

**Capstone Project Report**  
**The Battle of Neighborhoods**

**By**  
**Samira Gholizadeh**

15.June.2020

## 1. Business Problem section

### ➤ Introduction

One of the most common business problems that can affect the success of a business is location. Some hotels or cinemas are next to each other or close to highways, schools or shopping malls. There are some factors to influence the customer targeted group along with business purposes. Where each business should locates? Should they locate close to schools, highways or far away? Hotels are the most convenient place for tourists and visitors the stay. Before reservation, some people look at the price and locations, and others look at the amenities.

### ➤ Business Problem

As the topic for our project should be related to "battle of neighborhoods", I've decided to compare two cities of the New York City. In this scenario, it is urgent to adopt machine learning tools in order to assist visitors to find their suitable hotels. As a result, the business problem we are currently posing is: how could we provide support to visitors to reserve a suitable hotel in New York in this uncertain economic and financial scenario? If any investor is going to build new hotel, where is the best location?

To solve this business problem, we are going to cluster Brooklyn and Queen District of the New York City neighborhoods in order to recommend venues and the current average price of hotels where visitors can reserve their rooms. We will recommend profitable venues according to amenities and essential facilities surrounding such bars, restaurants, gym and fitnesses and etc.

### ➤ Data section

Since we have already looked at New York and Toronto's neighborhood data, I am going to accomplish my project on comparing two cities of New York. Besides of the New York neighborhoods dataset, I will use the population and capita of each neighborhood for the city. I used the neighborhoods dataset provided in the lab to bring in location (latitude, longitude) information of the two boroughs. And use FourSquare to generate maps. Data for capita and population were extracted from ([https://en.wikipedia.org/wiki/Boroughs\\_of\\_New\\_York\\_City](https://en.wikipedia.org/wiki/Boroughs_of_New_York_City)).

**Table 1. Population and capita**

	Borough	Population	GDP	per capita	square miles	persons /sq.mi
0	Bronx	1471160	28.787	19570	42.10	34653
1	Brooklyn	2648771	63.303	23900	70.82	37137
2	Manhattan	1664727	629.682	378250	22.83	72033
3	Queens	2358582	73.842	31310	108.53	21460
4	Staten Island	479458	11.249	23460	58.37	8112

## ➤ Methodology section

The Methodology section will describe the main components of our analysis and predication system. The Methodology section comprises four stages:

- a) Collect Inspection Data
- b) Explore and Understand Data
- c) Data preparation and preprocessing
- d) Modeling

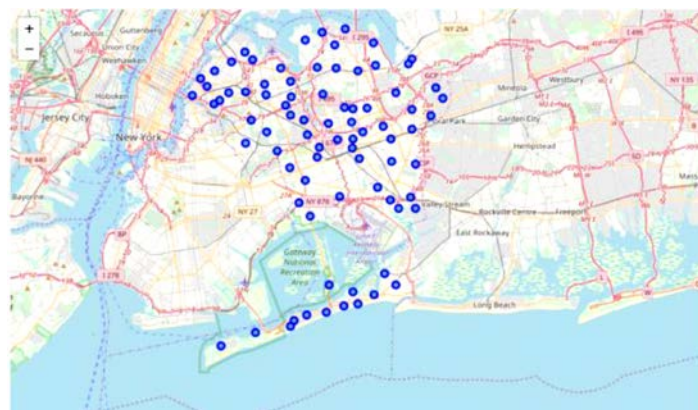
From the modelling of the datasets, data from Foursquare the following details can be considered:

### 1. Neighborhoods distribution

Comparing Brooklyn's (Map 1) and Queens' (Map 2) neighborhoods distribution, we can see that Brooklyn is one big area while Queens are one big area plus one small piece separated by Gateway National Recreation Area.



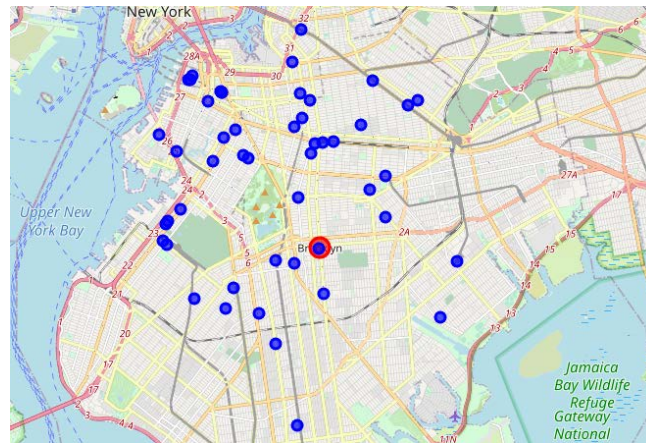
Brooklyn (Map 1)



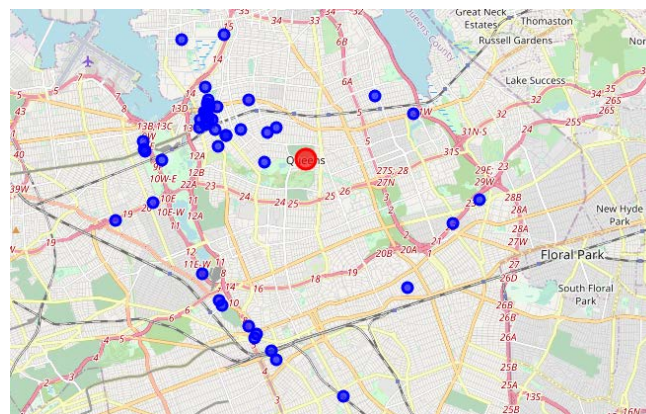
Queen (Map 1)

## 2. Current hotels distribution

From the current hotels distribution at Brooklyn (Map3) and Queens (Map 4), we can find that hotels are more evenly spread out in Queen than Brooklyn.



Brooklyn (Map3)



Queens (Map 4)

## 3. Hotels types

When I group by all the current hotels at Brooklyn and Queens by categories, the results are quite similar. The difference is in Queen there is Asian and Indian restaurants as well as 29 hotels in the area while Brooklyn shows 27 hotels and 2 hostels. Although, Queens' hotels spread out hotels in the area, but it seems it has visitors from Asia.

## “Brooklyn Hotels”

[24]:

	name	address	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	crossStreet	neighborhood
categories														
	Breakfast Spot	1	0	1	1	1	1	0	1	1	1	1	1	0
	Building	1	1	1	1	1	1	1	1	1	1	1	1	1
	Convenience Store	1	1	1	1	1	1	1	1	1	1	1	1	1
	Dive Bar	1	0	1	1	1	1	0	1	1	1	1	1	0
	General Travel	1	0	1	1	1	1	0	1	1	1	1	1	0
	Gym / Fitness Center	1	1	1	1	1	1	1	1	1	1	1	1	0
	Historic Site	1	0	1	1	1	1	0	1	1	1	1	1	0
	Hookah Bar	1	0	1	1	1	1	1	1	1	1	1	1	0
→	Hostel	2	1	2	2	2	2	1	2	2	2	2	2	0
→	Hotel	27	22	27	27	27	27	24	27	26	27	27	27	11
	Hotel Bar	3	1	3	3	3	3	1	3	3	3	3	3	0

## “Queen Hotels”

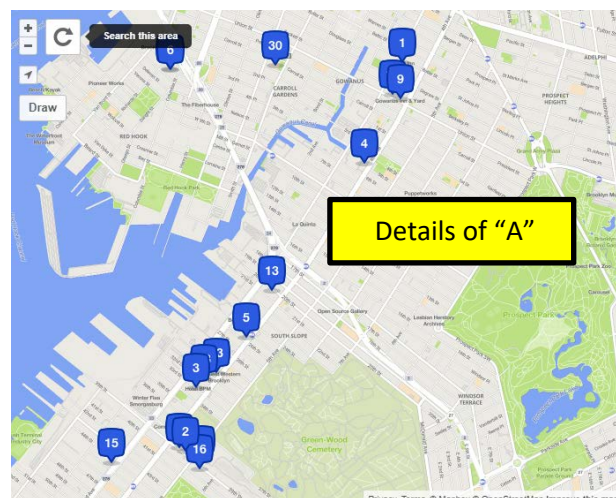
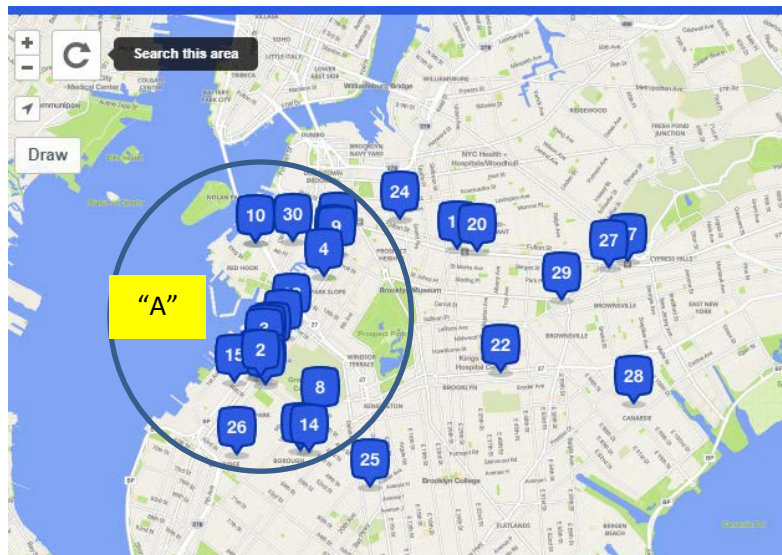
[36]:

	name	address	lat	lng	labeledLatLngs	distance	postalCode	cc	city	state	country	formattedAddress	crossStreet	neighborhood
categories														
	Apres Ski Bar	1	1	1	1	1	1	0	1	1	1	1	1	0
→	Asian Restaurant	1	1	1	1	1	1	1	1	1	1	1	1	0
	Bed & Breakfast	1	1	1	1	1	1	1	1	1	1	1	1	0
	Building	1	0	1	1	1	1	0	1	1	1	1	1	0
	Convention Center	1	0	1	1	1	1	0	1	0	1	1	1	0
	Event Space	1	1	1	1	1	1	1	1	1	1	1	1	1
→	Hotel	29	25	29	29	29	29	27	29	28	29	29	29	8
	Hotel Bar	1	0	1	1	1	1	1	1	1	1	1	1	0
→	Indian Restaurant	2	0	2	2	2	2	2	2	2	2	2	2	0
	Motel	1	0	1	1	1	1	0	1	0	1	1	1	0
	Music	1	0	1	1	1	1	0	1	1	1	1	1	0



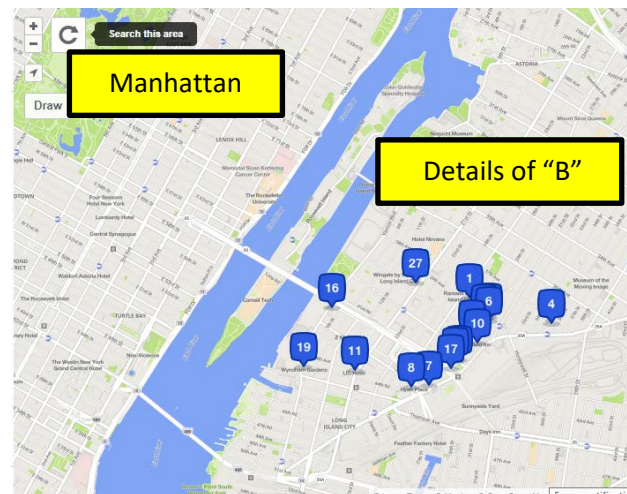
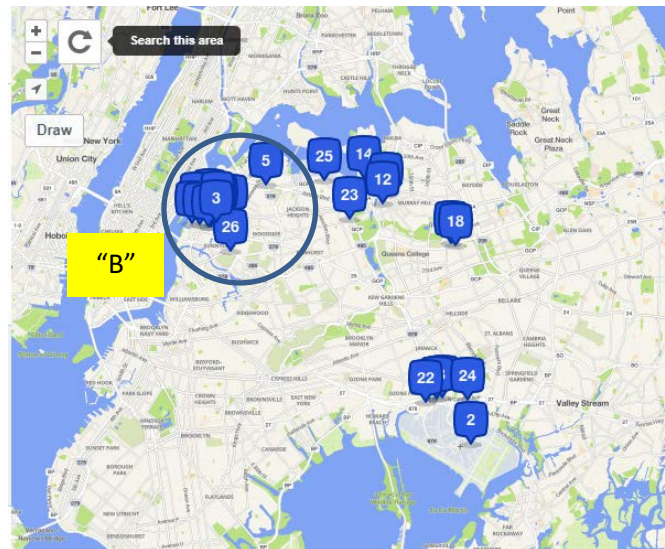
#### 4. Results and Discussion Section

It is interesting to note that, all hotels have been considered to be close to subway stations to facilitate access to other areas. Using the current restaurants dataset I gathered from FourSquare, I generate a trending map for each area. From the maps, we can clearly say that the most hotels at Brooklyn are on the 4<sup>th</sup> Avenue and close to subway stations facing Bay ridge channel.



**Brooklyn District**

While in Queen district, where hotels are more scattered, mainly placed and focused on the area that are close to Manhattan and have access to Manhattan subway as well as the highways.



## Queen District

Obviously, there are more factors we need to take into concern for the decision making. For example, price, facilities, safety and etc. For this project, we just simply bring in one set of data to analyze and predicting.

The surprising finding so far is it seems hotels close to subway stations and facing oceans are more popular at these areas. Does this mean there are more hotels or hostels likely to be accepted and liked by people? If we expand the area, will it be still the same results?

## 5. Conclusion:

To sum up, after comparing the neighborhoods along with current hotels distributions and categories, it apparent that 2 best locations to build new hotels is in Brooklyn on 4ht Avenue and in Queen close to Manhattan district.