Capstone Project
NYC Airbnb Booking Analysis

SARANYA N

CONTENTS

- **Problem Statement**
- **Data Summary**
- **Doing Basic Examination**
- **Cleaning The Dataset**
- **What Hosts And Areas Tell U**
- **What Can We learn from predictions , price , locations , reviews,listings**
- **Distribution lists with availability**
- **The Busiest Host**
- **Facts About Traffic**
- **Conclusion**

PROBLEM STATEMENT

Airbnb, Inc. is an American vacation rental online marketplace company based in San Francisco, United States. The company has huge data of guests and hosts (Guesthouse/Hotel Owners) of New York City. This data can be analyzed and used to discover understandings such as:

- **What can we learn about different hosts and areas?**
- **What can we learn from predictions? (ex: locations, prices, reviews, etc)**
- **Which hosts are the busiest and why?**
- **Is there any noticeable difference of traffic among different areas and what could be the reason for it?**

DATA SUMMERY

let us go over the columns that we are going to need to solve the problem.
- **host_id**
- **host_name**
- **neighbourhood_group**
- **neighbourhood**
- **longitude**
- **latitude**
- **room_type**
- **price**
- **minimum_nights**
- **reviews_per_month**
- **calculated_host_listings_count**

EXAMINATION

- **The dataset contains 48895 rows and 16 columns**

- **Datatypes of various columns are integer, float and objects**

- **Null values removeing**
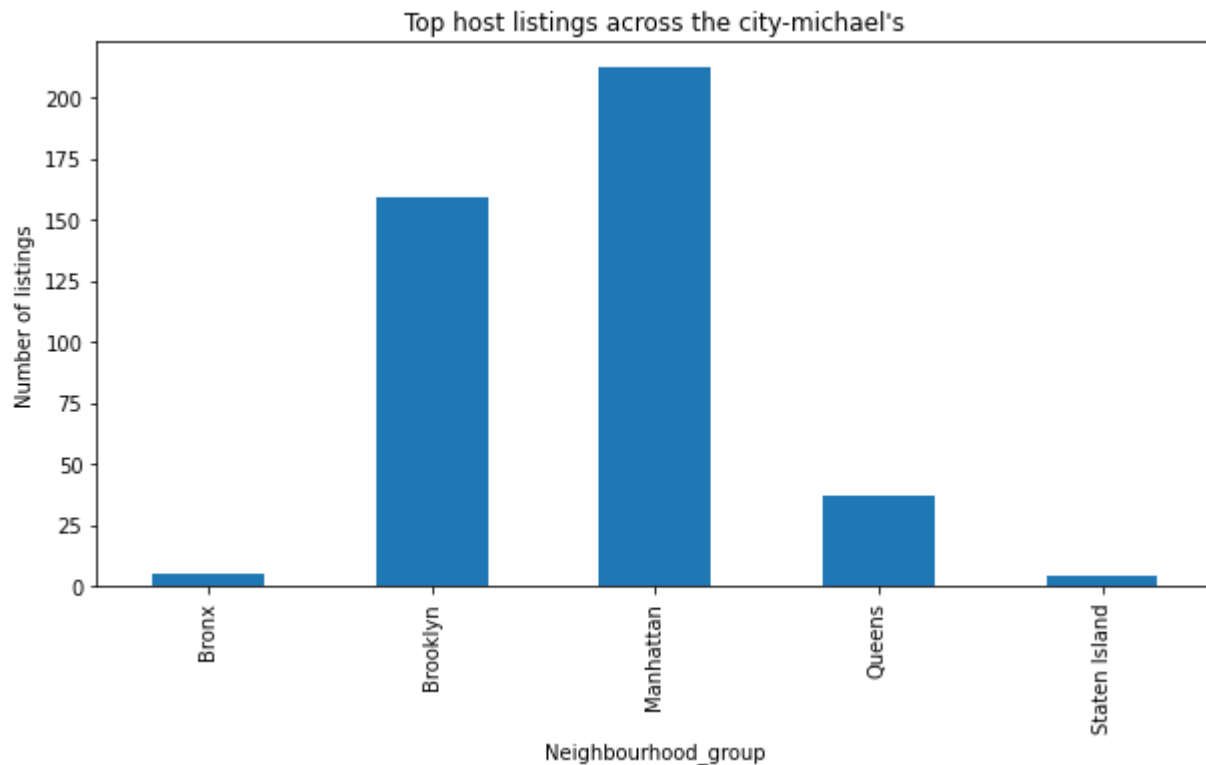
- **Data cleaning steps required**

CLEANING THE DATA

- **Checking null values**

- **Duplicate check**
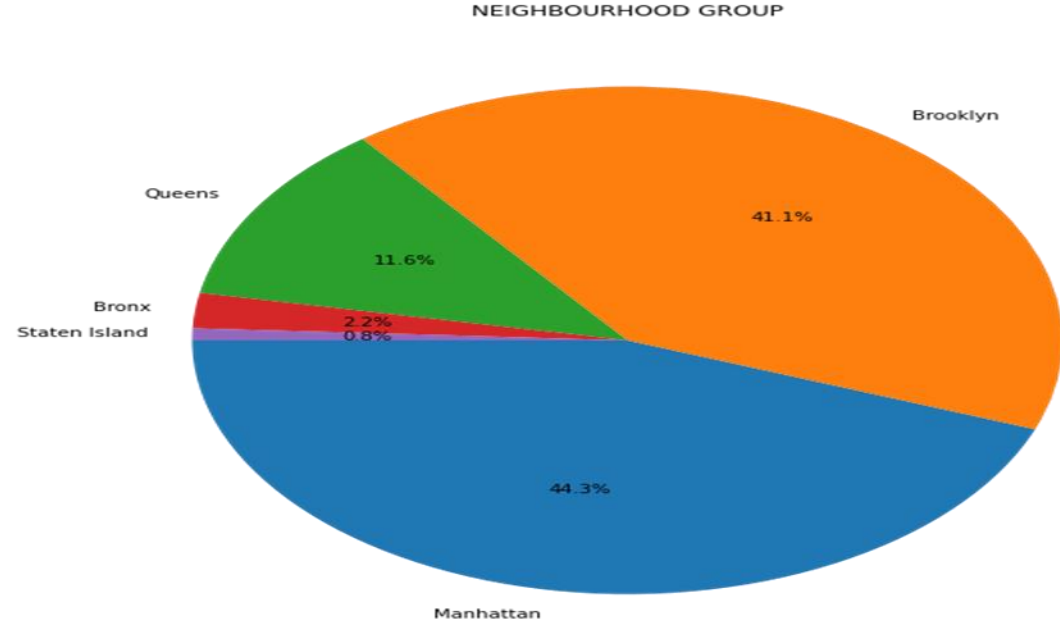
- **Make data ready for analysis**

Lets check the most no of listings in NYC .

- **FIRST WE FIND TOP 10 HOSTS MAXIMUM LISTING**

- **Then where the top host has his listings throughout the city.**

- **Plot the graph with number of listings and neighbourhood groups**
- **From the data we undestood that the top host(michaels) most of the properties located in manhatten**

Host listing throughout the city

Top host listings across the city-michael's

NEIGHBOURHOOD GROUP

Brooklyn 41.1%

Queens 11.6%

Bronx 2.2%

Staten Island 0.8%

Manhattan 44.3%

From the above pie chart we can understand that airbnb listing are near Manhattan,and brooklyn has the highest share of hotels

count of each room types in entire NYC
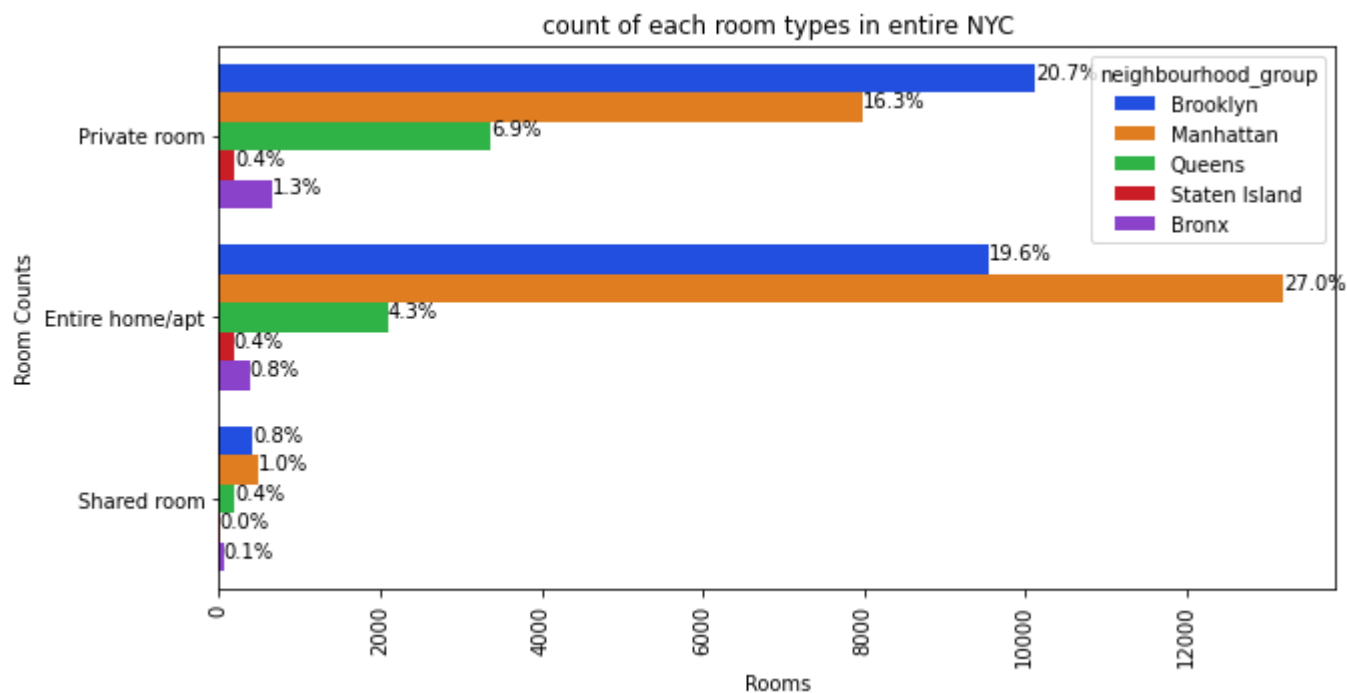
From the above graph we observe that Manhattan has more listed properties with entire home/apt around 27% of total listed properties followed by bookklyn with around 19.6%

in case of private room , in booklyn its 20.7% of the listed properties followed by manhattan with 16.3% . while 6.9% of private rooms are from Queens

Only a few of the total listed have shared rooms listed on Airbnb where there's rare or negligible shared rooms in Staten Island and Bronx.

from the above findings , we understand that Brooklyn , Queens,Bronx has more private room types while Manhattan has the highest no of listings in entire NYC (ie more entire home/apt room types)
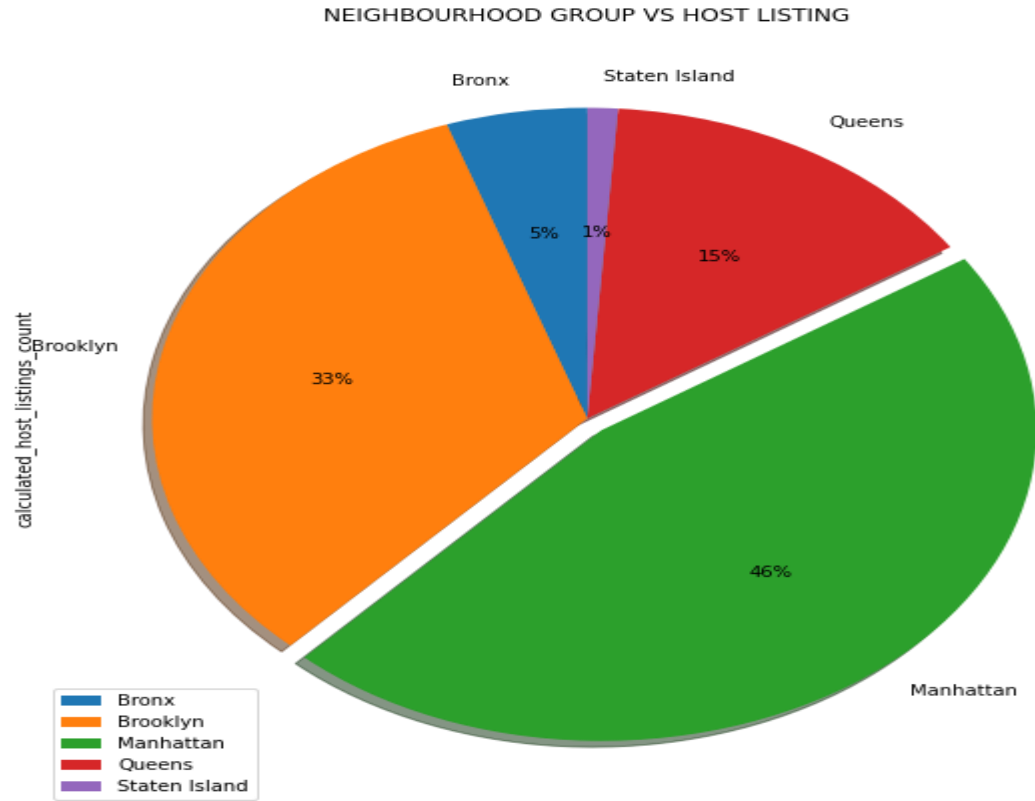
# what can we learn from predictions ? locations , prices, reviews, etc...

AREA VS NUMBER OF REVIEWS

PRICE VS NUMBER OF REVIEWS
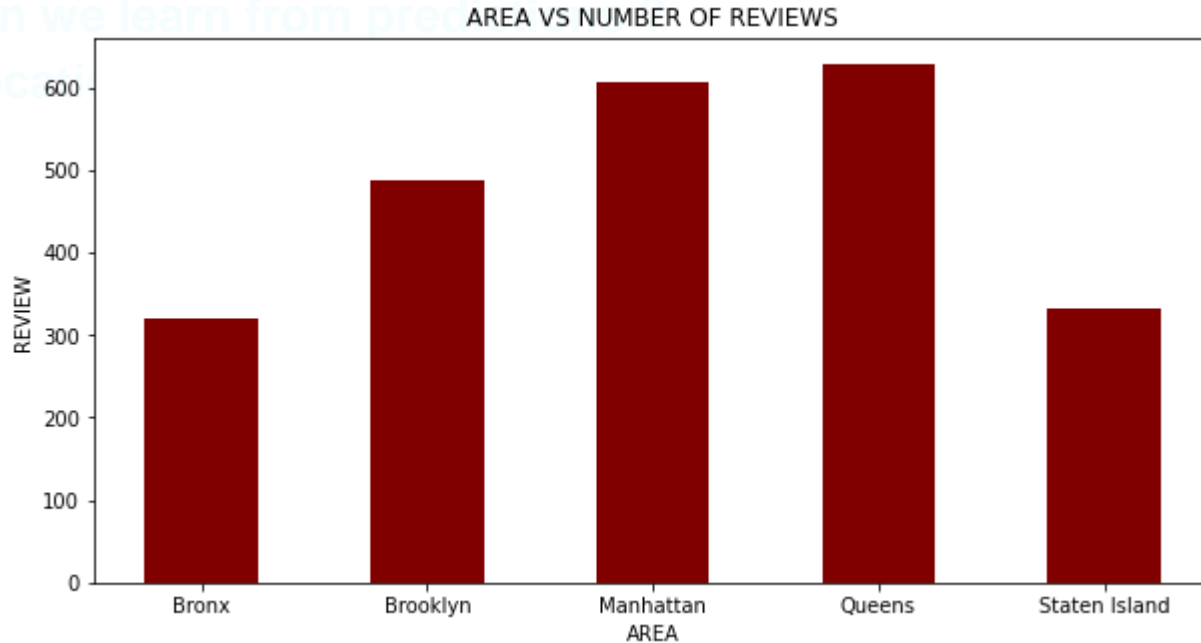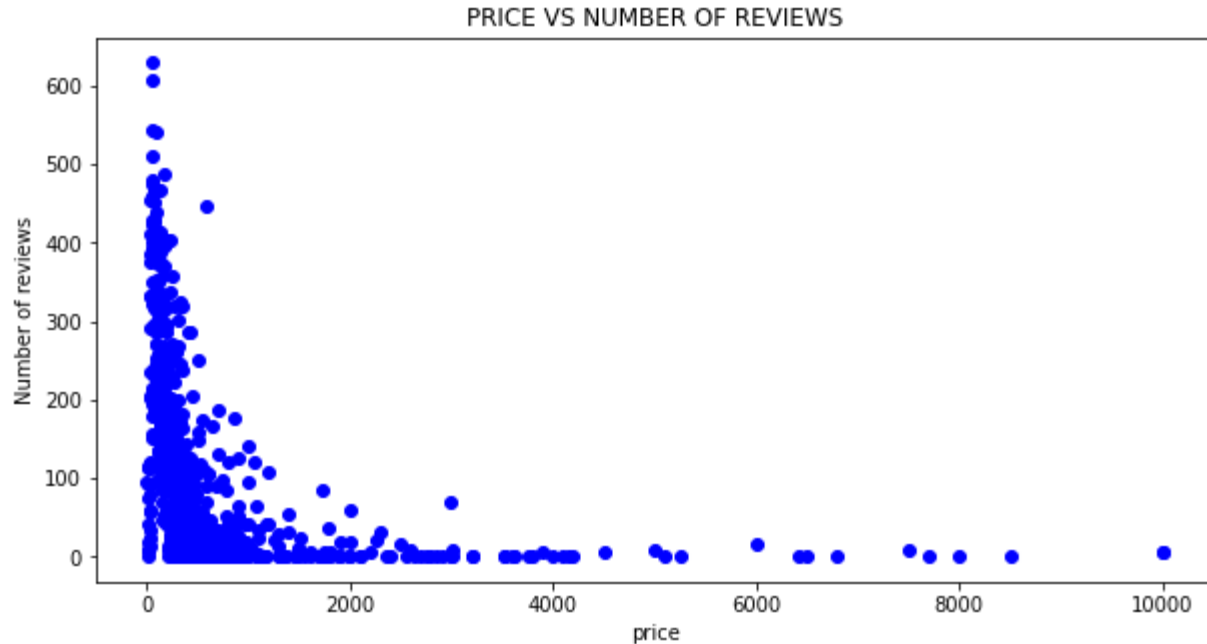
From the above analysis we can say that most people prefer to stay in place where price is less

NEIGHBOURHOOD GROUP VS HOST LISTING

Bronx — 5%
Staten Island — 1%
Queens — 15%
Brooklyn — 33%
Manhattan — 46%

calculated_host_listings_count

Legend:
- Bronx
- Brooklyn
- Manhattan
- Queens
- Staten Island

from this pie chart we can understand that the most number of listings are from *manhattan*

lets look at the distribution of the listings with availability.

NEIGHBOURHOOD GROUP VS ROOM AVAILABILITY
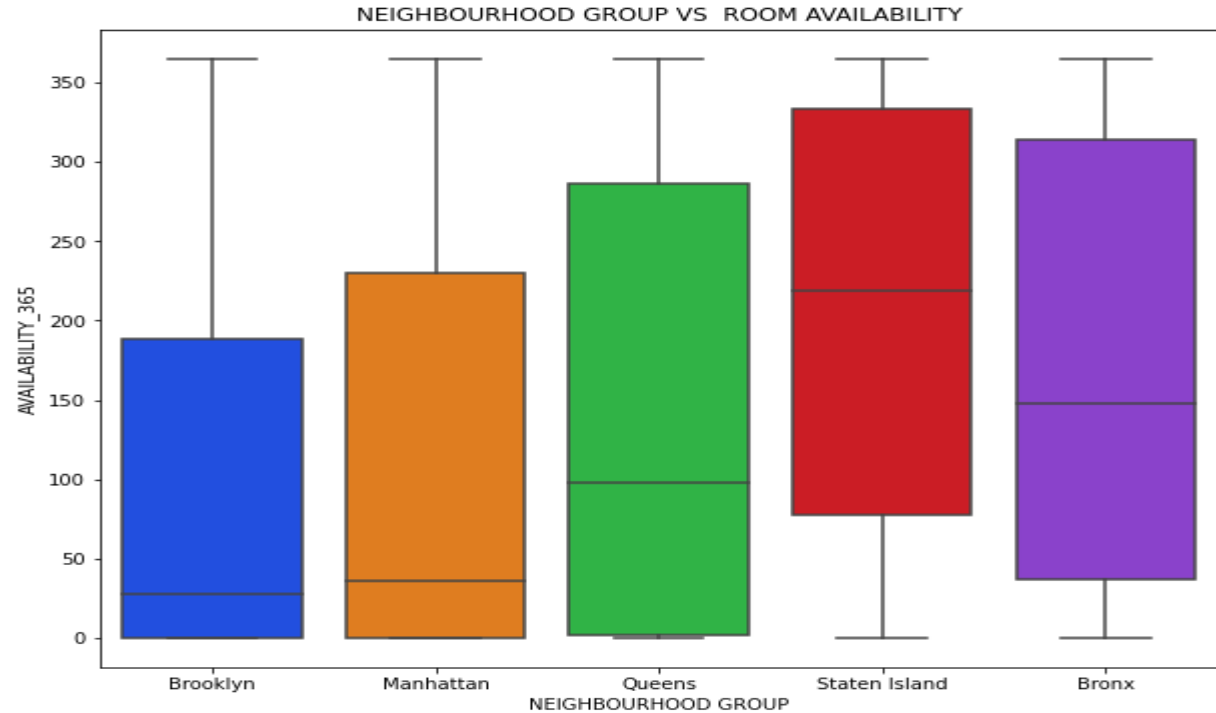
FROM THE ABOVE GRAPH WE CAN FIND THAT STATEN ISLAND SEEMS TO BE MORE AVAILABLE THROUGHOUT THE YEAR IT IS AVAILABLE MORE THAN 300 DAYS IN A YEAR.
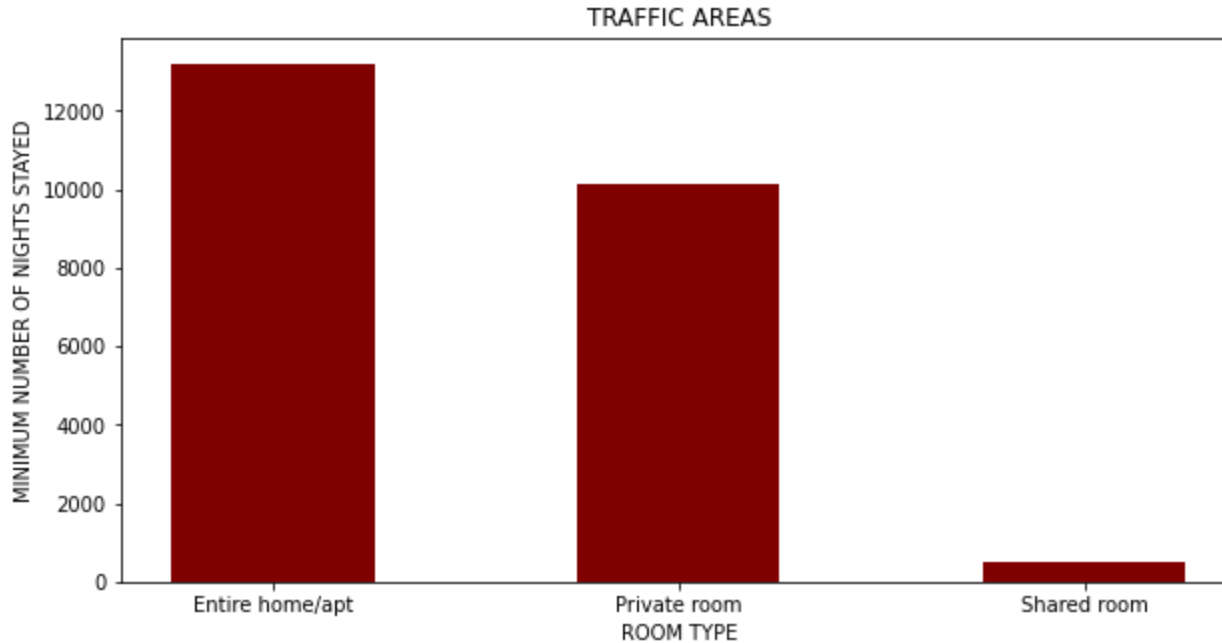
WHICH HOSTS ARE THE BUSIEST

BUSIEST HOSTS

-
- FROM THE ABOVE ANA
  BUSIEST HOSTS AR D
- D
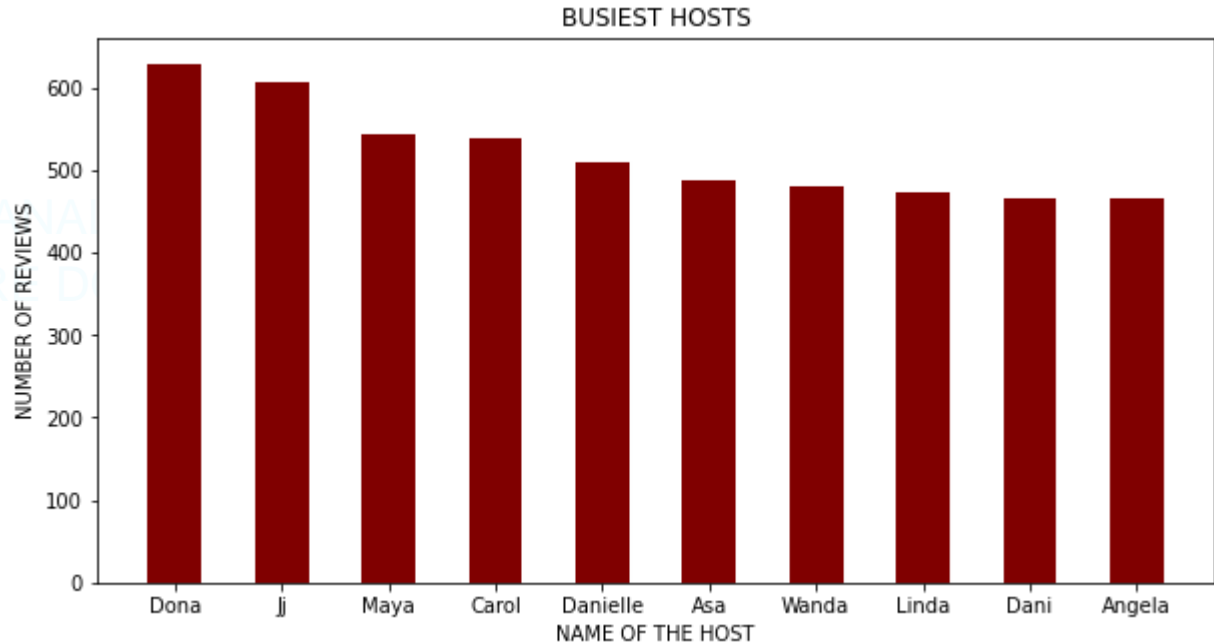
DONA is the most reviewed host with more than 600 reviews .s
o we can undersatand that dona is the most busiest host .

Is there any noticeable difference of traffic among different areas and what could be the reason for it

from this graph we can analisys that most people prefer entire home/apt and then private room which are in manhattan ,queens ,brooklyn

now just find how many nights people stayed in each neighbourhood group

Min Stays in different neighbourhood_group listed on Airbnb
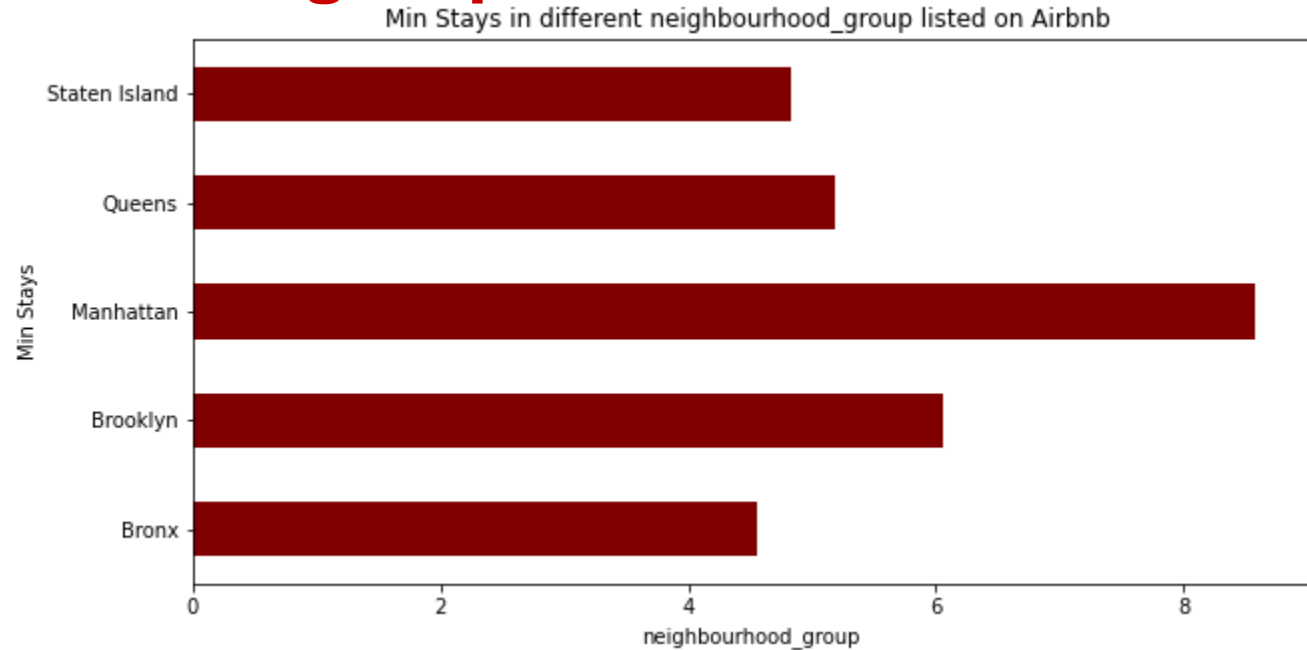
from this observations people mostly living in manhattan on an aver
age of more than 8 nights followed by guests who stayed in brookly
n where average stay is 6 nights

conclusions

1.The people who prefer to stay in Entire home or Apartment they are going to stay bit longer in that particular Neighbourhood only.

2.The people who prefer to stay in private rooms they wont stay longer as compared to home or apartments

3. Most visitors don't prefer shared rooms , they tend to visit private room or entire home

4. Most people prefer to pay less price, ie most of them prefer budget friendly rooms

5. If there are more number of reviews for perticular neighbourhood group that means that place is a tourist place

6. If people are not saying more than one night that means they are travellers.

7. Room availability is high in staten island , so that may be less busiest area

8. Manhattan and brooklyn are the two distinguished , expensive posh areas of NYC

Thank you