➔ Select the best neighborhood for a Fortune 500 company to move its headquarters (next: HQ) based on venues

Data

- The data used for this project will be acquired from the respective cities Wikipedia website pages
- The datasets consists for each neighborhood of:
  - Postal codes
  - Neighborhoods names
  - Latitude and longitude information
- Foursquare API search feature will be used to collect neighborhood venue information
- In addition to Foursquare, various Python packages will be used to create maps and machine learning models to further provide insights

I used the following datasets from these websites:

1. Toronto Neighborhoods: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
2. Toronto Latitude and Longitude: http://cocl.us/geospatial_data
3. New York City neighborhoods: https://geo.nyu.edu/catalog/nyu_2451_34572
4. New York City Latitude and Longitude: Python Geolibrar

Methodology

1. **HTTP request** → to pull the location information (latitude and longitude) to Foursquare API server using zip codes.

2. **Enabling Foursquare API Search future** → 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 over an interactive leaflet map.

4. **Comparative analysis** → using Python's scientific libraries Pandas, NumpPy and Scikit-learn so that two randomly picked neighborhoods would be carried out to derive the desirable insights from the outcomes

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

Used Python packages

1. Pandas → library for Data Analysis
2. NumPy → Library to handle data in a vectorized manner
3. JSON → Library to handles JSON files
4. Geopy → to retrieve location data
5. Requests → Library to handle http requests
6. Matplotlib → Python Plotting Module
7. Sklearn → Python machine learning library
8. Folium → Map rendering library

Results

Scarborough Borough in Toronto

I used k-means to group the neighborhoods in Scarborough into 3 clusters.

Cluster_0 has 15 neighborhoods and has a wide variety of venues such as:

- Train station
- Coffee shop
- Smoke shop
- Gym
- Grocery store
- Park
- Playground
- Pharmacy
- Breakfast spot
- Bank
- Soccer field
- Gas station
- Latin American Restaurant, Hakke restaurant, Vietnamese restaurant, Indian, Korean BBQ Restaurant, Mexican
- Bus line
- -Ice Cream Shop
- -Fast Food
- -General entertainment
- -College Stadium
- -Fried chicken Joint

Results

Scarborough Borough in Toronto

Cluster_1 has 1 neighborhood with the following venues:

- Hakka restaurant
- Thai Restaurant
- Athletics & Sport
- Bakery
- Bank
- Gas Station
- Fried Chicken Joint
- Caribbean Restaurant
- College Stadium
- Gym

Results

Scarborough Borough in Toronto

Cluster_2 has also neighborhood with the following venues:

- General Entertainment
- Skating Rink
- Cafe
- College Stadium
- Vietnamese Restaurant
- Clothing Store
- Gym
- Grocery Store
- Gas Station
- Fried Chicken Joint

Scarborough borough (17 neighborhoods)

Results

Queens Borough in New York City

I used k-means to group the neighborhoods in Scarborough into 5 clusters.

Cluster_0 has 6 neighborhoods with venues:

- Pizza Place -Home service
- Bank -All kinds of restaurant: Sushi, Falafel
- Bakery -Supplement store
- Mobile shop -Eye doctor
- Grocery stroe -Discount store
- Sport Bar -Farm
- Convenience store
- Food truck -Arts & crafts store
- Shoe store -Playground
- Bus stop -Brewery
- Dog run -Fried Chicken Joint
- Gym/Fitness Center
- Donut shop
- Yoga Studio

Results

Queens Borough in New York City

I used k-means to group the neighborhoods in Scarborough into 5 clusters.

Cluster_1 has 1 neighborhood with venues:

- Gym/Fitness Center
- Yoga Studio
- Pizza Place
- Convenience Store
- Park
- Thai Restaurant
- Pharmacy
- Optical Shop
- Chinese restaurant

Results

Queens Borough in New York City

I used k-means to group the neighborhoods in Scarborough into 5 clusters.

Cluster_2 has 70 neighborhoods with venues:

- Restaurants: Middle eastern, Seafood, Mediterranean, Indian, Peruvian, South American, Mexican, Thai, Sushi, Fast food, Filipino, Falafel, Italian, Spanish
- Baseball field
- Park
- Dance Studio
- Playground
- Hotel
- Tennis Courts
- Bus Line
- Rest Area
- Yoga Studio
- Fish Market
- Farm
- Bank
- Bakery
- Shoe store

Results

Queens Borough in New York City

I used k-means to group the neighborhoods in Scarborough into 5 clusters.

Cluster_3 has 3 neighborhoods with venues:

- Restaurants: Falafel, Caribbean, Mexican, Pizza Place
- Donut Shop
- Indie Movie Theater
- Lake
- Supermarket
- Grocery Store
- Dosa Place
- Basketball Courts
- Performing Arts Venue
- Clothing store
- Mobile Phone Shop

Results

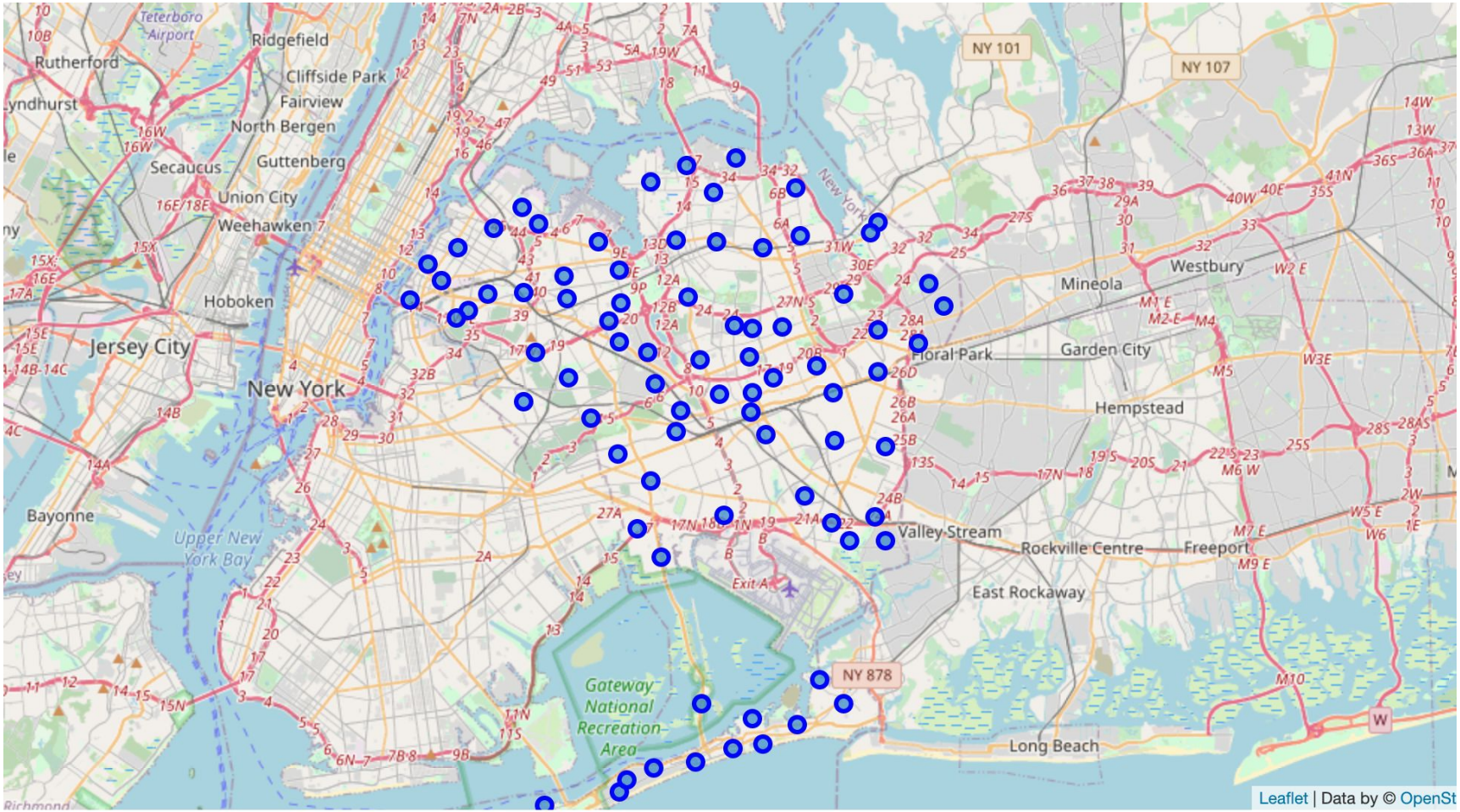
Queens Borough in New York City

I used k-means to group the neighborhoods in Scarborough into 5 clusters.

Cluster_4 has 1 neighborhood with venues:

- Restaurants: South American, Pizza Place
- Deli / Bodega
- Automotive Shop
- Playground
- Rental Car Location
- Spa
- Grocery Store
- Basketball Court

Queens borough (81 neighborhoods)



Discussion

Scarborough borough

Toronto has 11 boroughs and 103 neighborhoods.

Scarborough borough has total 80 distinct venues in 55 categories.

Queens borough

New York City has 5 boroughs in 206 neighborhoods.

Queens has 1733 distinct venues in 277 categories.

Queens borough has a significant more number of venues and neighborhoods than Scarborough.

Conclusion

In conclusion, based on the quantity of (the variety of) venues, i would advise the Fortune 500 company, to relocate their HQ's to the Queens (New York City) instead of Scarborough (Toronto) .

Queens offer way more choices for restaurants, gyms, grocery stores and extracurricular activities for the employees (individuals and families).

New York