

# Capstone Project -1

## EDA on Airbnb Booking Data Analysis

Presented by  
YASHWATI PATEL

# Content:

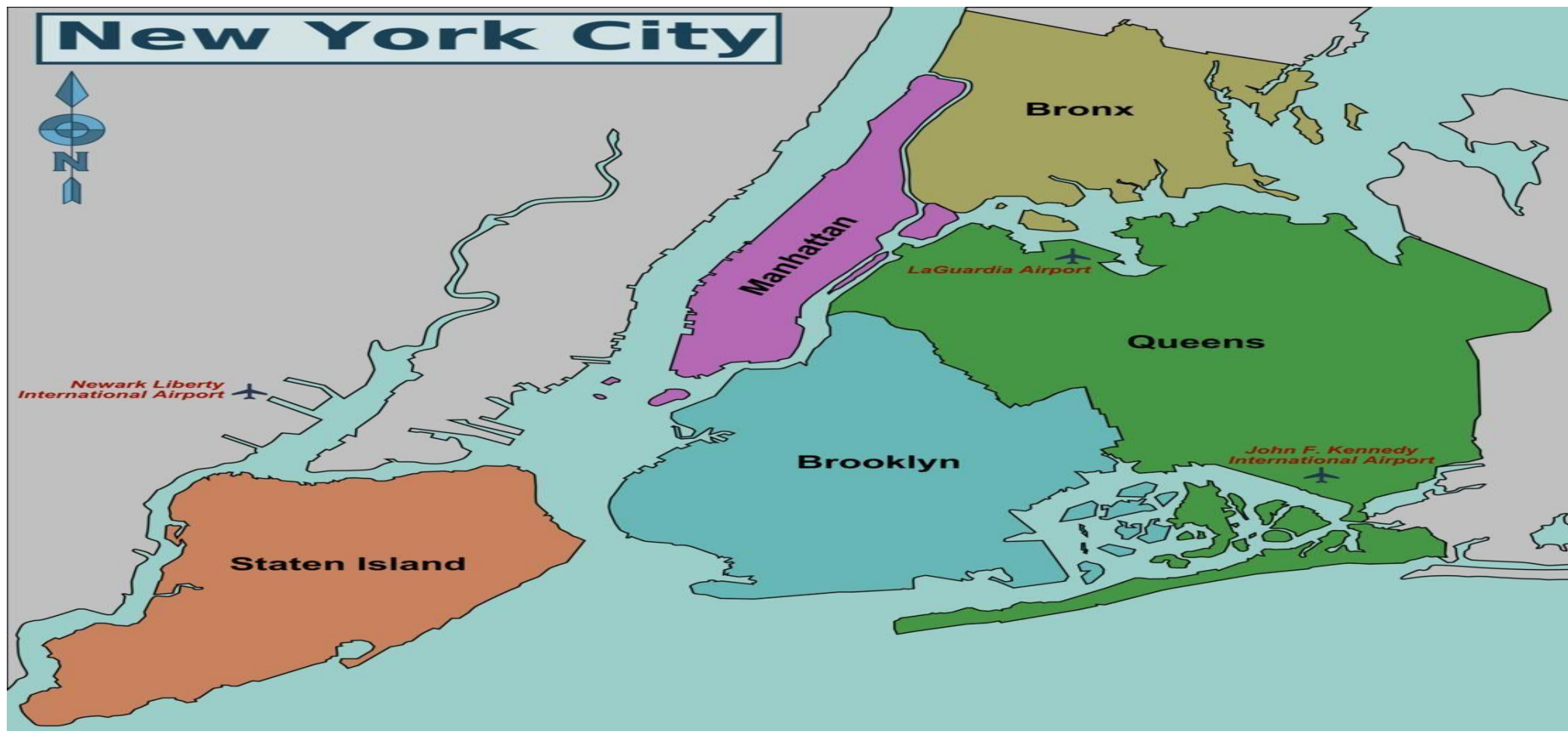
- Introduction
- Problem statement
- Data details
- Data Preparation & Data Wrangling
- Exploring and Visualizing Data
- Conclusion

# Introduction:

## Airbnb:

**Airbnb, Inc. is an American company that operates an online marketplace for lodging, primarily homestays for vacation rentals, and tourism activities. Based in San Francisco, California, the platform is accessible via website and mobile app. Airbnb does not own any of the listed properties; instead, it profits by receiving commission from each booking. The company was founded in 2008 by Brian Chesky, Nathan Blecharczyk and Joe Gebbia.**

❖ Map of newyork city



## Problem statement:

- Key objects what We will be find out from this project:
- what can we learn about different hosts and areas?
- what can we learn from predictions?(ex: locations,prices,reviews, etc)
- which hosts are busiest and why?
- Is there any noticeable difference of traffic among different areas and what could be the reason for it?

# Data details

## Numeric Data

- Id
- Host Id
- Latitude
- Longitude
- Price
- Minimum nights
- Number of reviews
- Reviews per month
- Calculated host listings count
- Availability 365

## Categorical Data

- Name
- Host name
- Neighbourhood group
- Neighbourhood
- Room type
- Last review

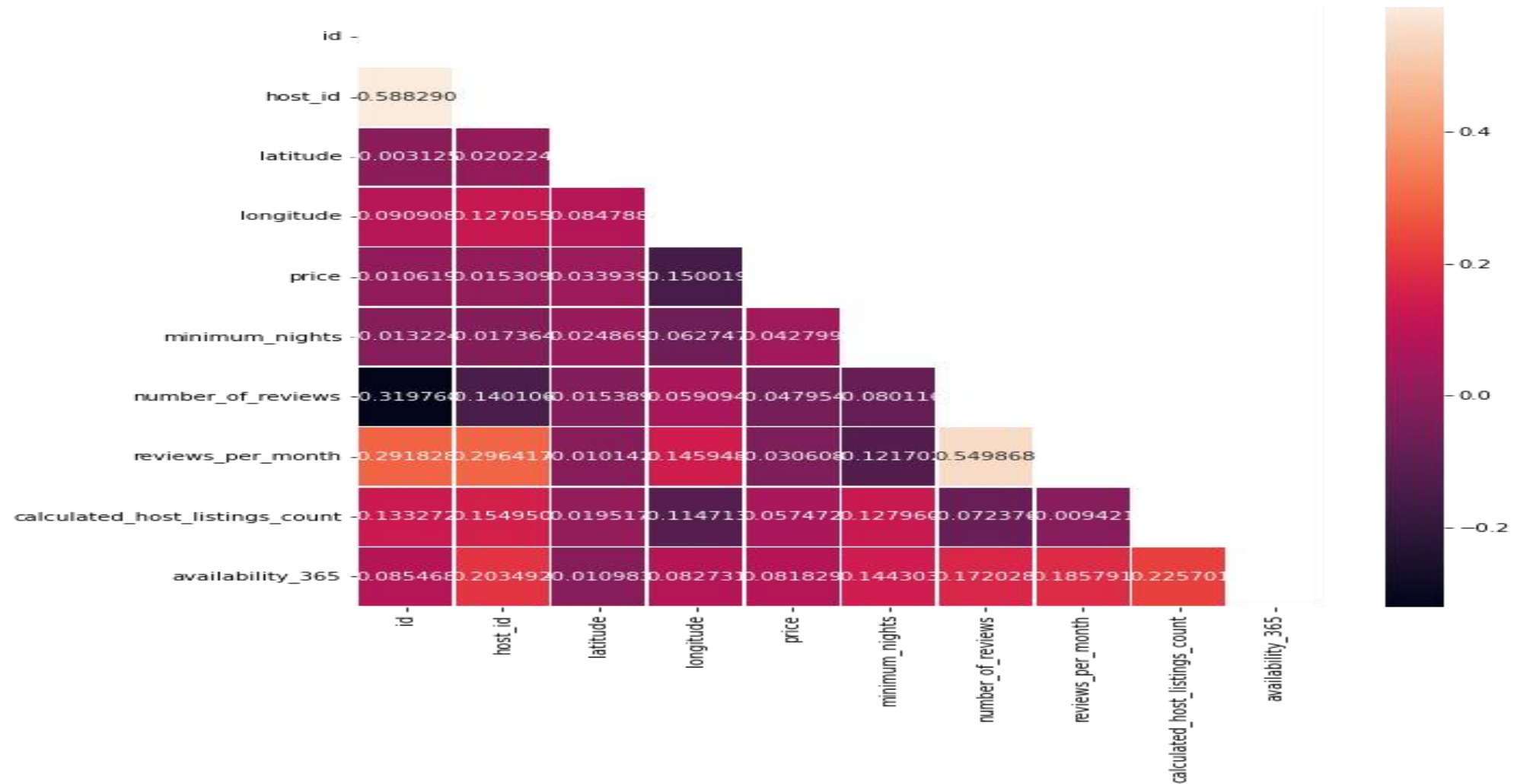
# ❖ Data preparation and Data wrangling

- **Data exploration:** visualization of data is made to analyze and understand data.
- **Dealing with missing values:** In datasheet we can remove null values or replace them as NaN value.
- **Reshaping data:** The data is modified and manipulated according to the requirements.
- **Filterinng data:** The unnecessary rows and columns are are filtered.
- **Others:** Other functions also used to making efficient data.

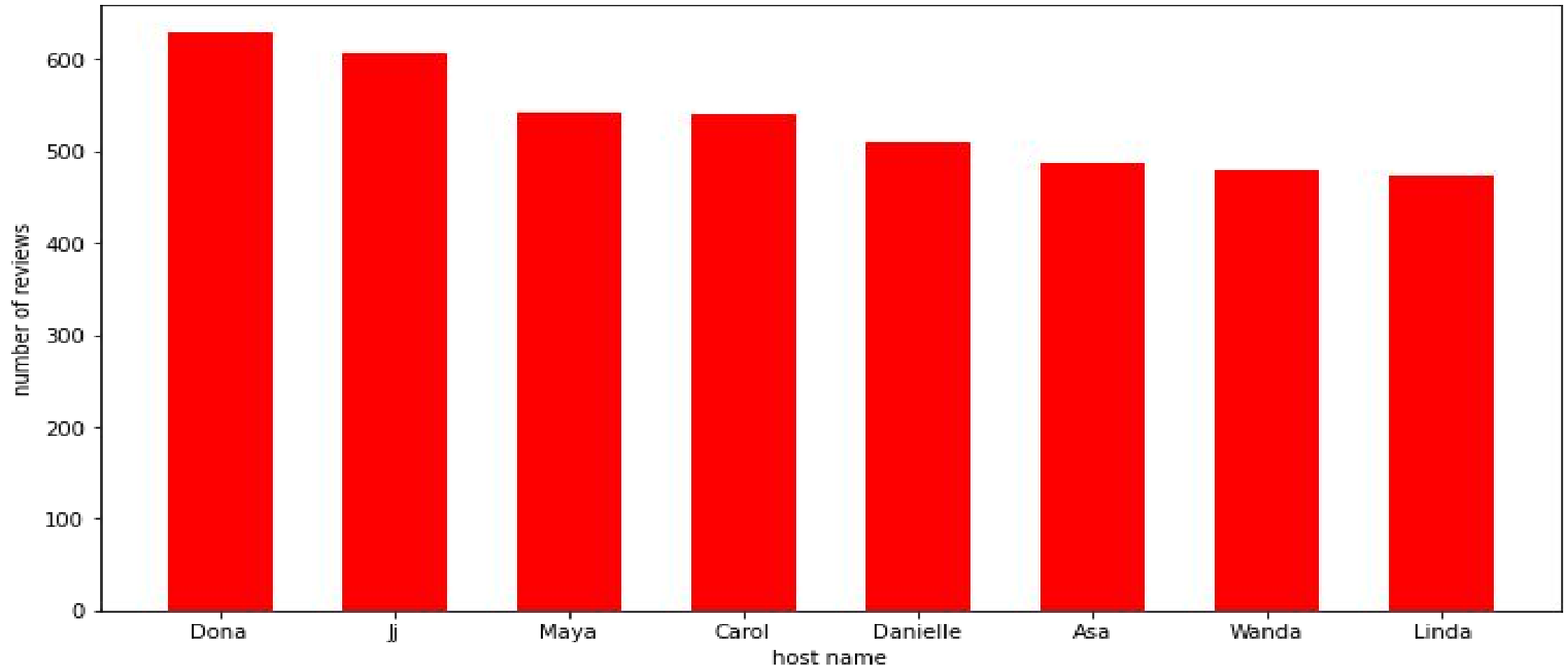
# *DATA VISUALIZATION*



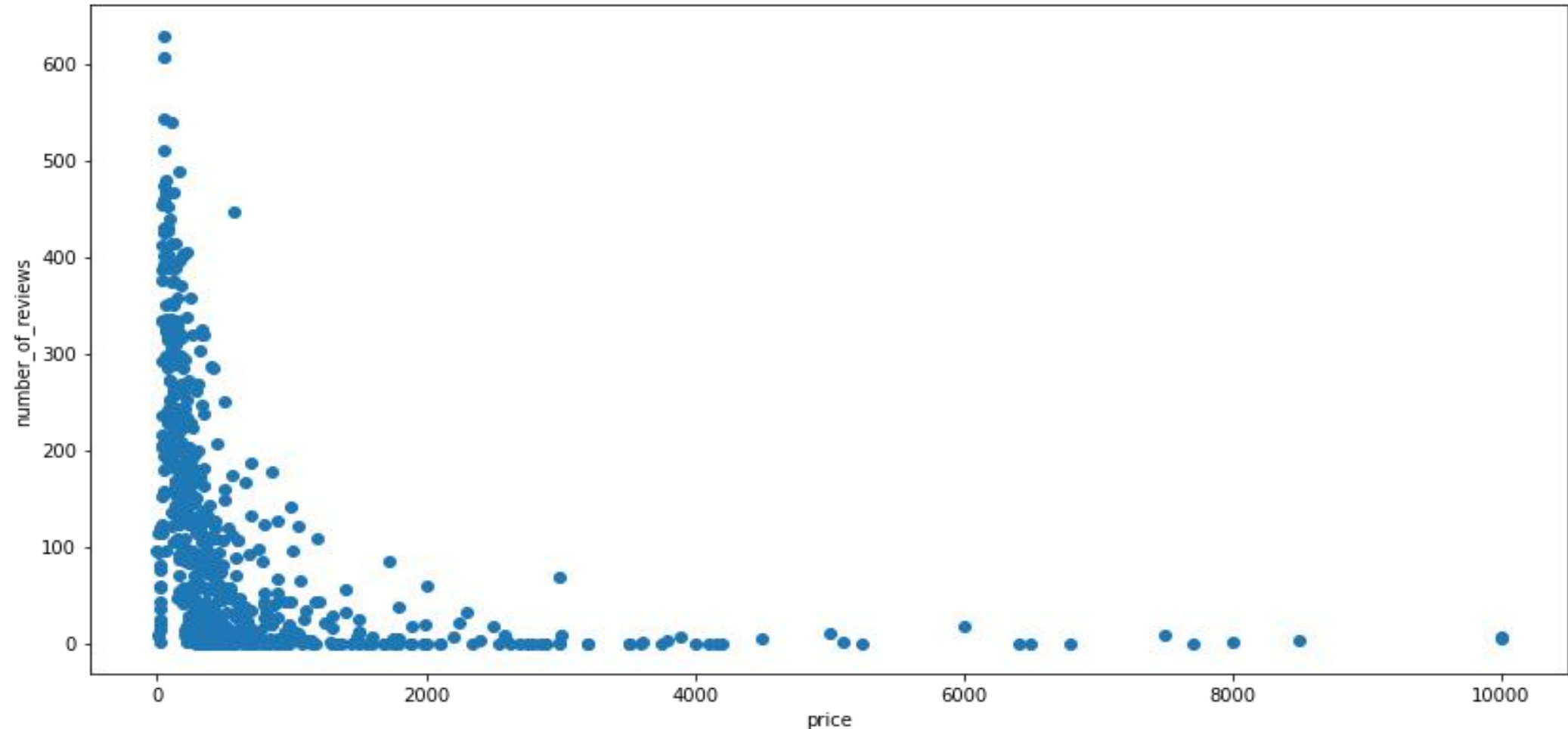
# Data Correlation Matrix



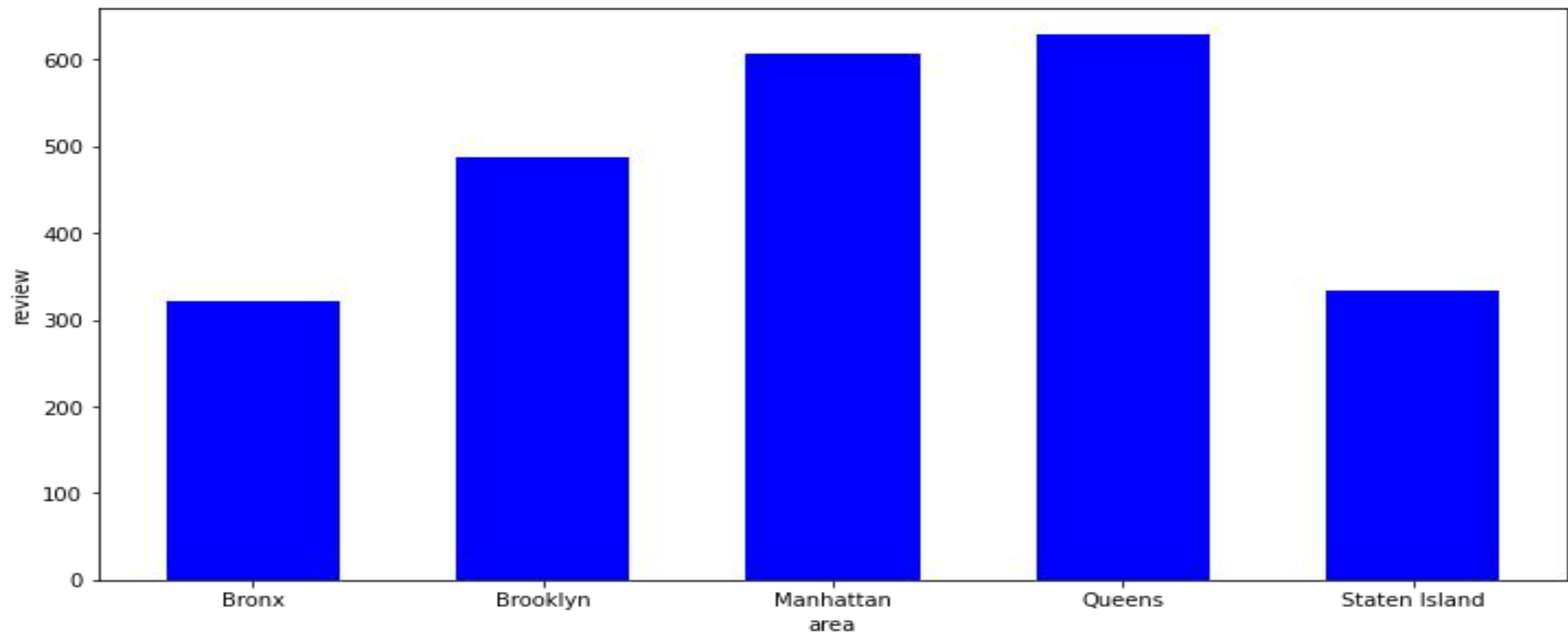
- From this plot we can say that most of busiest hosts are Dona, Ji, Maya, Carol and Danielle because these hosts listed room type as entire home/apt and private room which is preferred by most of people.



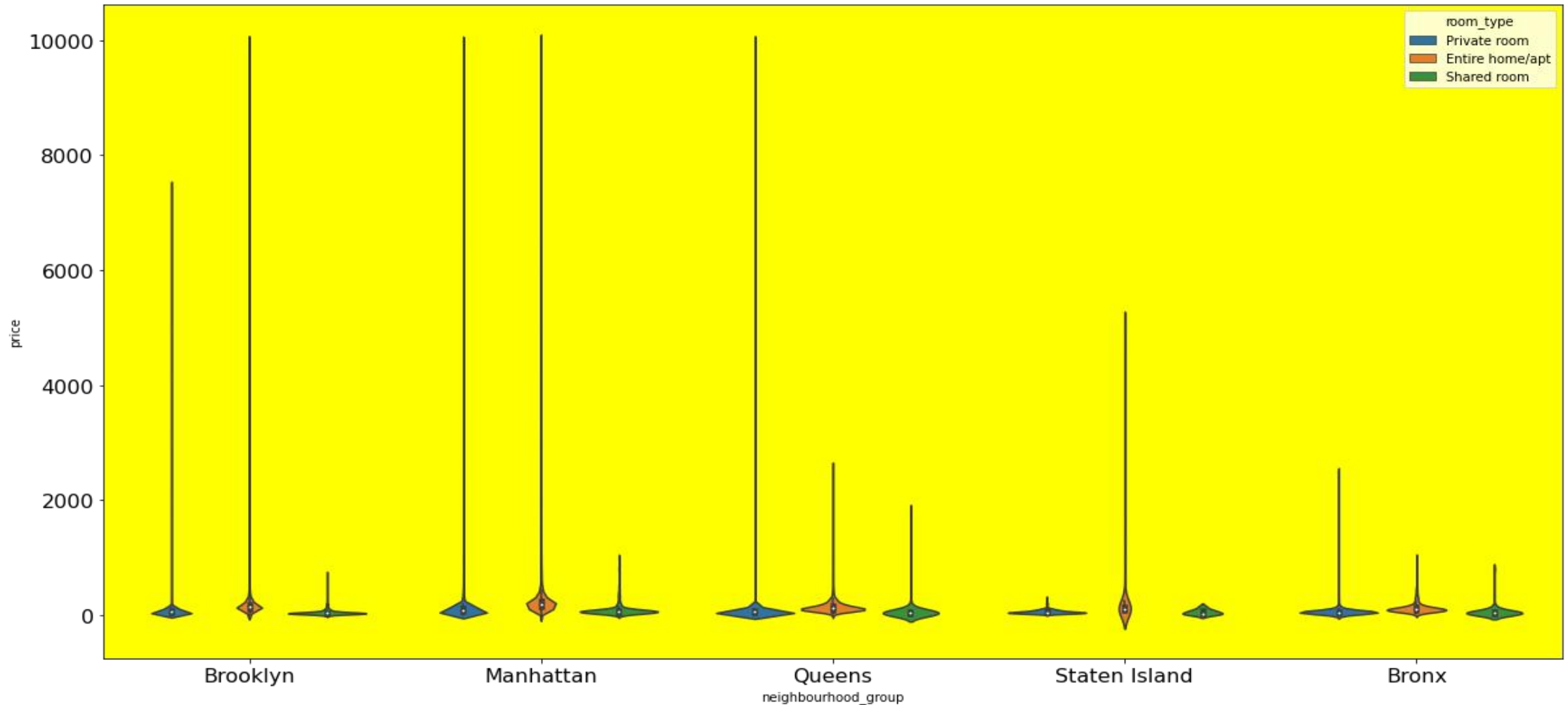
- From the analysis we can say that most people prefer to stay in which areas where prices is less.



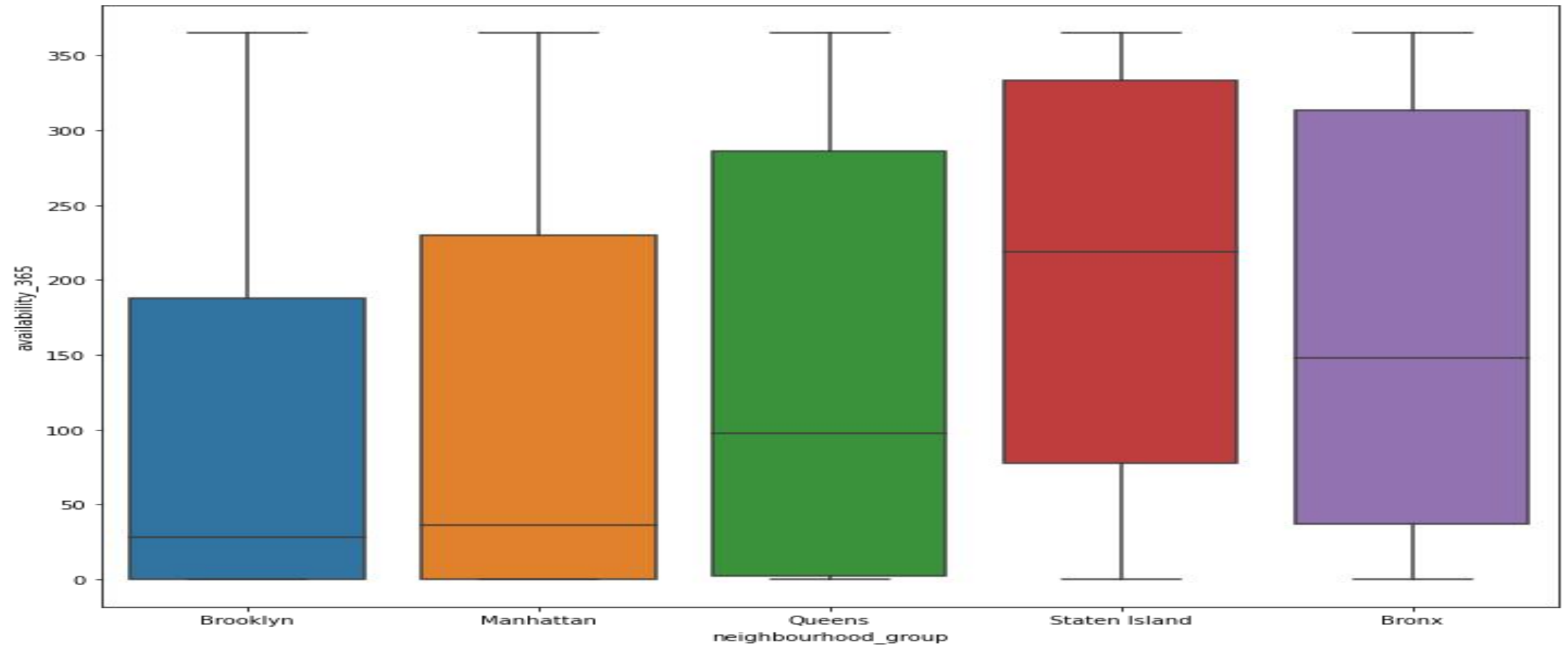
- Neighbourhood group vs number of reviews



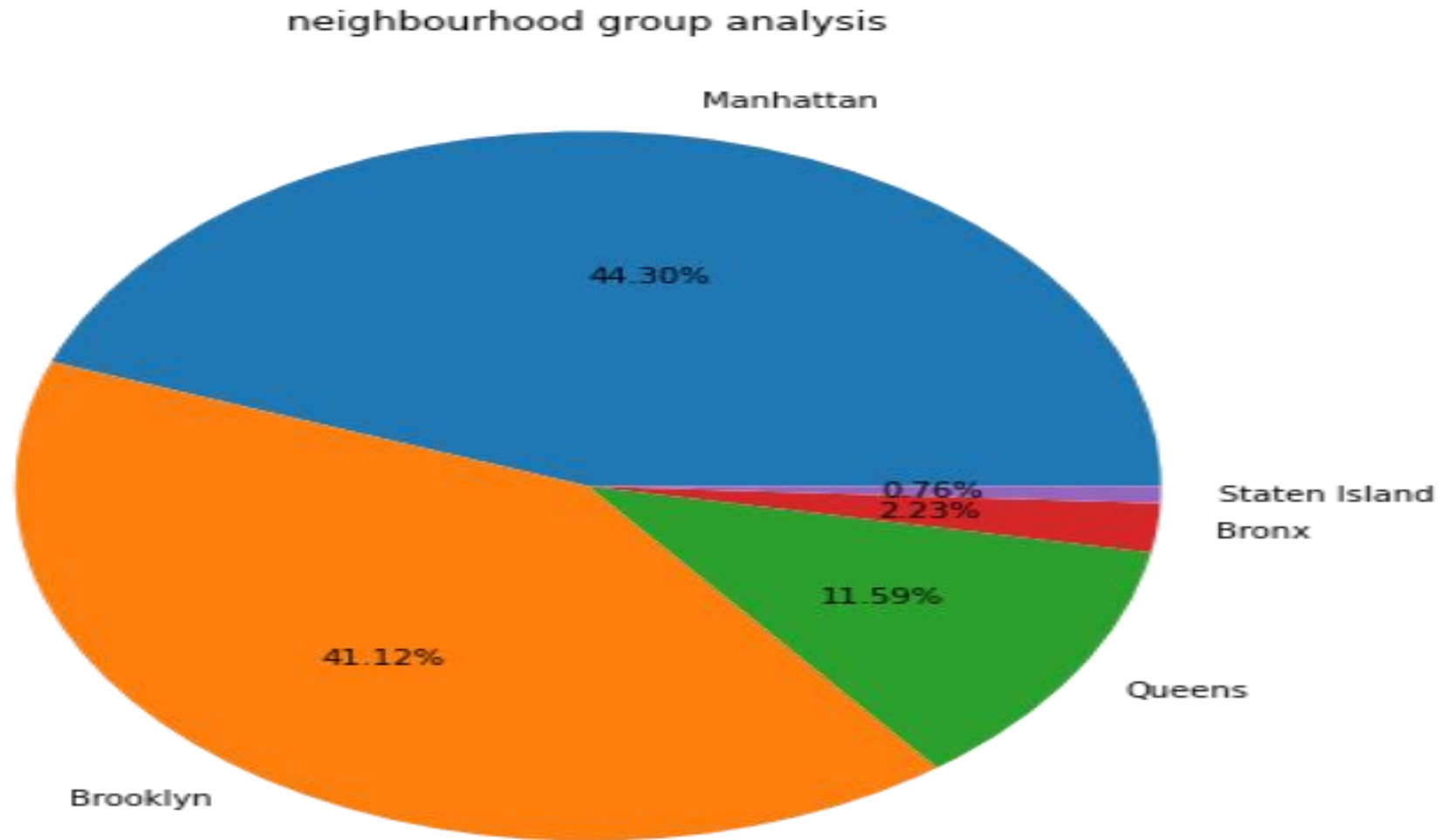
- Here , from this plot we can say that most of rooms are in prices below 2000 in all neighbourhood group. Manhattan,Queens and Brooklyn has most number of rooms in private and entire home/apt type rooms.



- In this plot we can find that island has the highest mean availability value compared to others.



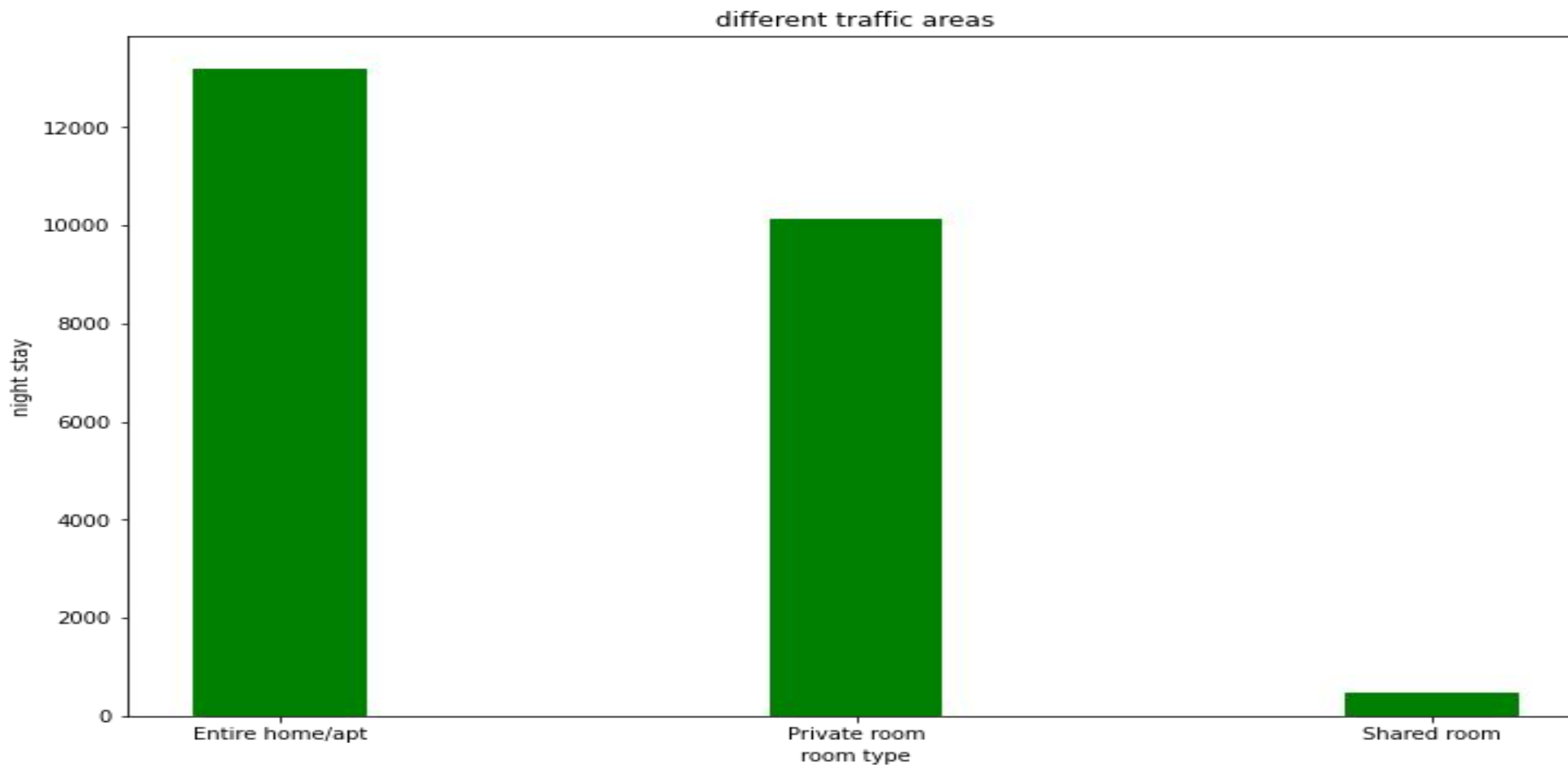
## ❖ Neighbourhood Group Analysis



- From the above analysis, We have to check out the percentage of each neighbourhood group. we can see that many rooms are rented out in two areas Manhattan 44.3% , Brooklyn 41.1% and Staten island cover very low area as compare to other neighbourhood group.



- In this analysis we find that most of people preferred entire home and private room for night staying.



# ❖ Conclusion

- **Most Airbnb services is located in Manhattan and Brooklyn due to capital income state.**
- **In the category of room type we analysis that entire home/apt and private room are the most preferred room.**
- **sondar(NYC) is the busiest host among all.**
- **The neighbourhood group 'Manhattan' has highest number of listings in entire NYC.**
- **Manhattan is the most focused place in New York for hosts to do their business.**

*Thank You !*

