

# Capstone Project

## Airbnb Exploratory Data Analysis

Submitted By: Karan Tiwari

# Introduction

**Since 2008, Airbnb has helped travelers and host a lot to find new places to go and offer a more distinctive, customized way to see the globe. Today, Airbnb is a unique service that the entire globe uses and recognizes. For the business, data analysis on the millions of listings offered through Airbnb is essential. These listings generate billions of data, which can be analyzed and used for a variety of purposes, including security, business decisions, understanding customer and host behavior and much more.**

# Steps Involved

- **1 Importing the dataset and libraries.**
- **2 Getting an overview of the dataset.**
- **3 Cleaning the data set and getting rid of null values and outliers.**
- **4 Data Analysis**
- **5 Data Visualization**
- **6 Conclusions**

# Purpose Of Analysis

- **The purpose of this study is to estimate demand for Airbnbs based on factors such as location, pricing, and so on, and then produce projections based on that data. While some neighborhoods retain a high price throughout the year, others are more costly during specific seasons. I looked at cost, neighborhood, room type, and popularity in the dataset. In my presentation, I delved into further depth regarding the conclusions.**

# Basic information of dataset

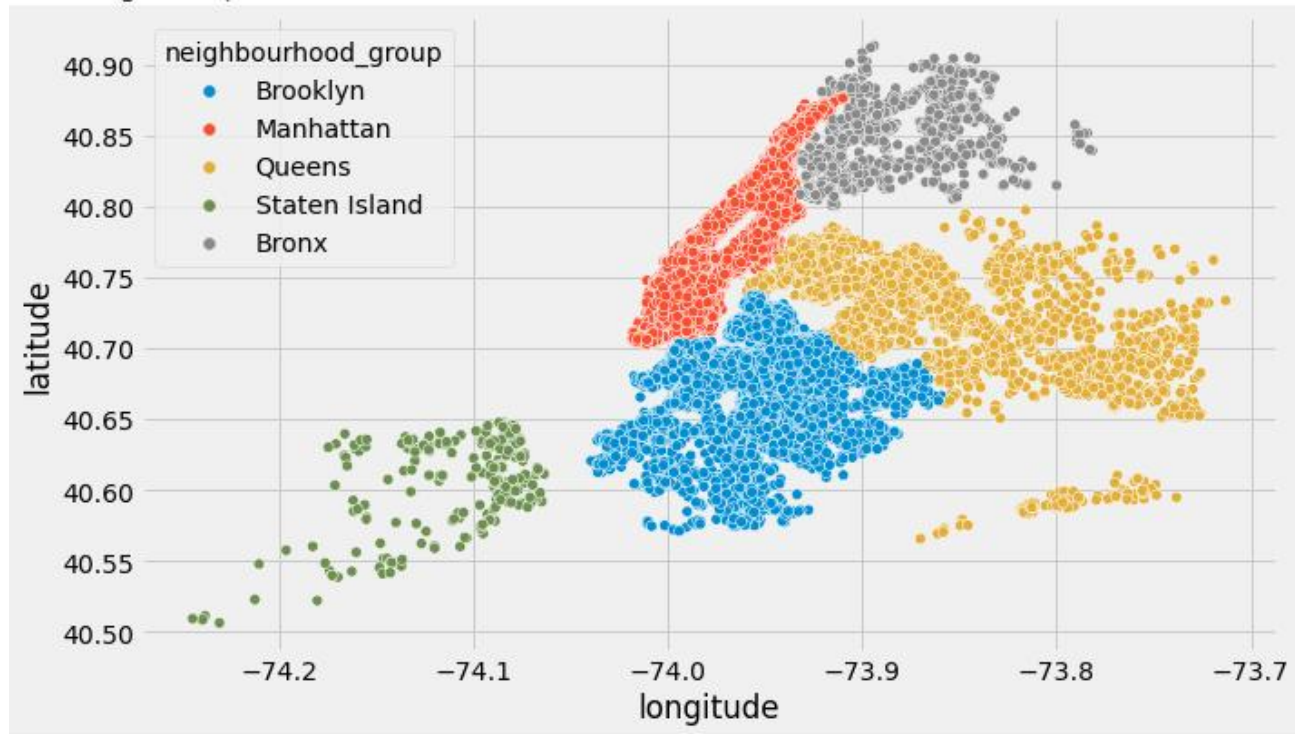
This dataset has around 48,895 rows and 16 columns

```
# Checking null values
df.isna().sum()
```

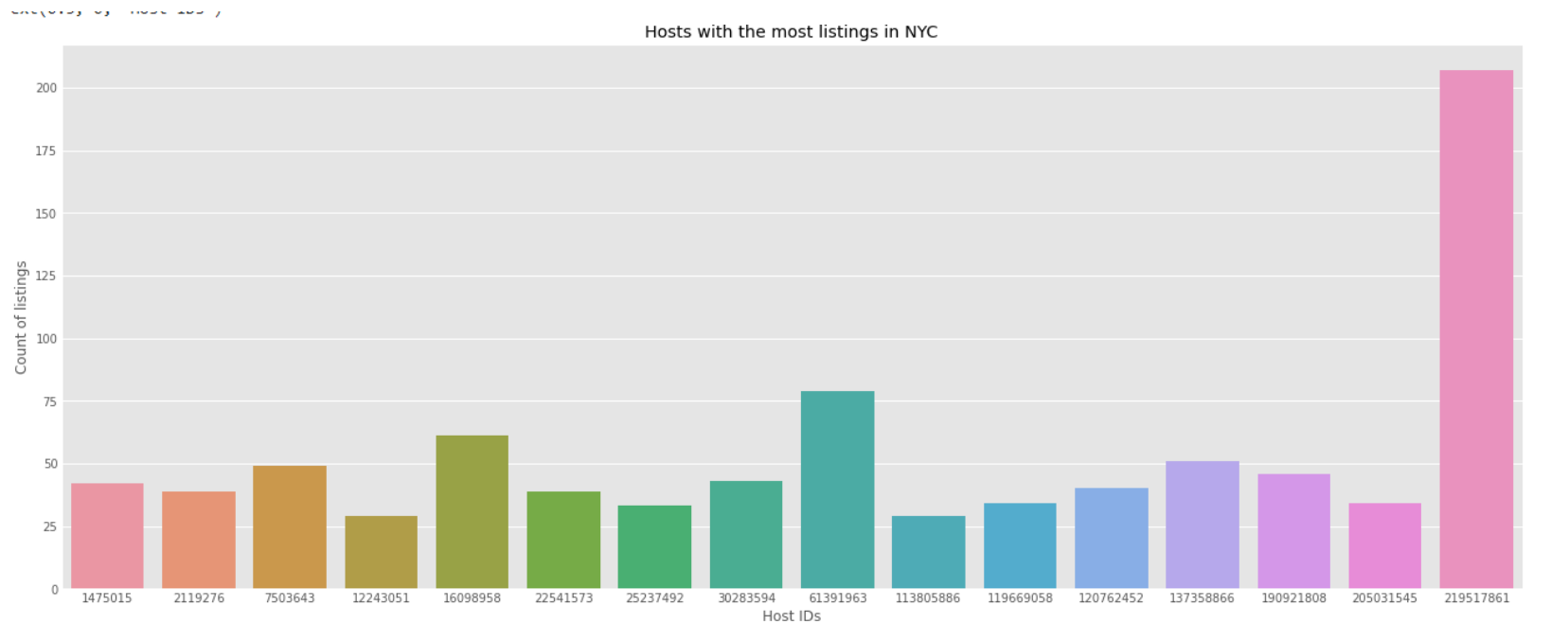
```
id          0
name        16
host_id     0
host_name   21
neighbourhood_group  0
neighbourhood  0
latitude    0
longitude   0
room_type   0
price       0
minimum_nights  0
number_of_reviews  0
last_review 10052
reviews_per_month 10052
calculated_host_listings_count  0
availability_365  0
dtype: int64
```

	id	host_id	latitude	longitude	price	minimum_nights	number_of_reviews	reviews_per_month	calculated_host_listings_count	availability_365
<b>count</b>	3.882100e+04	3.882100e+04	38821.000000	38821.000000	38821.000000	38821.000000	38821.000000	38821.000000	38821.000000	38821.000000
<b>mean</b>	1.810081e+07	6.424582e+07	40.728129	-73.951149	142.332526	5.869220	29.290255	1.373229	5.166611	114.886299
<b>std</b>	1.069372e+07	7.589752e+07	0.054991	0.046693	196.994756	17.389026	48.182900	1.680328	26.302954	129.529950
<b>min</b>	2.539000e+03	2.438000e+03	40.506410	-74.244420	0.000000	1.000000	1.000000	0.010000	1.000000	0.000000
<b>25%</b>	8.721444e+06	7.029525e+06	40.688640	-73.982460	69.000000	1.000000	3.000000	0.190000	1.000000	0.000000
<b>50%</b>	1.887286e+07	2.837092e+07	40.721710	-73.954810	101.000000	2.000000	9.000000	0.720000	1.000000	55.000000
<b>75%</b>	2.756746e+07	1.018905e+08	40.762990	-73.935020	170.000000	4.000000	33.000000	2.020000	2.000000	229.000000
<b>max</b>	3.645581e+07	2.738417e+08	40.913060	-73.712990	1000.000000	1250.000000	629.000000	58.500000	327.000000	365.000000

## Scatter plot of airbnbs in different neighborhoods on map

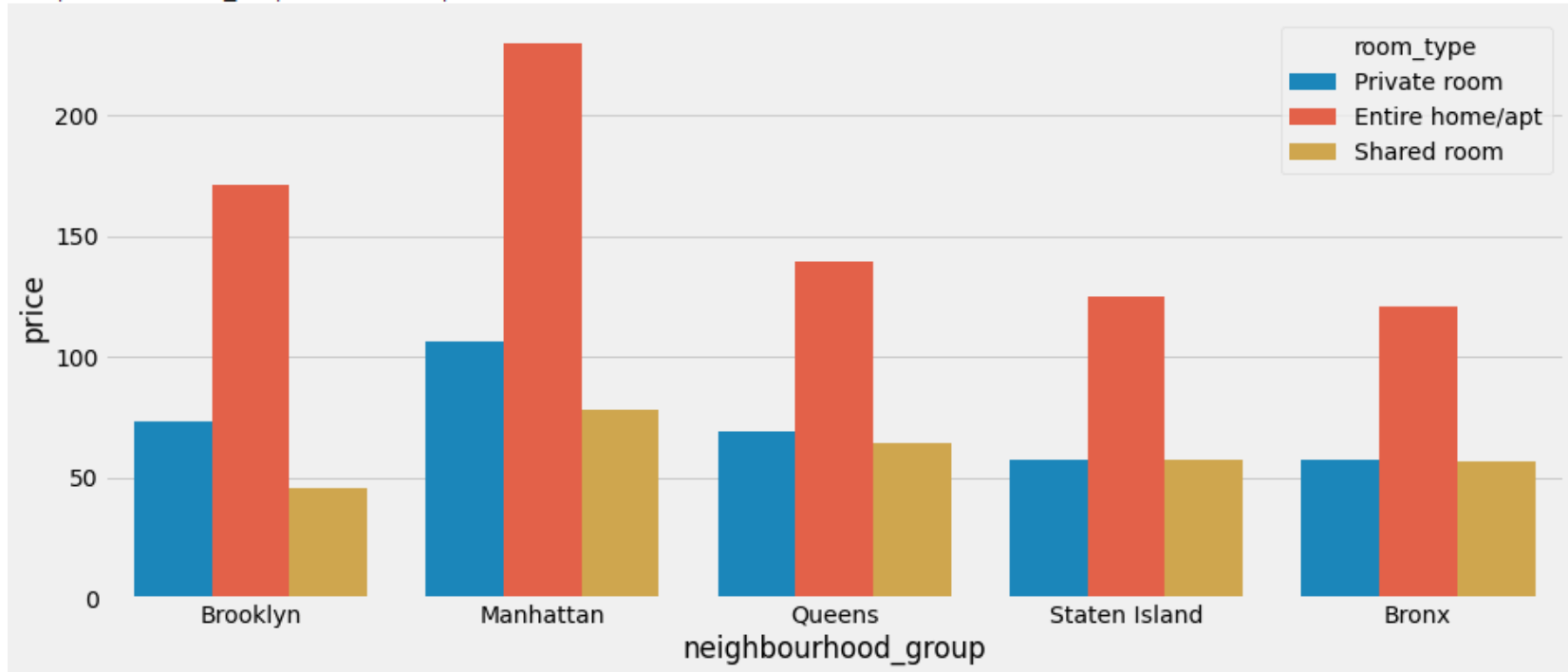


# Top Host's listing



**219517861 has most listing in NYC.**

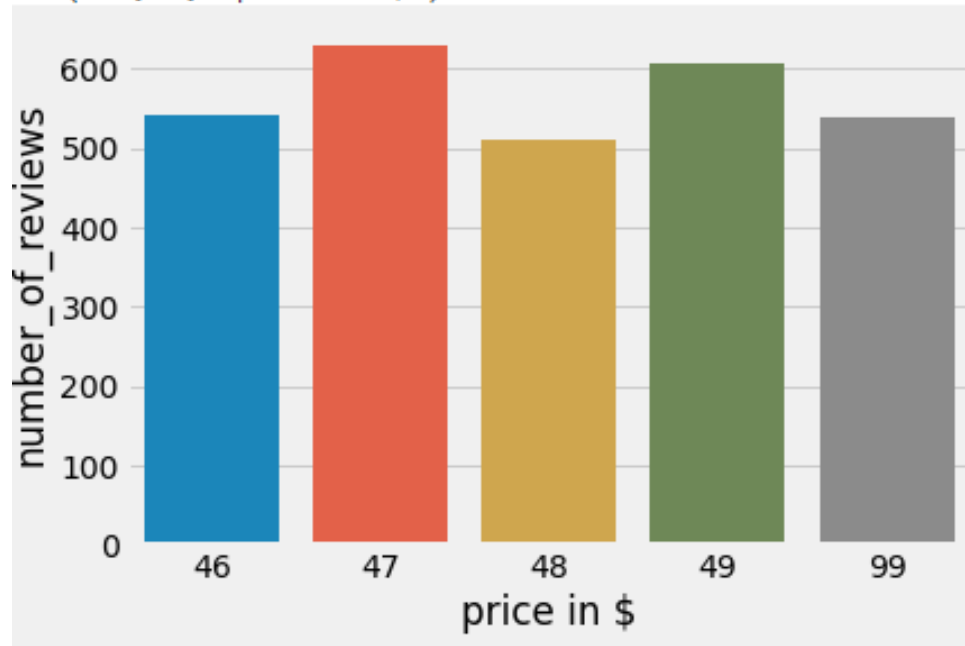
# Average price by room type



**Manhattan has very expensive rooms ,Staten island and Bronx have cheaper rooms.**

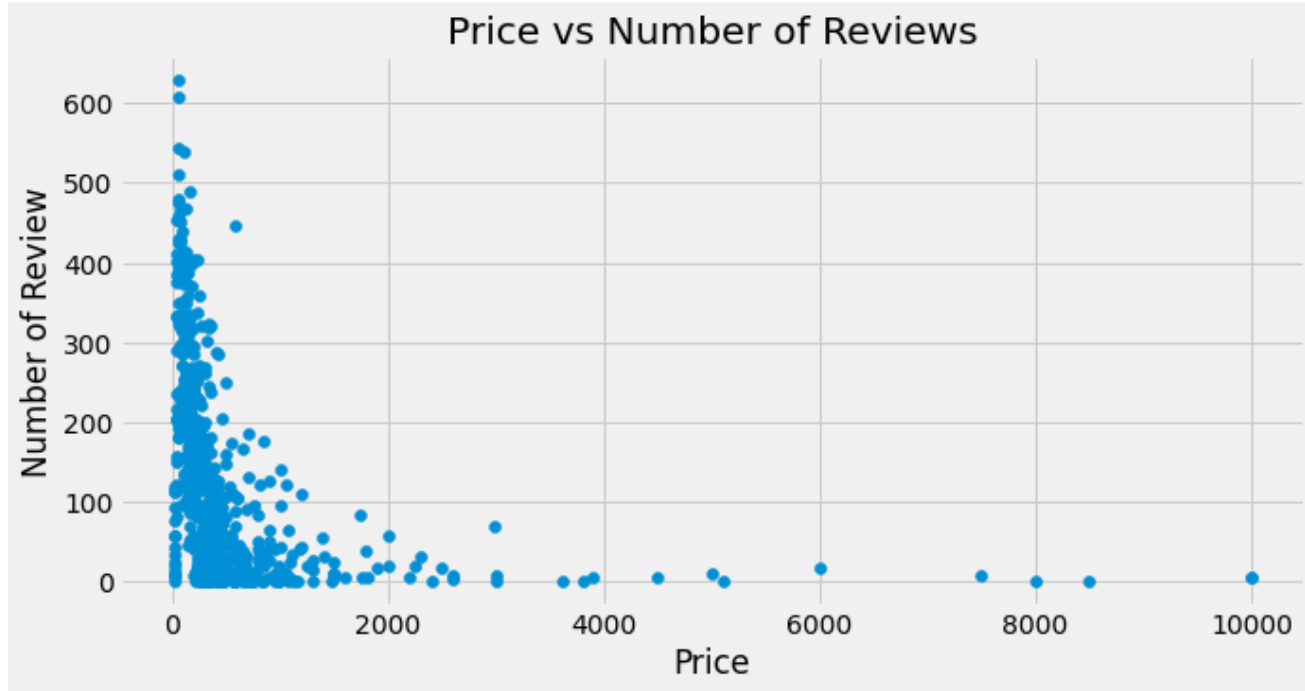


## Top 5 popular price range



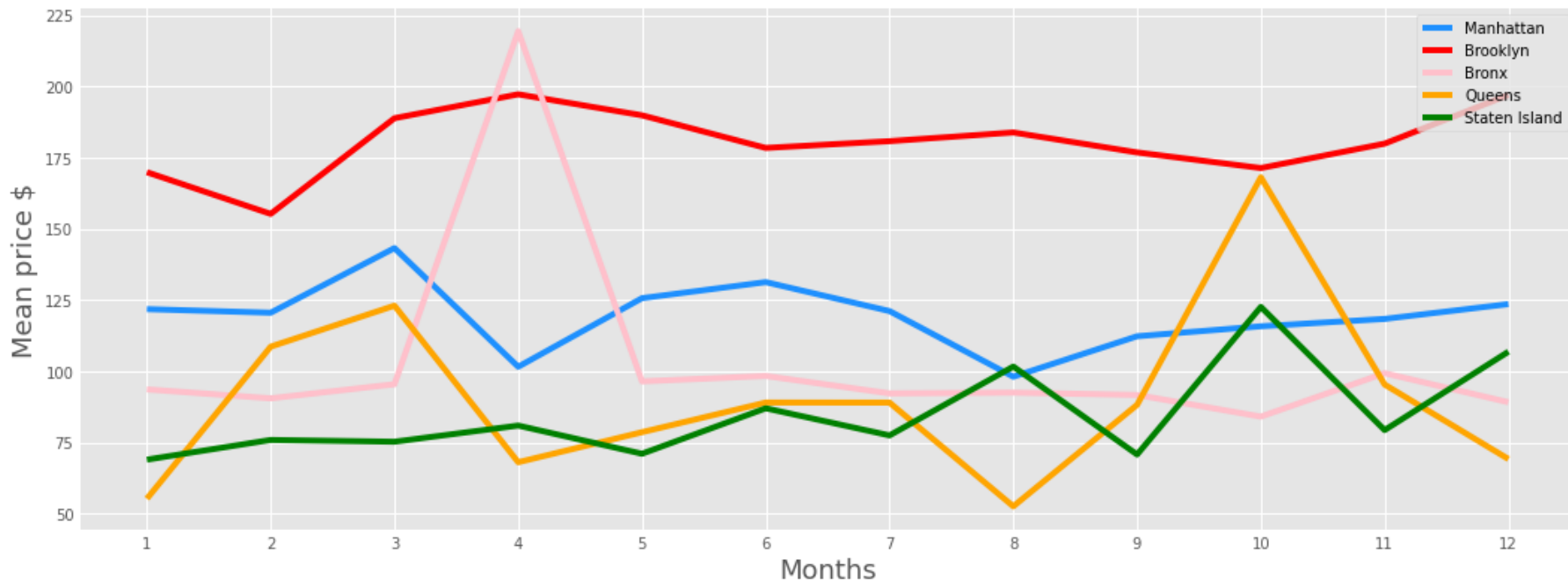
**Customers prefers to stay on those rooms only whose price range is between 40 to 100\$**

# Most frequently booked price range



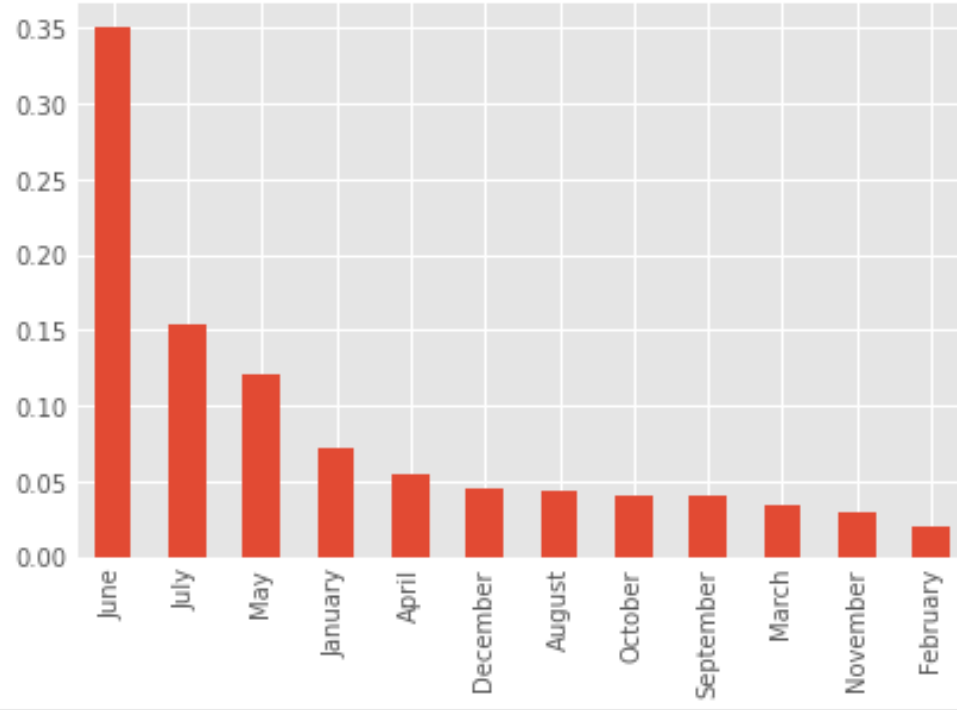
**Airbnbs with price less than 1000\$ are the Customers preference.**

# Mean prices of different neighborhoods through out the year



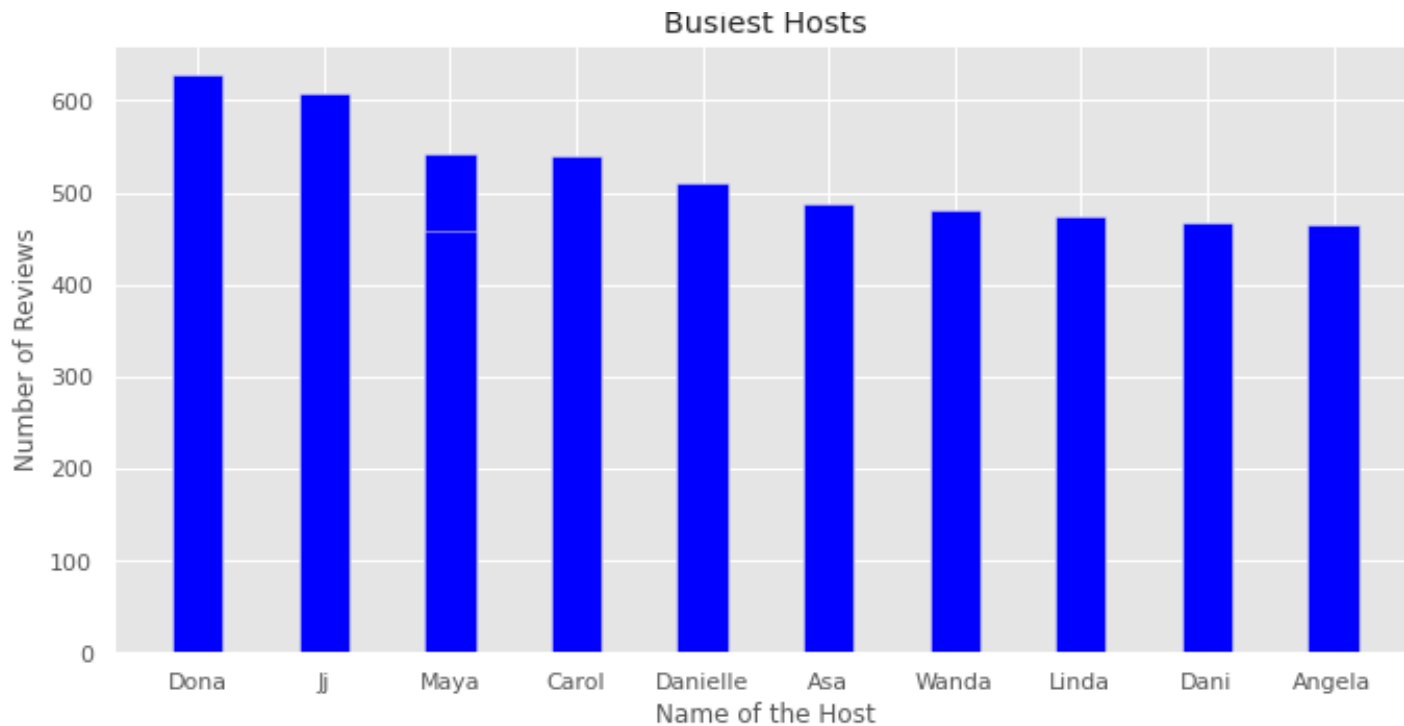
**Brooklyn maintains a high price throughout the year. Bronx is expensive in the month of April and Queens is expensive in the month of October.**

# Busiest month for airbnbs



**June and July are the busiest month, 50% of booking is done in the month of June and July.**

# Busiest Hosts



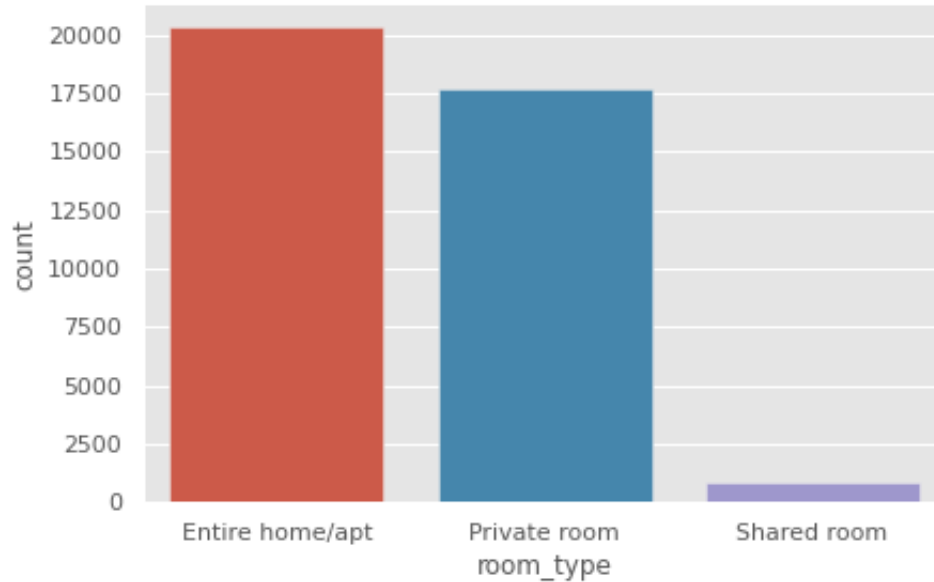
**These are the name of top 10 busiest host in NYC.**

# Insight into busiest hosts

	host_name	host_id	room_type	price	number_of_reviews
9383	Dona	47621202	Private room	47	629
16430	Jj	4734398	Private room	49	607
23759	Maya	37312959	Private room	46	543
5765	Carol	2369681	Private room	99	540
8179	Danielle	26432133	Private room	48	510
3650	Asa	12949460	Entire home/apt	160	488
35055	Wanda	792159	Private room	60	480
20953	Linda	2680820	Private room	55	474
7919	Dani	42273	Entire home/apt	120	467
2671	Angela	23591164	Private room	65	466
23758	Maya	37312959	Private room	45	459

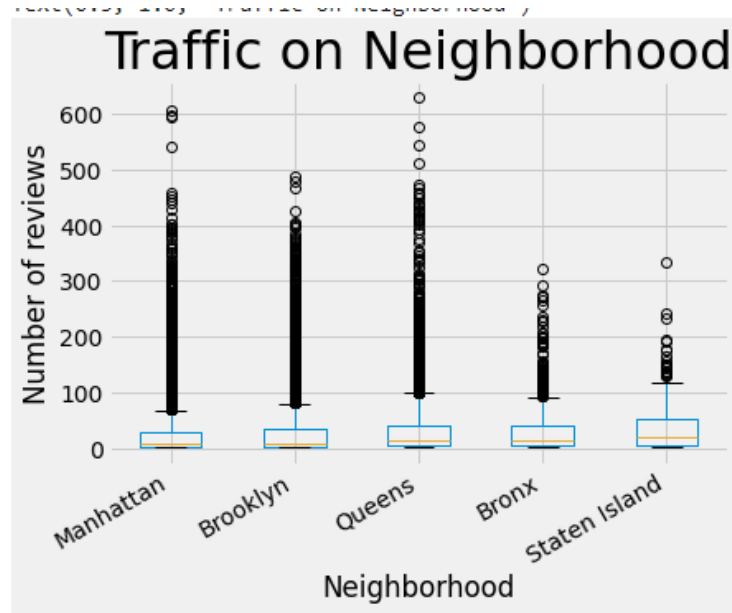
**On analyzing the table we can notice some similarities. Private rooms are booked most in Price range between 45-100 \$. Therefore we can conclude people prefer private rooms in price range 45-100\$.**

# People preference on room type



**Customers are mostly preferring Home/Apartment and private rooms ,only few customers prefers to stay in shared rooms.**

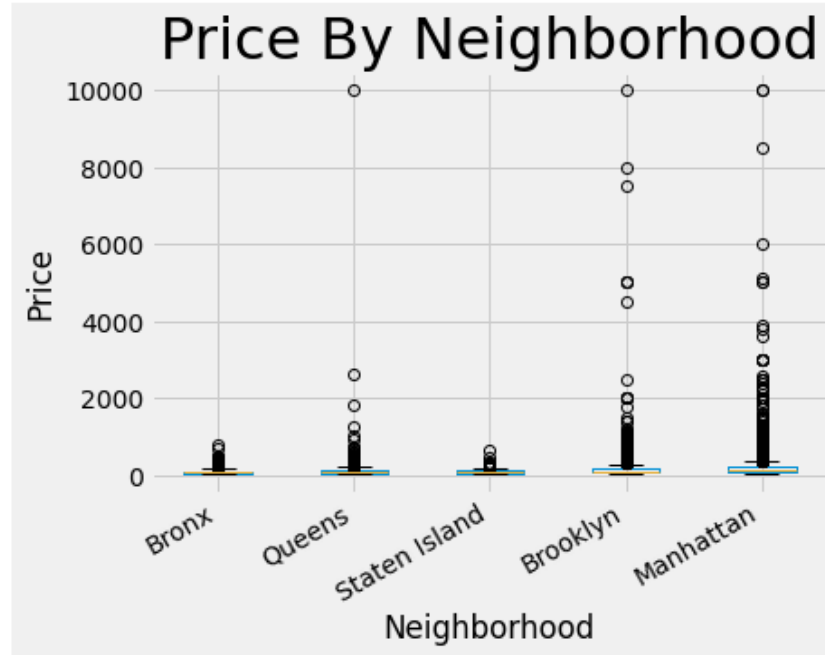
# Traffic areas



**Manhattan, Queens and Brooklyn have maximum traffic.**



# Price by Neighborhood



**Brooklyn and Manhattan are the most expensive places.**

# Conclusion

- **1. People prefer to stay in cheaper accommodations.**
- **2. If you are looking for cheaper stays then it is recommended to not travel to Bronx in the month of April and Queens in the month of October.**
- **3. June and July are the busiest months.**
- **4. Home apartment and Private rooms are most preferred room type**
- **5. Brooklyn and Manhattan experiences highest traffic. This is because they are tourist place. The average price in also Brooklyn and Manhattan are higher as compared to other neighborhoods.**