

MAPPING THE HOTELS IN NEW YORK CITY

1. INTRODUCTION

1.A. BUSINESS ISSUES

In this project I wanted to have an overview of the existing hotels in New York City. The number of hotels, their location and their repartition through the different boroughs of New York will be presented in this report.

1.B. STACKEOLDERS

This information will give an overview to future investors that may want to study the hotel market in the city of New York.

In this report the following questions will be answered:

- How many hotels exists in the city of New York?
- Where are located these hotels?
- What is the repartition through the different boroughs (Manhattan, The Bronx, Queens, Brooklyn and Staten Island)?

2. DATA

2.1. DATA SOURCES

The data that has been used to build this report was provided by Coursera in a JSON format that anyone can download in the following link: https://cocl.us/new_york_dataset.

The JSON file contained a variety of information such as the different boroughs and neighborhoods of New York City and its correspondent coordinates: latitude and longitude.

The JSON file was transformed into a readable data frame using the pandas library. The data frame contained 306 references. Here below an extract of the first 5 rows:

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

This information was completed with the Foursquare API (<https://foursquare.com/>) in order to know the venues of each neighborhood.

The Foursquare API extended the former data frame with the venue name, its coordinates and its category (pharmacy, shop, hotel, etc.). Here below an extract of the extended data frame with some venues of the neighborhood of Wakefield in The Bronx:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	Lollipops Gelato	40.894123	-73.845892	Dessert Shop
1	Wakefield	40.894705	-73.847201	Carvel Ice Cream	40.890487	-73.848568	Ice Cream Shop
2	Wakefield	40.894705	-73.847201	Walgreens	40.896528	-73.844700	Pharmacy
3	Wakefield	40.894705	-73.847201	Rite Aid	40.896649	-73.844846	Pharmacy
4	Wakefield	40.894705	-73.847201	Dunkin'	40.890459	-73.849089	Donut Shop

2.2. DATA TREATMENT

The data didn't need any special cleaning as there were not a single empty cell. The JSON file was already in good conditions and the foursquare API did not interfere in the good shape of the data frame.

The only modification of the original data frame was a split in several data frames to have an individual analysis of the five different boroughs.

3. METHODOLOGY

3.1. PROCESS AND INITIAL OUTPUTS

Once the data frame was ready, the first step consisted on split the original data frame into five different data frames, each one corresponding to a borough (Manhattan, The Bronx, Queens, Brooklyn and Staten Island).

The next step was to connect to the Foursquare API and extract all the venues from the neighborhoods, for each borough data frame. The references were multiple in each data frame and this is one of the reasons why I decided to split the original data frame into five distinct data frames.

As an example, the rows in the data frames for the boroughs were as follows (each row corresponded to single venue that could be a pharmacy, a shop, a hotel, a restaurant, etc.):

BOROUGHS DATA FRAME	NUMBER OF ROWS
The Bronx	1 114
Manhattan	1 869
Brooklyn	2 151
Queens	1 864
Staten Island	836

Once the data frame with the venues for each borough were extracted then a filter was passed to retrieve only the hotels. The results are in the table here below:

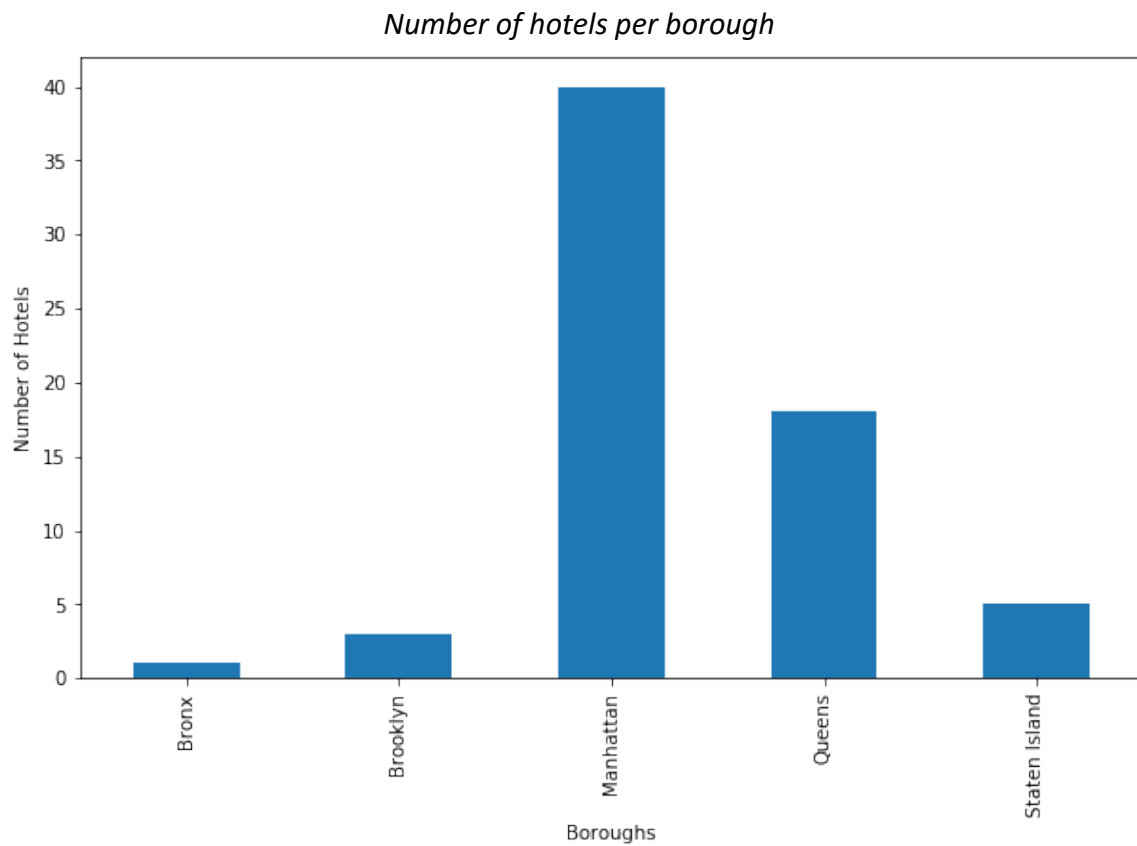
BOROUGHS DATA FRAME	NUMBER OF HOTELS
The Bronx	1
Manhattan	40
Brooklyn	3
Queens	18
Staten Island	5

Finally, I displayed a bar chart showing the neighborhoods that hosted the hotels, a bar chart for the total amount of hotels per borough and a map to show the locations of the hotels.

3.2. ANALYSIS AND DEFINITIVE OUTPUTS

As shown in the previous section Manhattan has the biggest number of hotels and The Bronx has the minimum with only 1 hotel.

The repartition of the hotels is very unregular with two main boroughs that had a leader position and 3 boroughs which number of hotels vary between 1 and 5.



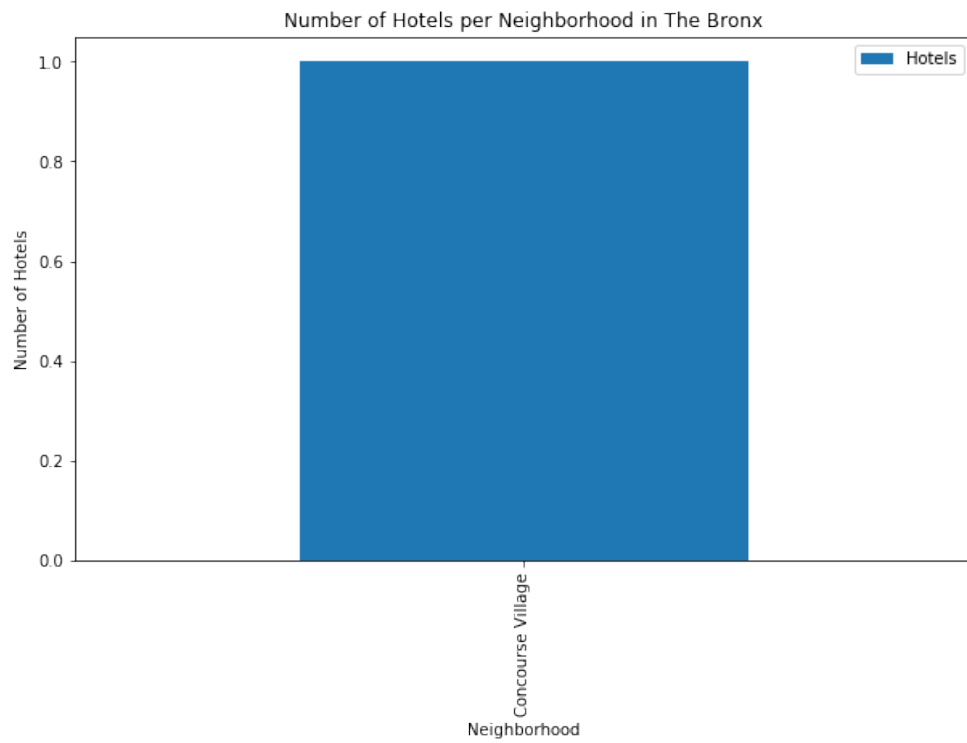
4. RESULTS

4.1. STATISTICAL REPARTITION

The repartition per neighborhood is as follows:

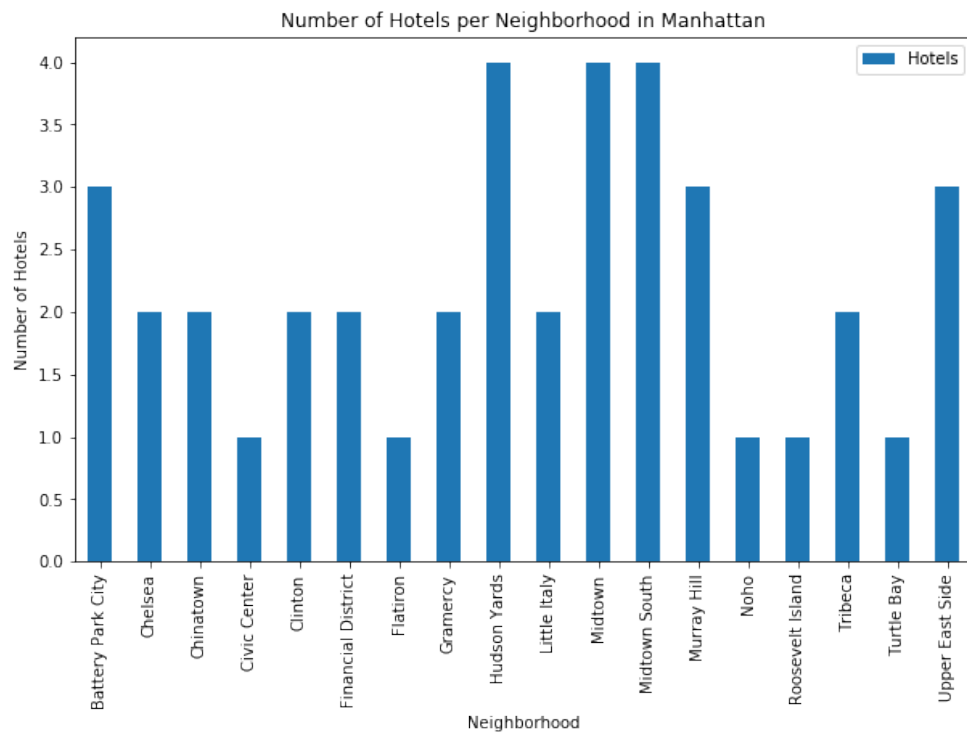
The Bronx:

Only one hotel in the neighborhood of Concourse Village.



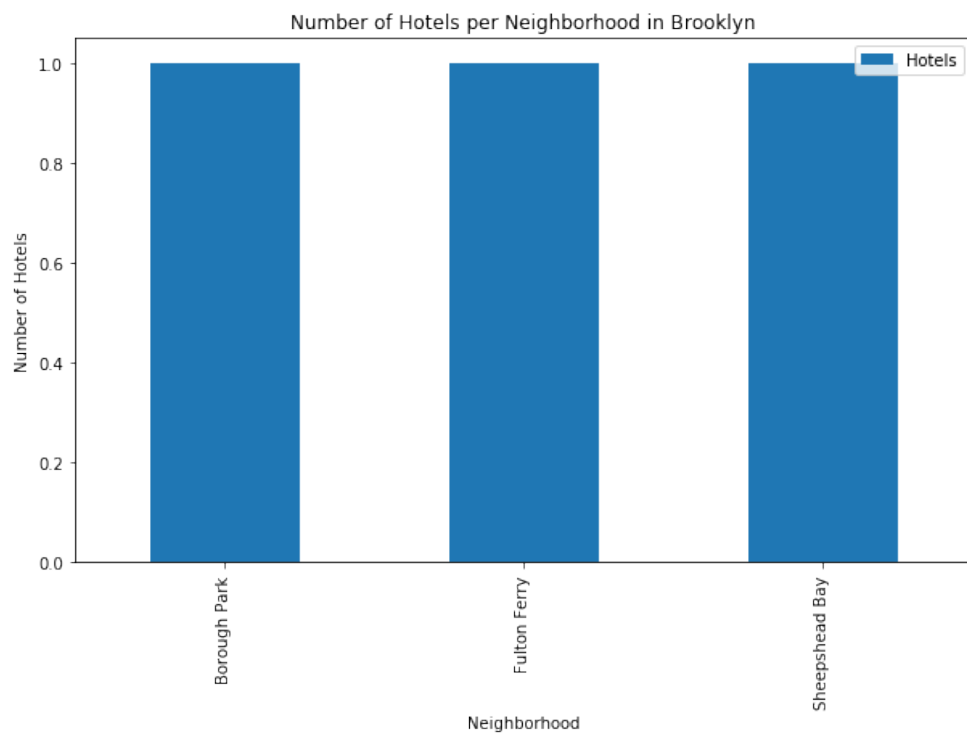
Manhattan:

The most popular borough with 40 hotels mainly based in Hudson Yards, Midtown and Midtown South.



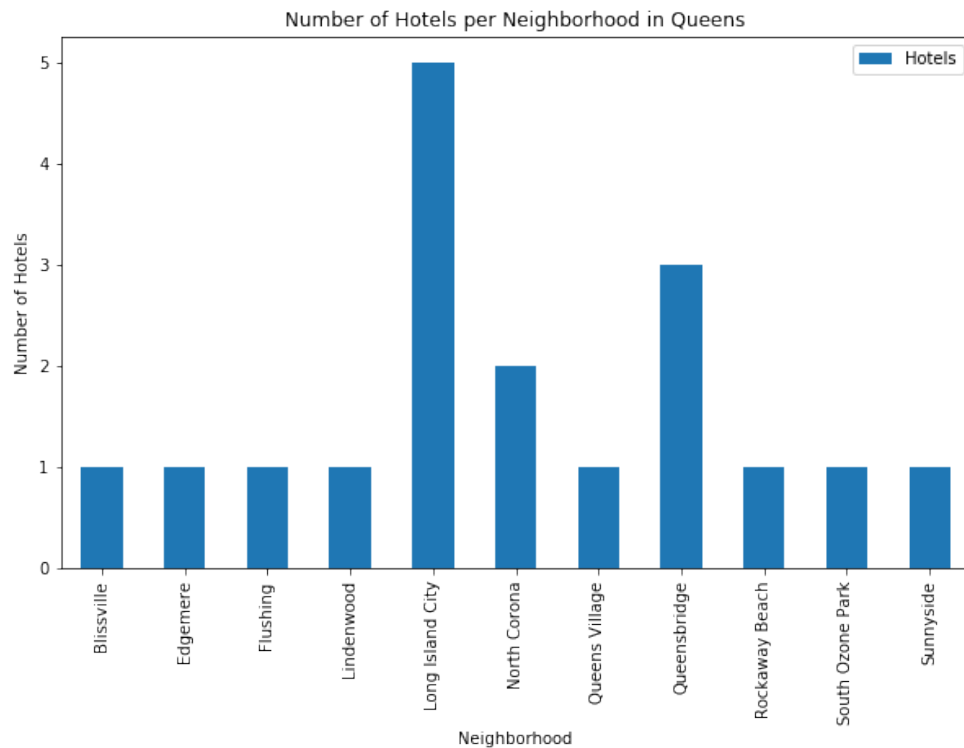
Brooklyn:

3 hotels in the neighborhoods of Borough Park, Fulton Ferry and Sheepshead Bay.



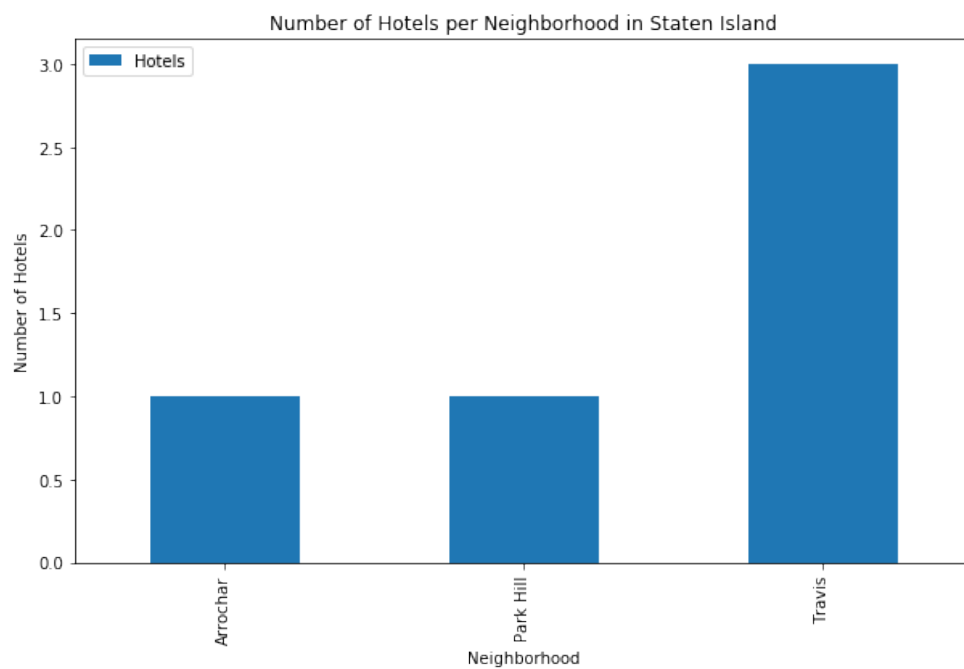
Queens:

The second most popular place for hotels with a total amount of 18. 25% of them are located in the neighborhood of Long Island City.



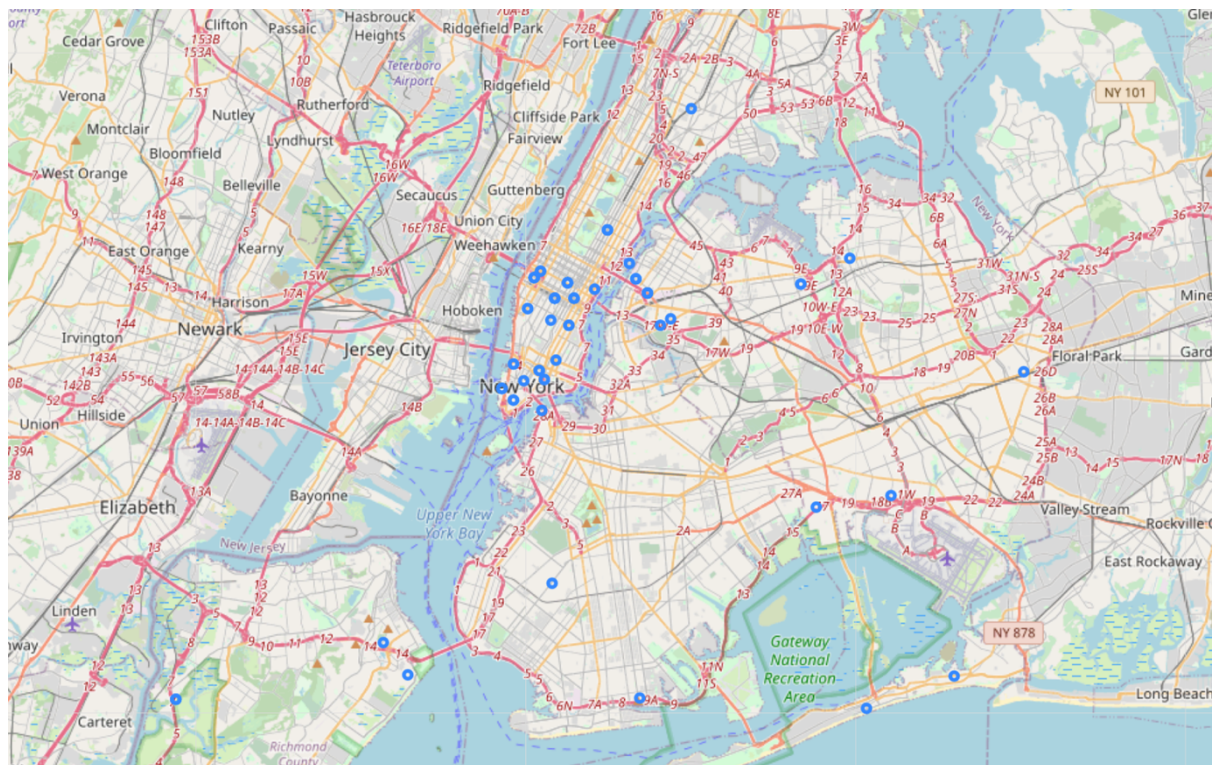
Staten Island:

5 hotels in this borough, most of them located in Travis neighborhood.



4.2. GEOGRAPHICAL REPARTITION

Geographical repartition of the hotels in New York City



As we can see in this map, most of the hotels are located in the Manhattan Island where they are very concentrated.

The hotels that are not in Manhattan seem to be very disperse around New York City.

5. DISCUSSION AND CONCLUSION

In this report you can see the repartition of the hotels around the city of New York. We can easily conclude that most of the hotels are located in Manhattan which seems to be the most centric part.

This report gives thus an overview to future investors that may want to have study market information. Since other boroughs than Manhattan or Queens seem to have very few hotels, an investment in boroughs that are not Manhattan or Queens are maybe riskier as they are not hotel areas.