Coursera Capstone Project — The Battle of Neighborhoods

Relocation Conundrum

By Ravindra Sariya

Introduction

In this world of globalization people often moves from one location to another due to various reasons. According to the 2016 Yearbook of Immigration Statistics, the United States admitted a total of 1.18 million legal immigrants (618k new arrivals, 565k status adjustments) in 2016. If we combine this number with internal migration in the country, then number would be even higher. When people move to a new location, they spend considerable amount of time in finding the place to live in the new city. They take help of their friends, relatives, colleagues etc. if they are already living in the destination city. But many of them still get disappointed and wants to change the rented home or apartment within a few months of relocation. It affects person on many other aspects such as school of their kids, their membership on gym and yoga centers. So, they spend time and money to know the neighborhood, which they would like once they relocate. Even, people take help of external agents to help them to find the good neighborhoods.

Problem Statement

There is an ABC global relocation consultant firm, who advises people for their relocation needs in worlds major metropolitan cities. They charge 50% of their fees to clients at the time of recommending the place before moving to the new location and next 50% after 3 months of relocation. Using past data, they found that only in 60% of cases they get remaining fees after three months of the relocation. They also noticed that if they recommend the place like the place where customer currently staying then probability of getting due fees increased to 95%.

Few days back, they got a client who is moving to New York from Toronto. He is currently staying at The Beaches West neighborhood in Toronto. He suggested that he likes The Beaches West and India Bazar neighborhoods of Toronto compared to other Neighborhoods. He also requested to recommend the neighborhoods in the borough which has low crime rates. Basically, his idea was to get a small list of neighborhoods which he can explore by himself and take a decision based on other factors such as cost, distance from his employer etc. He has some reservation about Staten Island and do not want to go there.

Data Requirements

For this project we need the following data:

- **New York City data that contains the list of Boroughs, Neighborhoods along with their latitude and longitude.**
  Data source: https://cocl.us/new_york_dataset
- **New York City crime data that contains list of all the reported crimes in year 2018 as per Boroughs**
  Data source: Downloaded the csv file from following location:
  https://data.cityofnewyork.us/Public-Safety/NYPD-Complaint-Data-Historic/qgea-i56i
- **2020 Projected population of different boroughs**
  Data source: Downloaded the csv file from following location:
  https://data.cityofnewyork.us/City-Government/2020-population/t8c6-3i7b
- **Toronto City data that contains the list Boroughs, Neighborhoods**
  Data source: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- **Geospatial data for Toronto to get latitude and longitude of all Boroughs and Neighborhoods**
  Data source: https://cocl.us/Geospatial_data
- **Data source: Fousquare API**

- Description: By using Foursquare api we will can explore all the venues in each neighborhood.

Methodology

As a consultant we wanted to recommend the neighborhoods in a borough which are similar in characteristics with current location of the client.
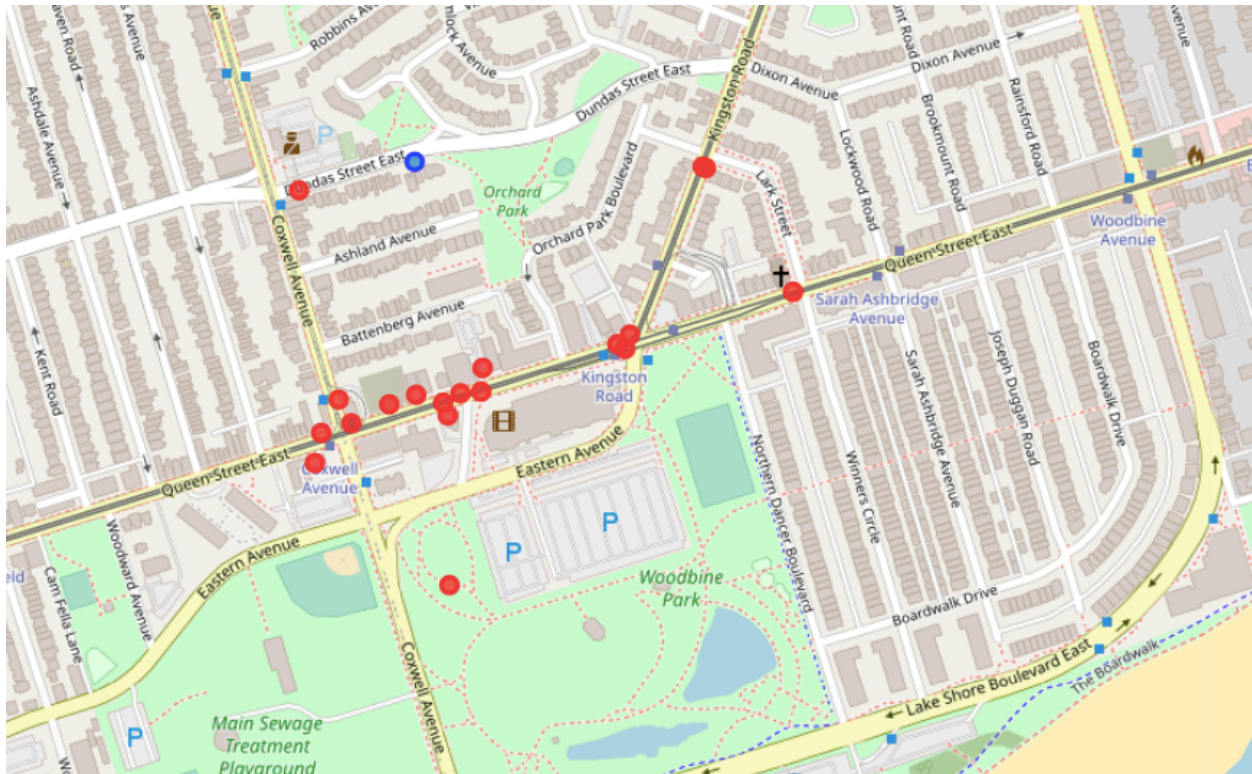
Exploring area around The Beaches West and India Bazaar

Client is from The Beaches West Toronto. Therefore, venues around the client location have been explored. First, we had web scraped the Toronto neighborhood data then obtained the coordinates of neighborhoods. We limited the data to The Beaches West and India Bazaar. Now, using the FourSquare API, all nearby venues being explored. We found 19 venues in these neighborhoods.

[17]:

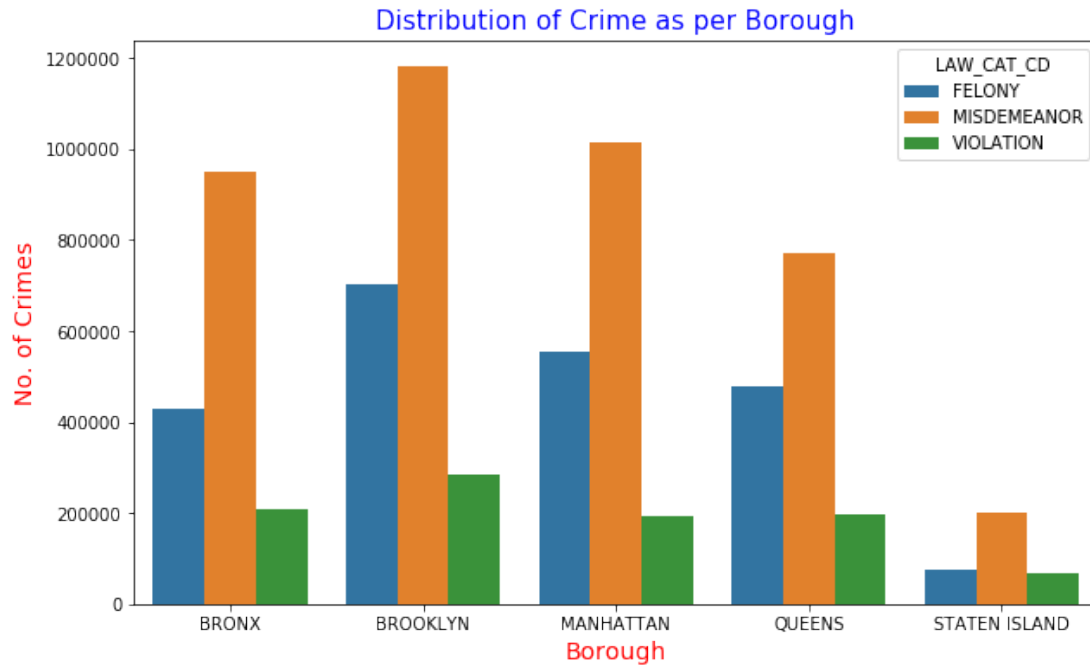|	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	India Bazaar, The Beaches West	43.668999	-79.315572	System Fitness	43.667171	-79.312733	Gym
1	India Bazaar, The Beaches West	43.668999	-79.315572	British Style Fish & Chips	43.668723	-79.317139	Fish & Chips Shop
2	India Bazaar, The Beaches West	43.668999	-79.315572	Brett's Ice Cream	43.667222	-79.312831	Ice Cream Shop
3	India Bazaar, The Beaches West	43.668999	-79.315572	The Burger's Priest	43.666731	-79.315556	Fast Food Restaurant
4	India Bazaar, The Beaches West	43.668999	-79.315572	O Sushi	43.666684	-79.316614	Sushi Restaurant
5	India Bazaar, The Beaches West	43.668999	-79.315572	Godspeed Brewery	43.672620	-79.319228	Brewery
6	India Bazaar, The Beaches West	43.668999	-79.315572	Woodbine Park	43.664860	-79.315109	Park
7	India Bazaar, The Beaches West	43.668999	-79.315572	LCBO	43.666732	-79.314966	Liquor Store
8	India Bazaar, The Beaches West	43.668999	-79.315572	Casa di Giorgio	43.666645	-79.315204	Italian Restaurant
9	India Bazaar, The Beaches West	43.668999	-79.315572	Pet Valu	43.666979	-79.314665	Pet Store
10	India Bazaar, The Beaches West	43.668999	-79.315572	Murphy's Law	43.667319	-79.312656	Pub
11	India Bazaar, The Beaches West	43.668999	-79.315572	The Tulip Steakhouse	43.666348	-79.316854	Steakhouse
12	India Bazaar, The Beaches West	43.668999	-79.315572	Burrito Bandidos	43.666445	-79.316447	Burrito Place
13	India Bazaar, The Beaches West	43.668999	-79.315572	Alliance Cinemas - The Beach	43.666747	-79.314685	Movie Theater
14	India Bazaar, The Beaches West	43.668999	-79.315572	Harvey's	43.666528	-79.315127	Restaurant
15	India Bazaar, The Beaches West	43.668999	-79.315572	Subway	43.666052	-79.316933	Sandwich Place
16	India Bazaar, The Beaches West	43.668999	-79.315572	KFC	43.666624	-79.315916	Fast Food Restaurant
17	India Bazaar, The Beaches West	43.668999	-79.315572	Pizzaiolo	43.668953	-79.311683	Pizza Place
18	India Bazaar, The Beaches West	43.668999	-79.315572	Jody's Deli	43.668946	-79.311635	Sandwich Place
19	India Bazaar, The Beaches West	43.668999	-79.315572	Supperworks	43.667737	-79.310452	Food & Drink Shop

MAP

Study on crime rates in different Boroughs of New York

As our client has requested for the borough which has low crime rates. We have obtained the yearly crime and population data of boroughs from city of New York official website.

Crimes are divided in three major categories Felony, Misdemeanor and violation. Out of these three, serious crimes fall under Felony. So, we have restricted our study on crime type Felony. We have obtained crime rates as per borough as shown in following table.

Crime rate for different boroughs:

	Borough	Felony_Crime	Population	Felony_Crime/Population
0	BRONX	427916	1446788	0.295770
1	BROOKLYN	703224	2648452	0.265523
2	MANHATTAN	555234	1638281	0.338913
3	QUEENS	476811	2330295	0.204614
4	STATEN ISLAND	74859	487155	0.153666

Distribution of Crime as per Borough

Type of crimes in Felony Category:

```
: OFNS_DESC
  ABORTION                            7
  ARSON                           17316
  BURGLARY                       246395
  CHILD ABANDONMENT/NON SUPPORT     450
  CRIMINAL MISCHIEF & RELATED OF 122513
  DANGEROUS DRUGS                 84491
  DANGEROUS WEAPONS               72593
  ENDAN WELFARE INCOMP               94
  FELONY ASSAULT                 280938
  FELONY SEX CRIMES                  14
  FORGERY                         74369
  GAMBLING                          138
  GRAND LARCENY                  634653
  GRAND LARCENY OF MOTOR VEHICLE 128648
  HOMICIDE-NEGLIGENT,UNCLASSIFIE    164
  HOMICIDE-NEGLIGENT-VEHICLE        106
  INTOXICATED/IMPAIRED DRIVING       64
  KIDNAPPING                         54
  KIDNAPPING & RELATED OFFENSES    2875
  KIDNAPPING AND RELATED OFFENSES     4
  MISCELLANEOUS PENAL LAW        172497
  NYS LAWS-UNCLASSIFIED FELONY     6988
  OTHER STATE LAWS (NON PENAL LA    183
  POSSESSION OF STOLEN PROPERTY   13058
  PROSTITUTION & RELATED OFFENSES   253
  RAPE                            21583
  ROBBERY                        262739
  SEX CRIMES                      17885
  THEFT-FRAUD                     76972
```
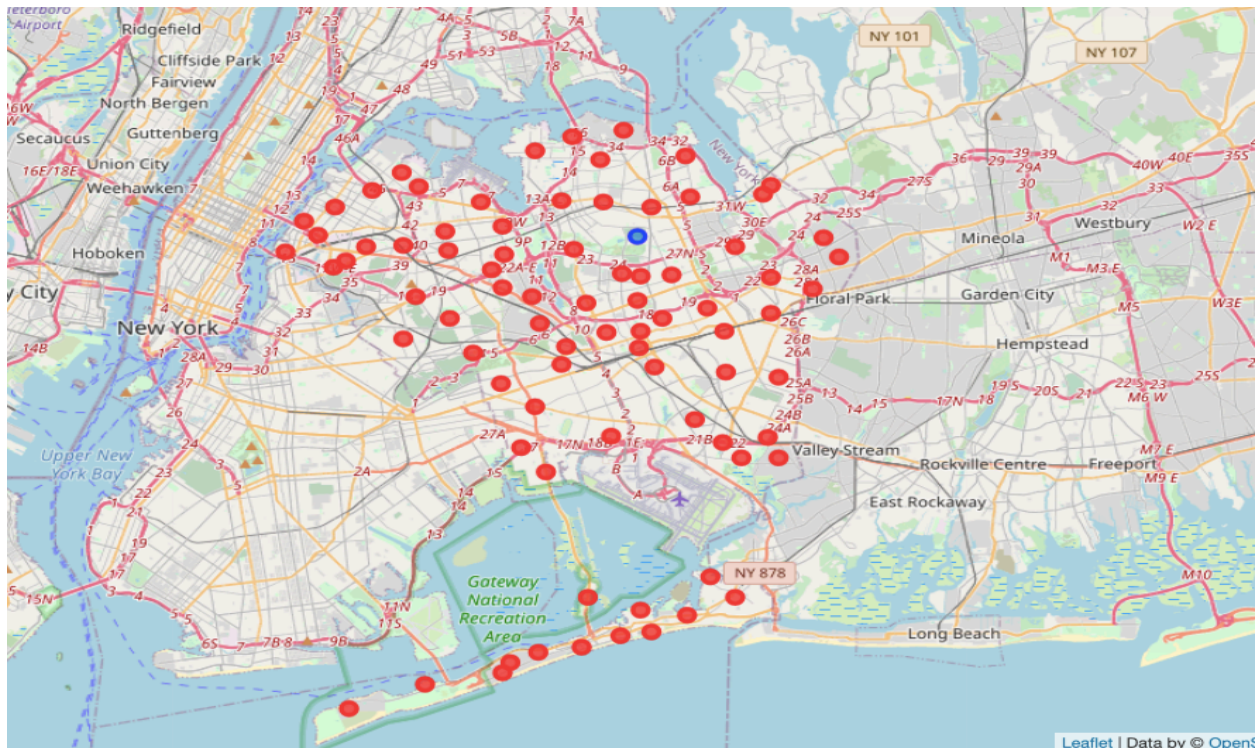
From above analysis, crime rate for serious crimes are lowest in Staten Island and highest in Manhattan. As our client has some reservation to not go for Staten Island, Queens is the next best choice for our analysis.

Explore venues around all the neighborhoods in Queens

We have decided to go for Borough Queens. This Borough has about 82 Neighborhoods. Map for the same is shown below.

We have fetched the New York Neighborhood data and geo location and filtered the same for Queens. We have used Four Square API to explore the neighborhoods. There are total 1534 venues in all the neighborhoods of Queens. Distribution of venues as per categories are as shown below.

	Venue Category	Venue
0	Accessories Store	2
1	Afghan Restaurant	2
2	American Restaurant	16
3	Antique Shop	1
4	Arepa Restaurant	3
5	Argentinian Restaurant	2
6	Arts & Crafts Store	2
7	Asian Restaurant	9
8	Athletics & Sports	2
9	Automotive Shop	2
10	BBQ Joint	4
11	Bagel Shop	20
12	Bakery	49
13	Bank	37
14	Bar	21
15	Baseball Field	4
16	Basketball Court	5
17	Beach	29
18	Beach Bar	2
19	Beer Garden	2

We have used one hot encoding method to convert categorical values into integer and obtained top 10 venues for all the neighborhoods.

Our goal is to find the list of neighborhoods in Queens which are similar with client's current location in Toronto. Therefore, we will merge the neighborhood details for The Beaches West and India Bazaar in the Queens neighborhoods details. Again, used one hot encoding to convert categorical values into integers.

K-Means Clustering

In K-means clustering it is observed that if there are multiple outliers then number of clusters would have only one values in them and majority of observations falls into single category. That would not lead us to our intended clustering results.
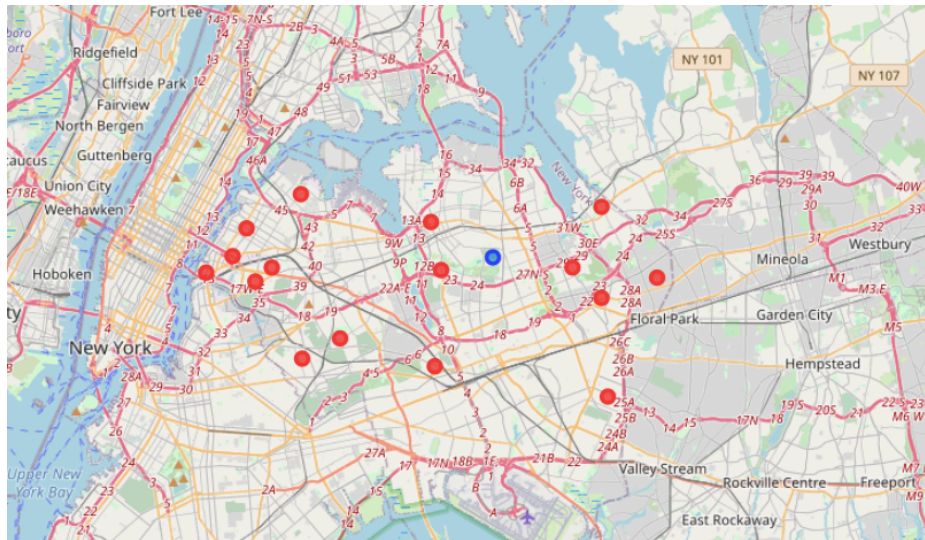
To avoid the same, we have done the clustering in two steps. First, to remove all the outlier neighborhood and get a set of neighborhoods which further can be divided into clusters.

In the first step, we could get 37 neighborhoods in one cluster including The Beaches West and India Bazaar which is client's current location. In next stage, we had again performed clustering on this group and found that The Beaches West and India Bazaar is parts of other 16 neighborhoods of Queens.

Results

Using clustering we have obtained the concise list of 16 neighborhoods which are similar with client's current neighborhood is present.

Map for all the sixteen neighborhoods:

Conclusion:

As per the requirement given by our client, we recommend him the list of final 16 neighborhoods in Queens. We have done our study based on client's current preference of neighborhoods in Toronto.

List of Sixteen neighborhoods:

#	Name of Neighbourhood
1	Queensboro Hill
2	Oakland Gardens
3	Bellaire
4	Floral Park
5	Little Neck
6	Flushing
7	Kew Gardens
8	Middle Village
9	Ridgewood
10	Hunters Point
11	Sunnyside
12	Long Island City
13	Sunnyside Gardens
14	Ravenswood
15	Steinway
16	Cambria Heights