

Battle of the Neighborhoods

Introduction

A while ago, Amazon dropped the plans for New York Headquarters. The choice of a headquarter location is a result of serious and comprehensive consideration. One important aspect is whether the location is attractive to labors needed.

Based on the assumption that an XYZ company is looking for a new headquarter location in either Toronto, Canada or New York City, New York. This project will focus on the similarities and dissimilarities between certain neighborhoods in the two cities, especially concerning living quality of potential employees, and determine which neighborhood is a better choice.

Data

To help making the decision, we will need neighborhood data for both cities. The data are available from Wikipedia website and Python modules. The detailed online sources are listed below.

New York: https://geo.nyu.edu/catalog/nyu_2451_34572

Toronto: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
http://cocl.us/Geospatial_data

A second source of data is the Foursquare data. By Foursquare API we will be able to use collected information of neighborhood venues. We assume that the data from Foursquare is sufficient enough for the purpose of this project.

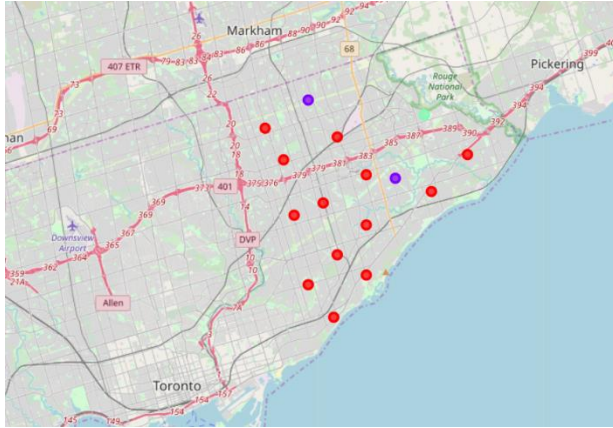
Methodology

By segmentation and clustering, we will compare the similarity and dissimilarity of both cities and provide classification information for decision. The detailed workflow is as follows.

1. Collect data for both cities into dataframe containing borough, neighborhood, geographical coordinate(latitude and longitude) information.
2. Data cleaning and feature extraction. We extract only boroughs of interest and, via the Foursquare API, search for the needed neighborhood venue information.
3. By Folium module, visualize clustering information for further analysis.
4. K-means clustering is used to form clusters of different categories of venues and the clusters will be used for further analysis.

Results

Scarborough in Toronto, Canada



There are 96 venues in Scarborough with 56 unique categories. The venues are divided into 3 categories by k-means clustering.

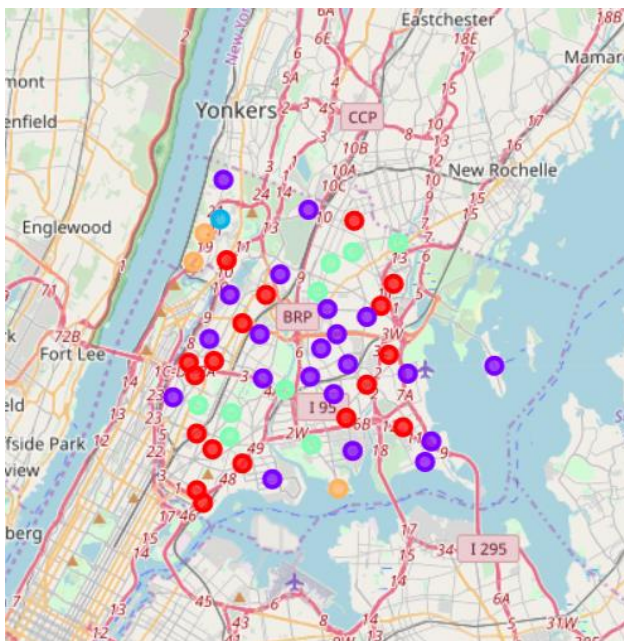
Cluster-1: 13 neighborhoods, 1st most common venue mainly including various restaurants.

Cluster-2: 2 neighborhood, 1st most common venue being coffee shop and playground.

Cluster-3: 1 neighborhood, 1st most common venue being fast food restaurant.

Among the 2nd and 3rd most common venue, Scarborough provides skating rink, convenience store, rental car location and so on.

Bronx in New York City



There are 1226 venues with 173 unique categories in Bronx. By k-means clustering, they are divided into 5 clusters.

Cluster-1: 18 neighborhoods, 1st most common venue mainly including Latin American restaurant, fast food restaurant, pizza place and pharmacy.

Cluster-2: 21 neighborhoods, 1st most common venue mainly including pizza place and Italian restaurant.

Cluster-3: 1 neighborhood, 1st most common venue being plaza.

Cluster 4: 9 neighborhoods, 1st most common venue mainly including grocery store and bus station.

Cluster 5: 3 neighborhoods, 1st most common venue being park.

Apart from the 1st common venues of the two cities, among the 2nd and 3rd common venues, Bronx also provides a relatively rich options of bank, gym, supermarket, pub, Chinese/Asian restaurant, Mexican restaurant, farmers/fish markets, donut shop, pub and coffee shops and so on.

Discussion and Conclusion

Bronx and Scarborough both provides various dining options. Comparing the two boroughs in the two cities, Bronx provides more option for office workers concerning daily life and social events. Scarborough fits better for college students with relatively rich dining options and limited option of stores.

As a conclusion, Bronx is a better fit for a new headquarter.