

EDA CAPSTONE PROJECT

AIRBNB BOOKING ANALYSIS

By

Soniya Kumawat

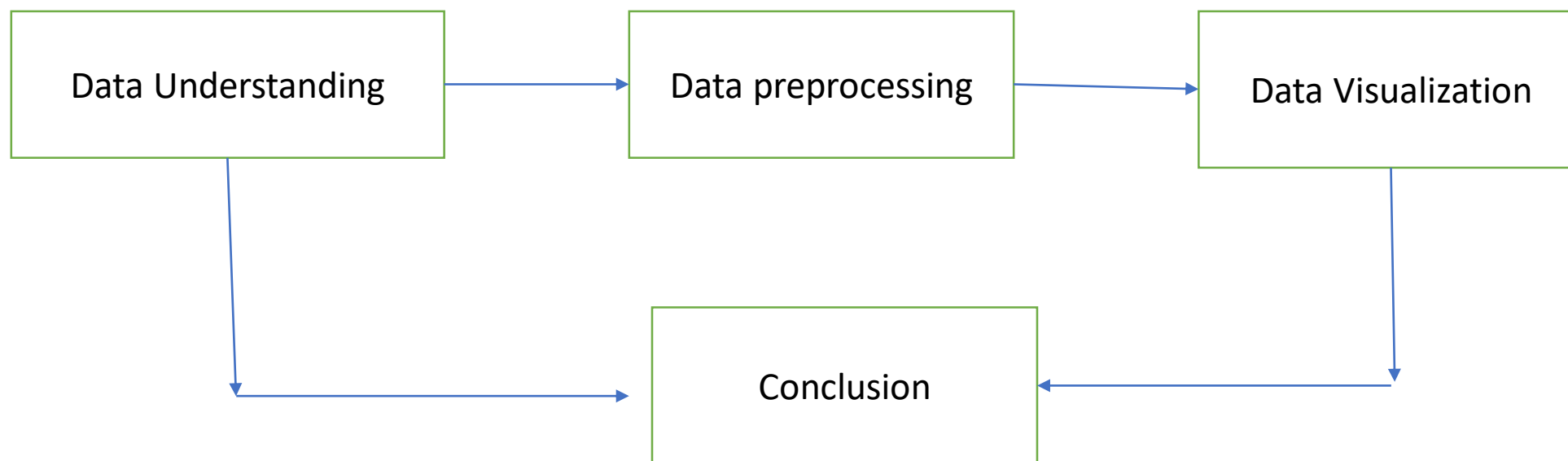
Introduction

- Since 2008, guests and hosts have used Airbnb to expand on traveling possibilities and present a more unique, personalized way of experiencing the world. Nowadays, Airbnb became one of a kind service that is used by the whole world. Data analysts become a crucial factor for the company that provided millions of listings through Airbnb. These listings generate a lot of data that can be analyzed and used for security, business decisions, understanding of customers' and providers' behavior on the platform, implementing innovative additional services, guiding marketing initiatives, and much more.

Problem Statement

- This dataset has around 48885 observation in it with 16 columns and it is a mix between categorical and numeric values.
- Question 1 What can we learn about different hosts and areas?
- Question 2 What can we learn from predictions? (ex:locations, prices, reviews, etc)
- Question 3 Which hosts are the busiest and why?
- Question 4 Is there any noticeable difference of the traffic among different areas? and what could be the reason for it

Methodology:



Data Understanding

- Id - Id of Each row
- Name - Name of Each room given by host
- Host_id - Id of each Host
- Host_name - Name of each Host
- Neighbourhood_group – Borough country Name
- Neighbourhood –All cities of Each Borough Country
- Latitude and Longitude –Geographical Data of Borough Country
- Room_type –Name of each room type
- Price –price of each rooms
- Minimum_Nights-Minimum nights to stay
- Number_of_review- Total reviews given by customer
- Last_reviews- Date of last reviews given by customer
- Reviews_per_month –Average review rate per month
- Calculated_host_listing_count –Count of Host listing in Airbnb
- Availabilty_365 – Number of days

Data preprocessing

- No duplicate data found
- Dropped the columns which we didn't want and have null values like name, last_reviews, reviews_per_month
- Fill the host_name null values with mode
- Groupby data for the desired result

Data Visualization

Top 5 hosts and area

	host_name	neighbourhood_group	calculated_host_listings_count
13217	Sonder (NYC)	Manhattan	327
1834	Blueground	Manhattan	230
9742	Michael	Manhattan	221
3250	David	Manhattan	202
9741	Michael	Brooklyn	168

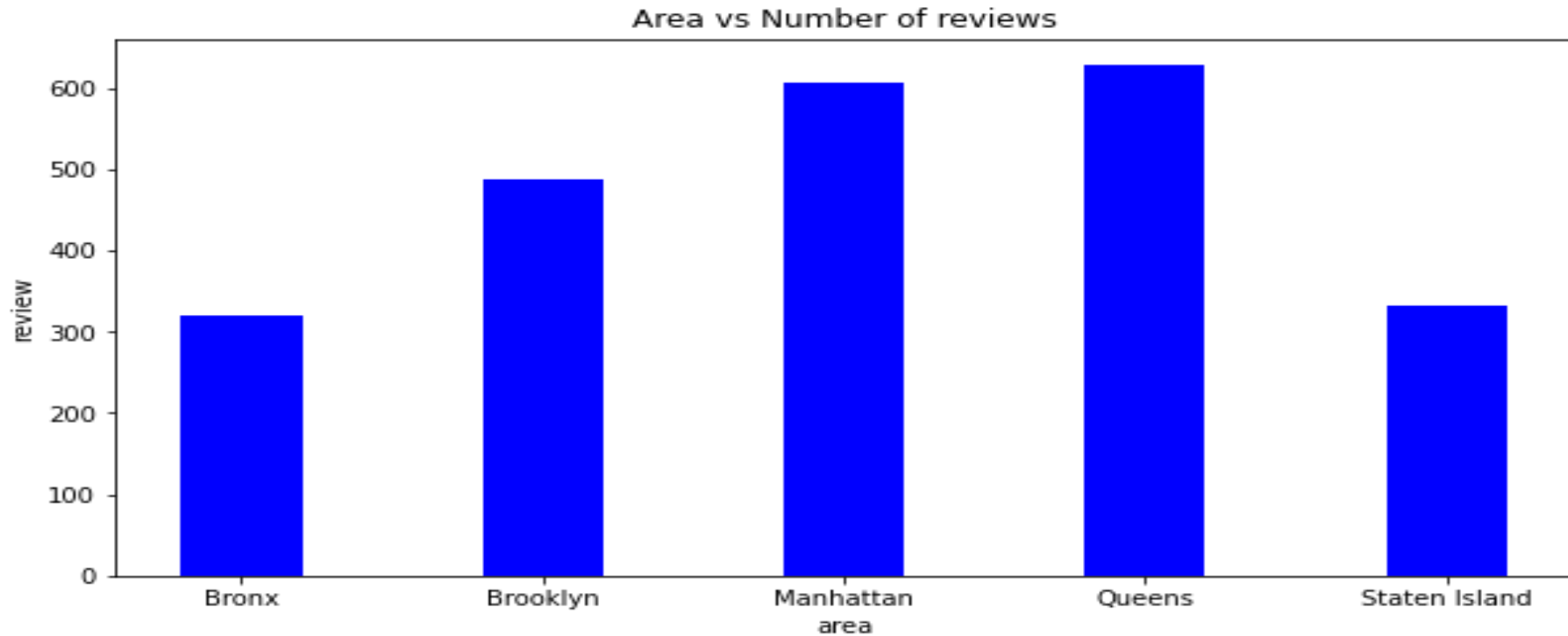
***Conclusion-As we can see most number of listings are from Sonder(NYC), Blueground, Michael,David and area is Manhattan
Manhattan is famous area to live***

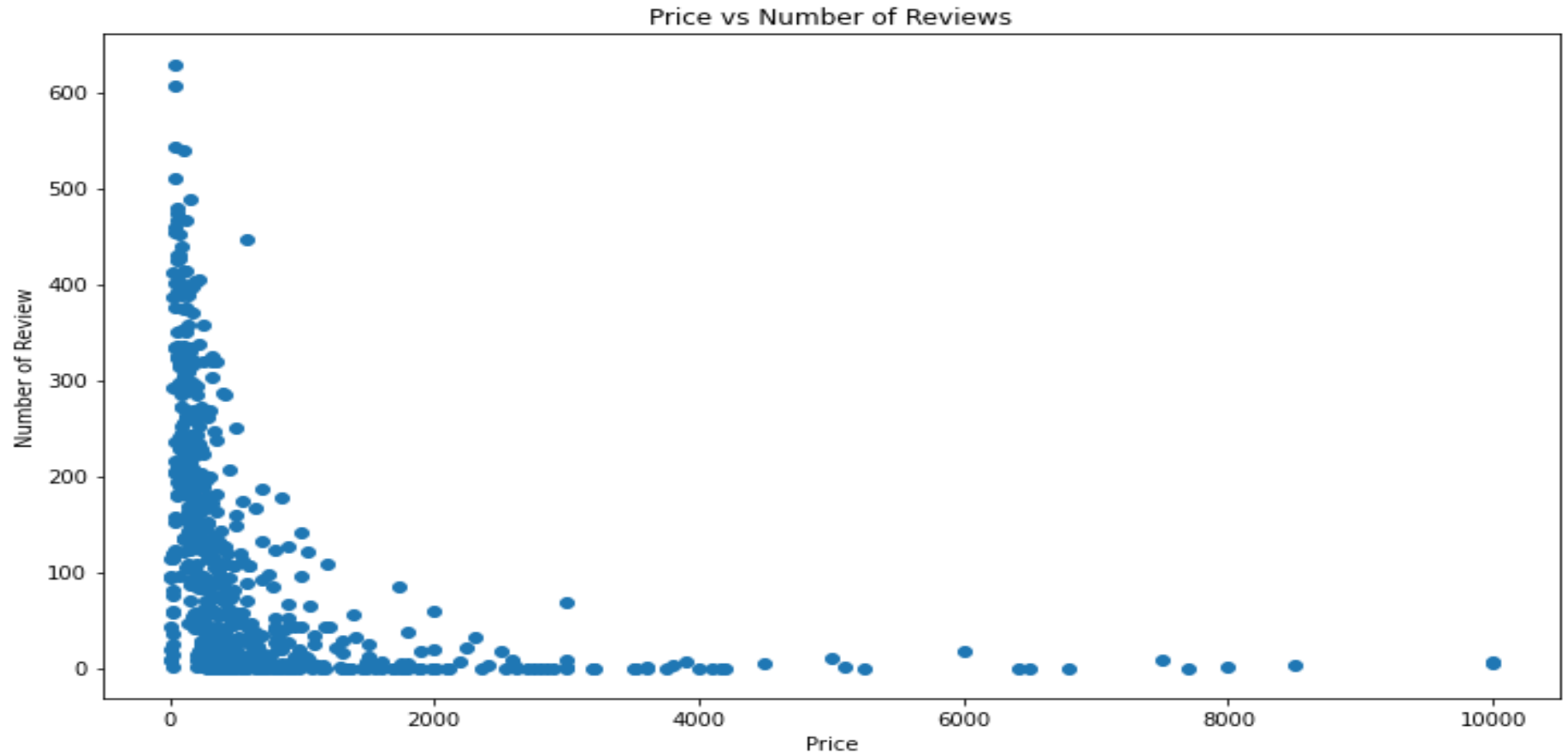
Top five locations and reviews

	Neighbourhod_group	Number_of_reviews
0	Bronx	321
1	Brooklyn	488
2	Manhattan	607
3	Queens	629
4	Staten Island	333

Maximum reviews are from Brooklyn, Manhattan,Queens

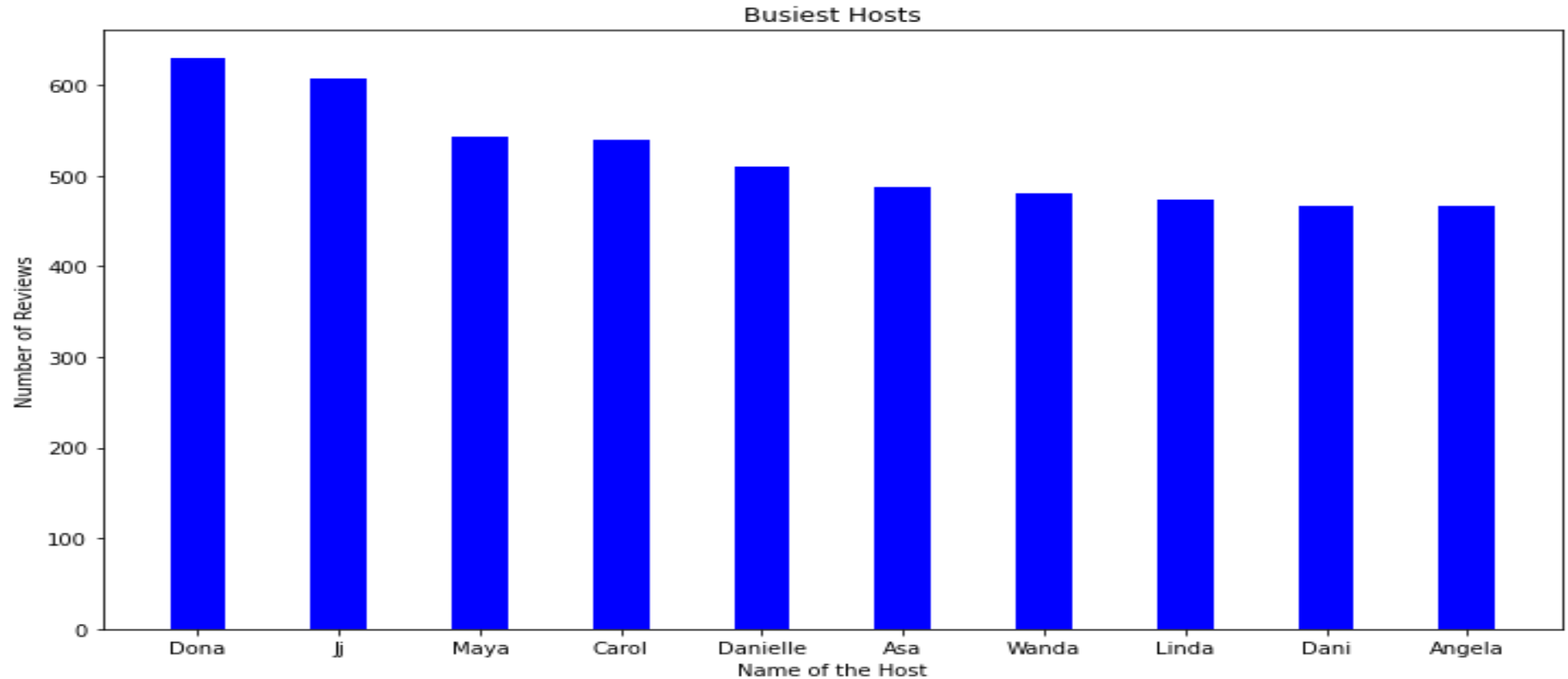
Data visualization of Top five locations and reviews





From the above Analysis we can say that most people prefer to stay in place where price is less.

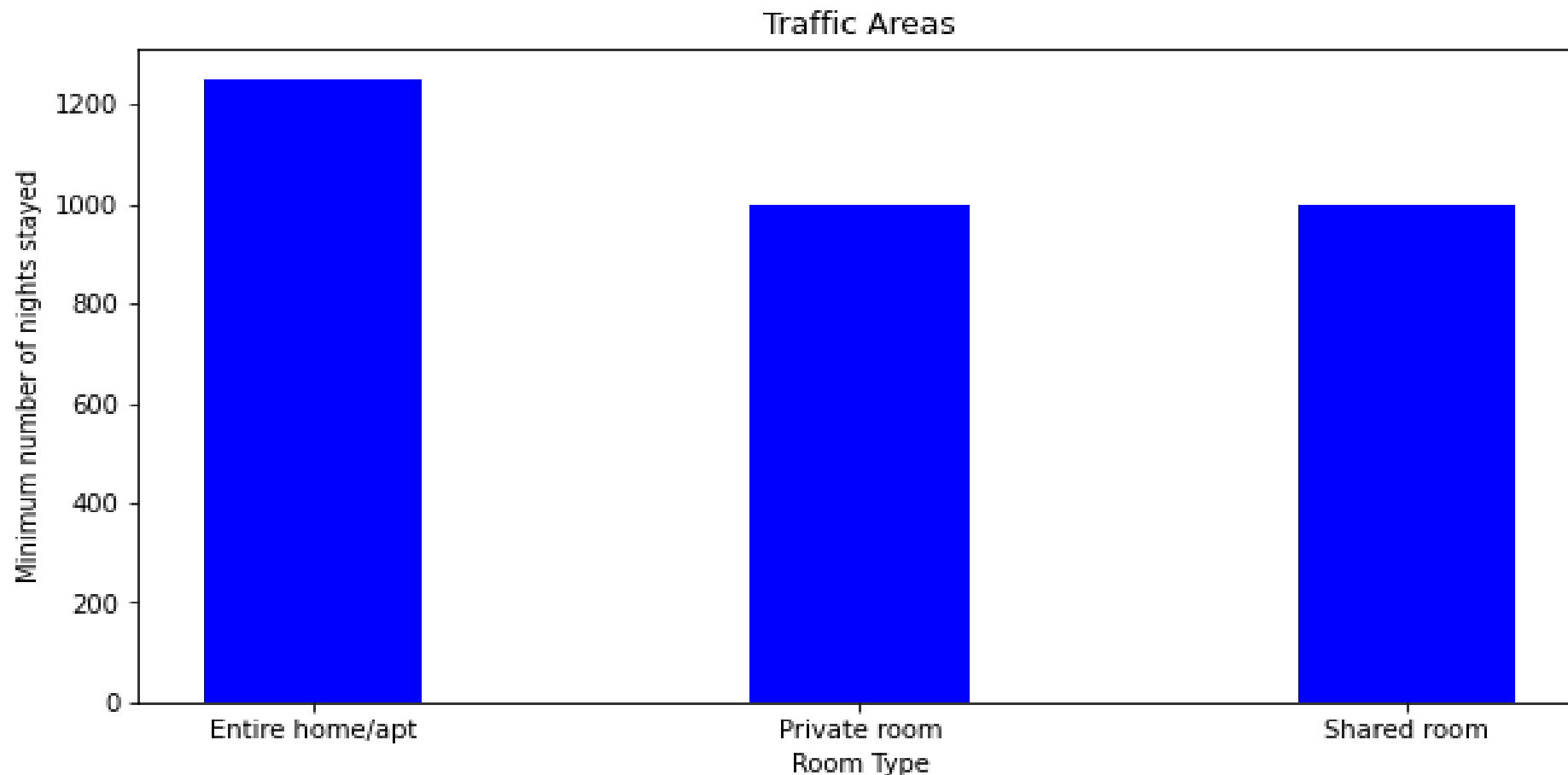
Visualization of Busiest host

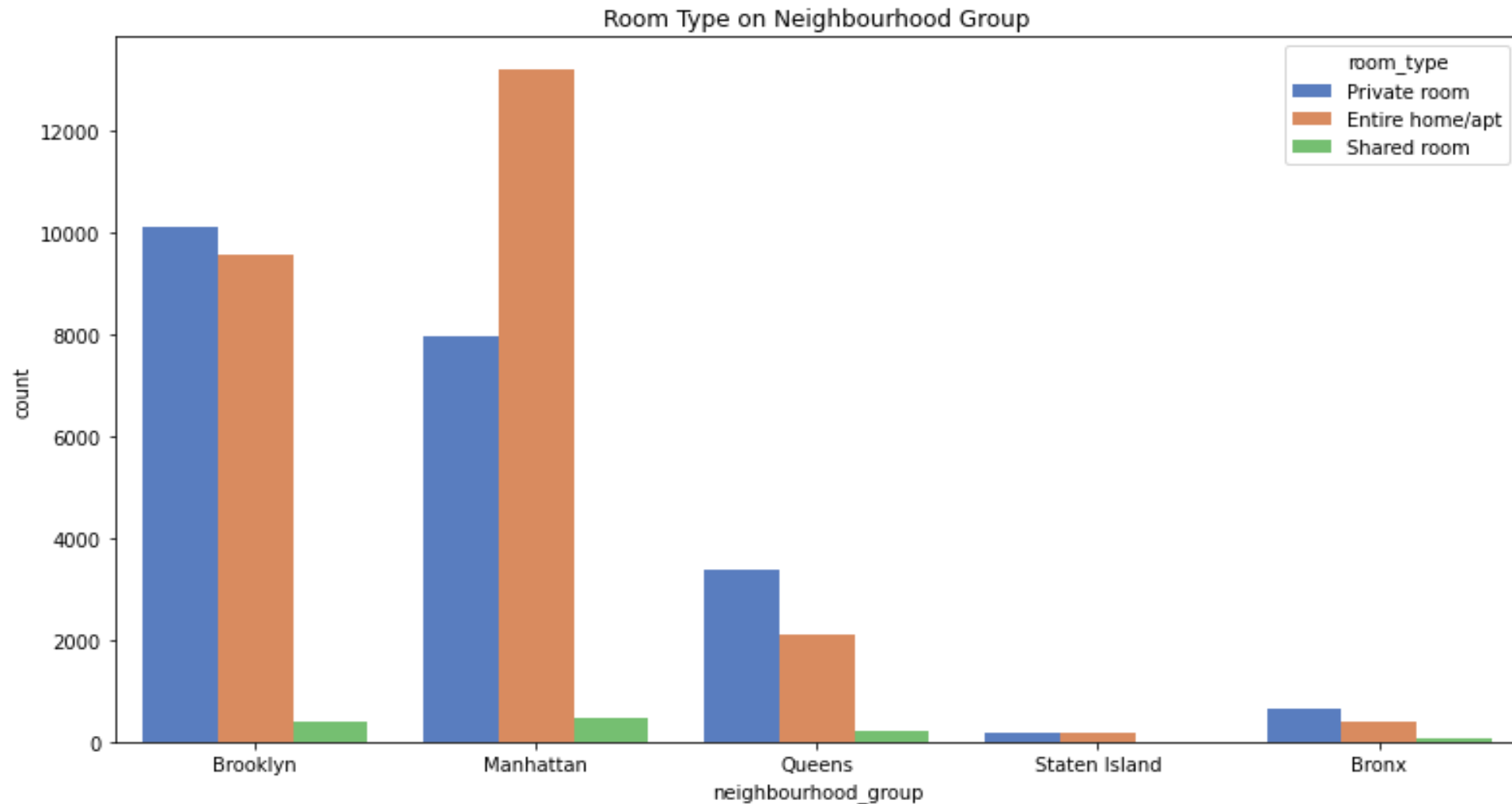


Busiest hosts are: Dona. Maya, Carol, Ji

Because these hosts listed room type as Entire home and Private room as per data given which is preferred by most number of people

Noticeable difference of traffic among different areas and room type used by people





From the Above Analysis We can Say that People are preferring Entire home/apt or Private room which are present in Manhattan, Brooklyn, Queens and people are preferring listings which are less in price.

Conclusion

- **The people who prefer to stay in Entire home or Apartment they are going to stay bit longer in that particular Neighbourhood only.**
- **The people who prefer to stay in Private room they won't stay longer as compared to Home or Apartment.**
- **Most people prefer to pay less price.**
- **If there are more number of Reviews for particular Neighbourhood group that means that place is a tourist place.**
- **If people are not staying more then one night means they are travellers.**

SUMMARY

- We have found hosts that take good advantage of the Airbnb platform and provide the most listings
- We found that our top host has 327 listings.
- We found that people are preferring home more.
- Most famous place to live is Manhattan
- The findings from the EDA can be informative to either hosts or tourists or both.
- Overall, we discovered a very good number of interesting relationships between features and explained each step of the process.

Thank You

