

# Capstone Project - 1

## EDA on Airbnb Booking Analysis

By

**Pankaj Kumar,  
DSN Naveen Kumar,  
Ankit Sharma**

**Data Science Trainee, AlmaBetter**

# Problem Statement

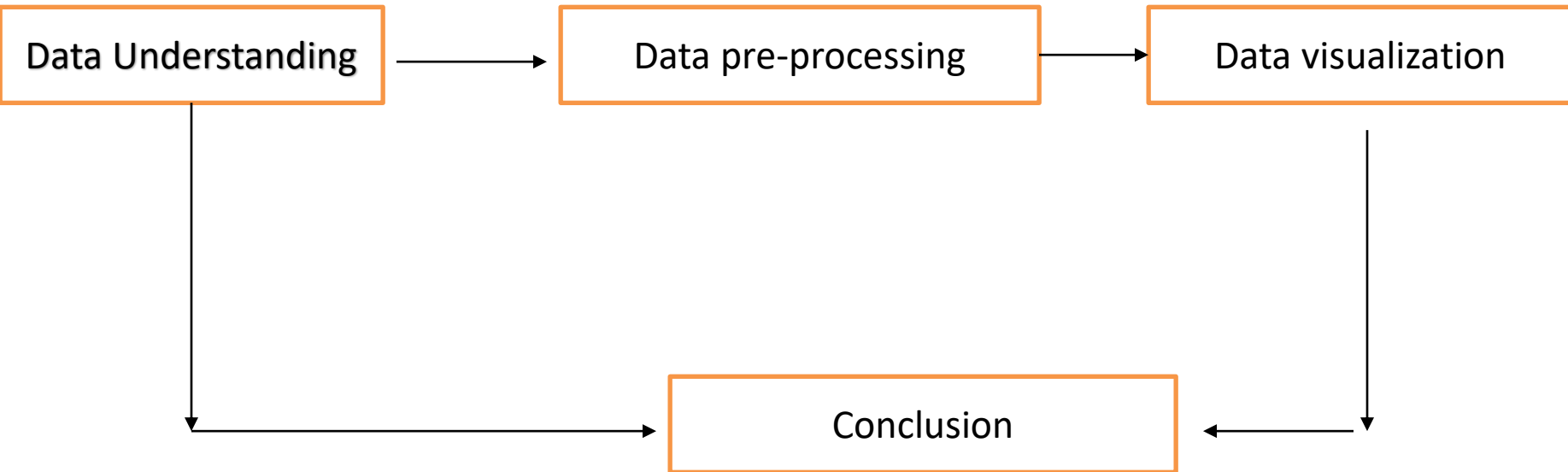
- ❑ One dataset is provided, with **information about Airbnb Bookings.**
- ❑ We must examine and evaluate the data in dataset in order to identify the important characteristics that influence Airbnb Bookings across the world.

## So, what factors influence Airbnb Bookings ?

An Airbnb booking Analysis Depends upon below mentioned factors:-

- ❑ Predictions like Location , Price, Reviews
- ❑ Different Hosts and Areas they choose
- ❑ Busiest Hosts
- ❑ Traffic among different areas.





# Data Summary

## Airbnb Booking Data

- |  |  |
|--|--|
| <input type="checkbox"/> ID                  | <input type="checkbox"/> Longitude                     |
| <input type="checkbox"/> Name                | <input type="checkbox"/> Room type                     |
| <input type="checkbox"/> Host ID             | <input type="checkbox"/> Price                         |
| <input type="checkbox"/> Host Name           | <input type="checkbox"/> Minimum Nights                |
| <input type="checkbox"/> Neighbourhood Group | <input type="checkbox"/> Calculated host listing count |
| <input type="checkbox"/> Neighbourhood       | <input type="checkbox"/> Availability around year      |
| <input type="checkbox"/> Latitude            |  |

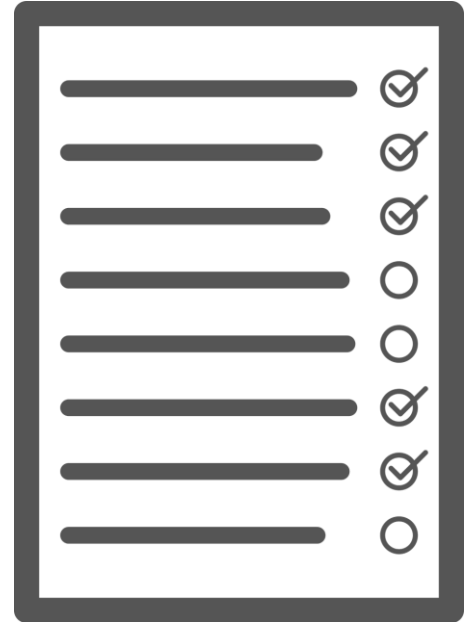
## User\_reviews

- ☐ Number of reviews
- ☐ Last review
- ☐ Reviews per month



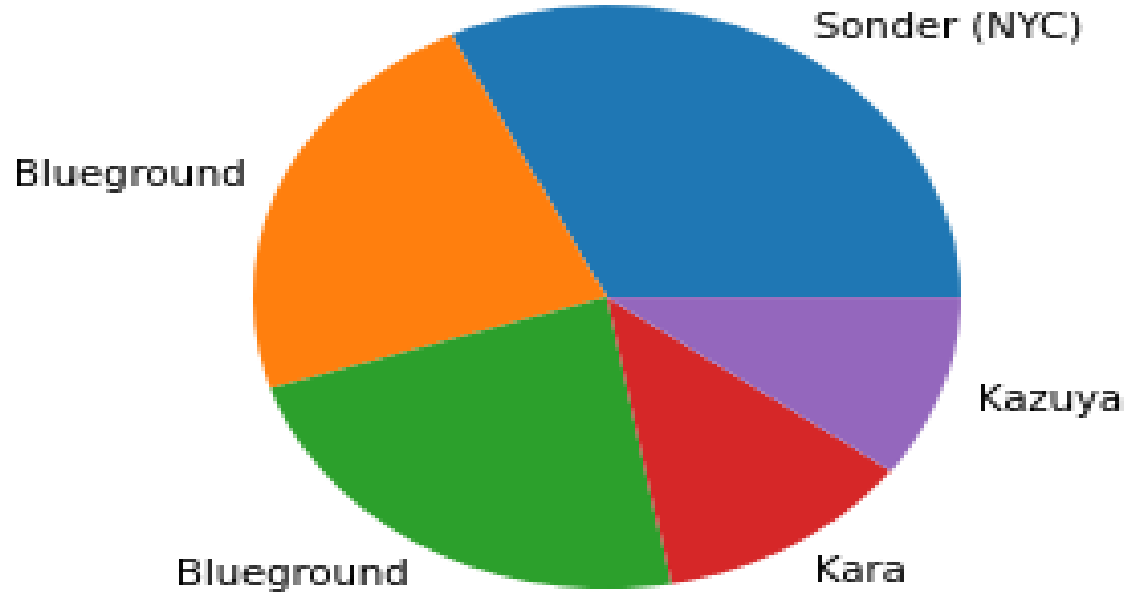
# Agenda

- ☐ Room Type Rating Analysis
- ☐ Categorical Analysis
- ☐ Airbnb Rating Analysis
- ☐ Busiest host and areas
- ☐ Average stay at a particular area
- ☐ Most Popular Areas
- ☐ Data realisation
- ☐ Reviews Analysis
- ☐ Challenges Faced
- ☐ Analysis Summary



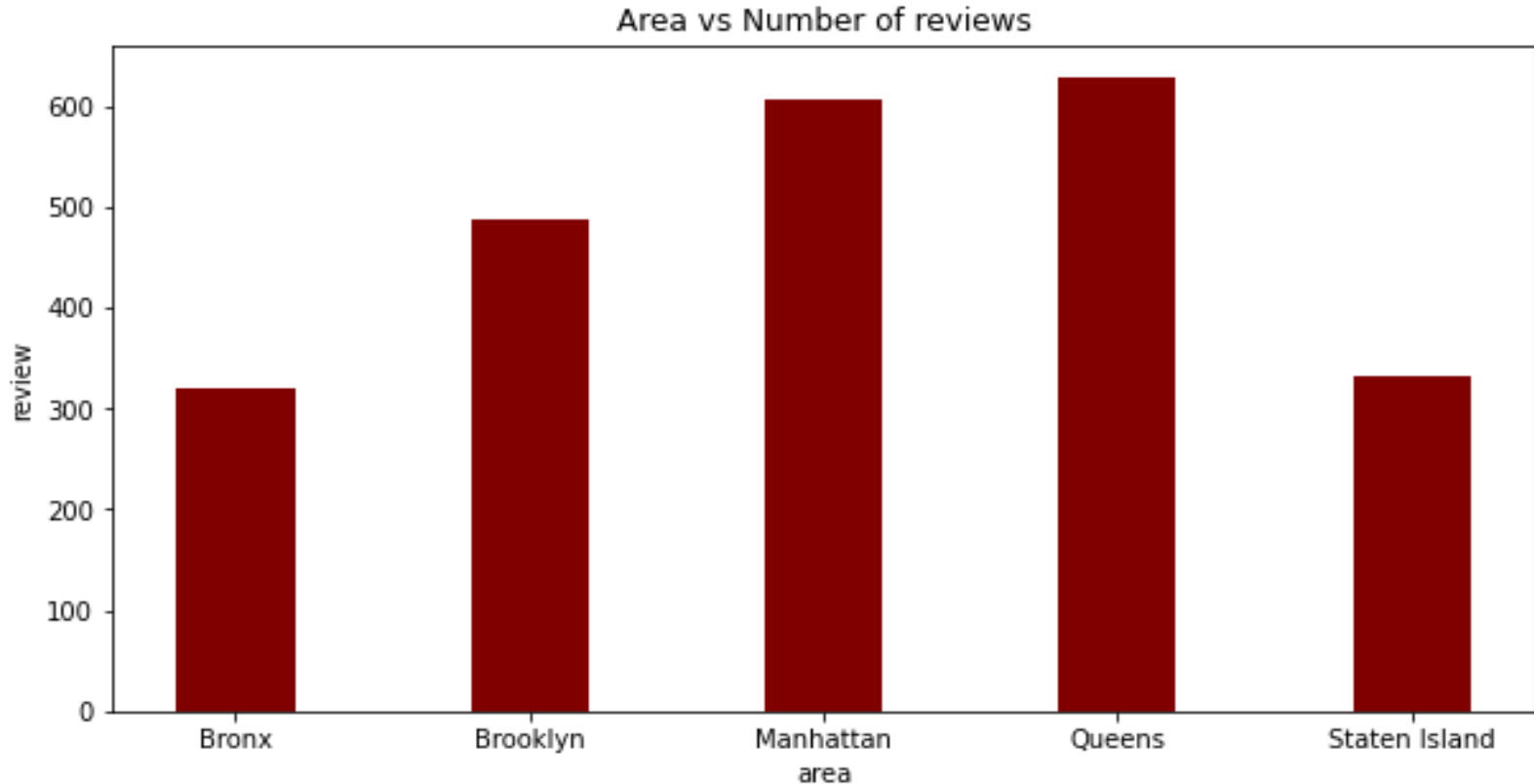
# Hosts Vs Area

☐ Most number of listings



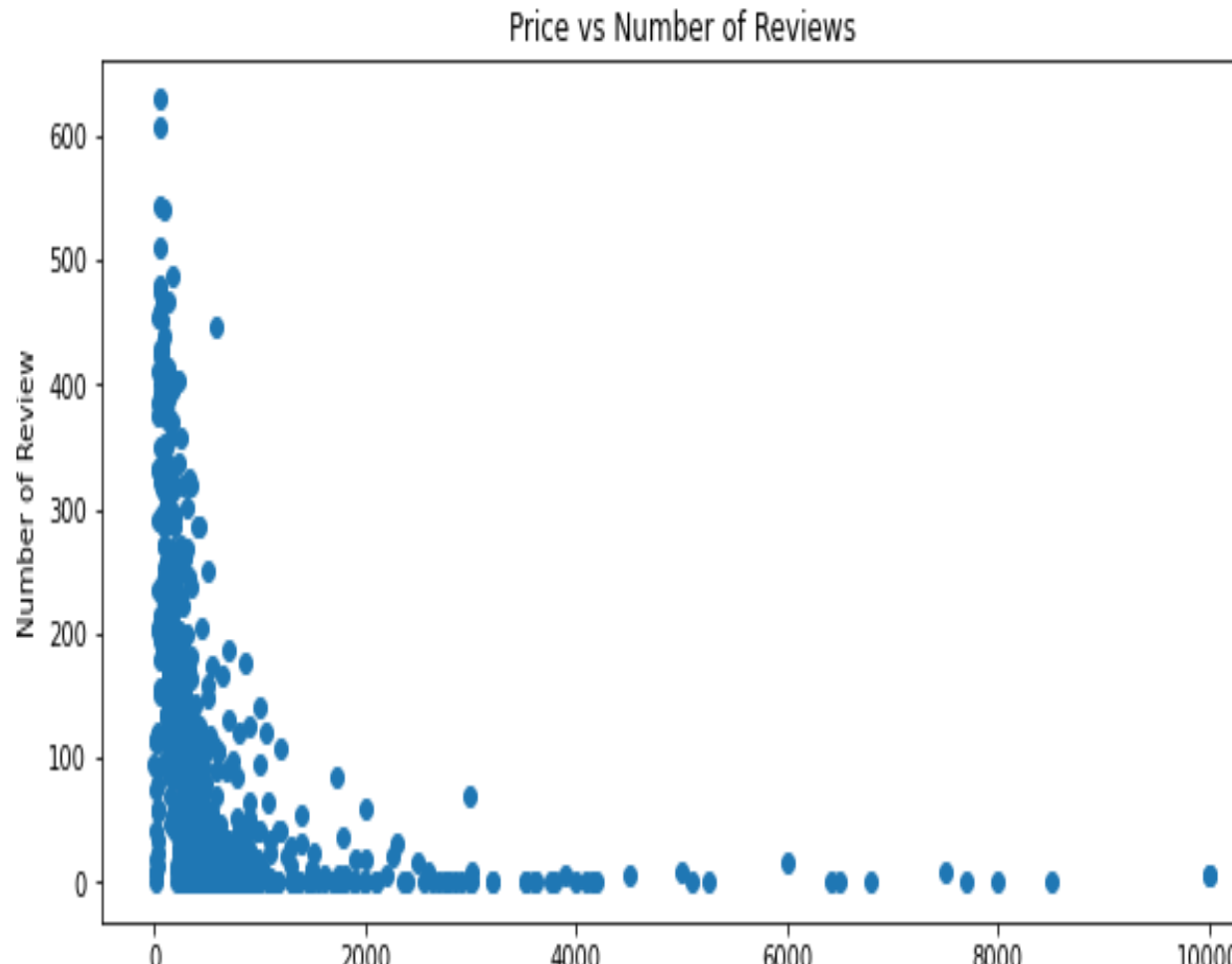
# Categorical Analysis

□ Area Vs number of reviews



# Categorical Analysis (Contd.)

- ❑ From the analysis, we can say most people prefer to stay in place where price is less.
- ❑ We can conclude from the graph that hotels with low prices have the highest number of reviews.

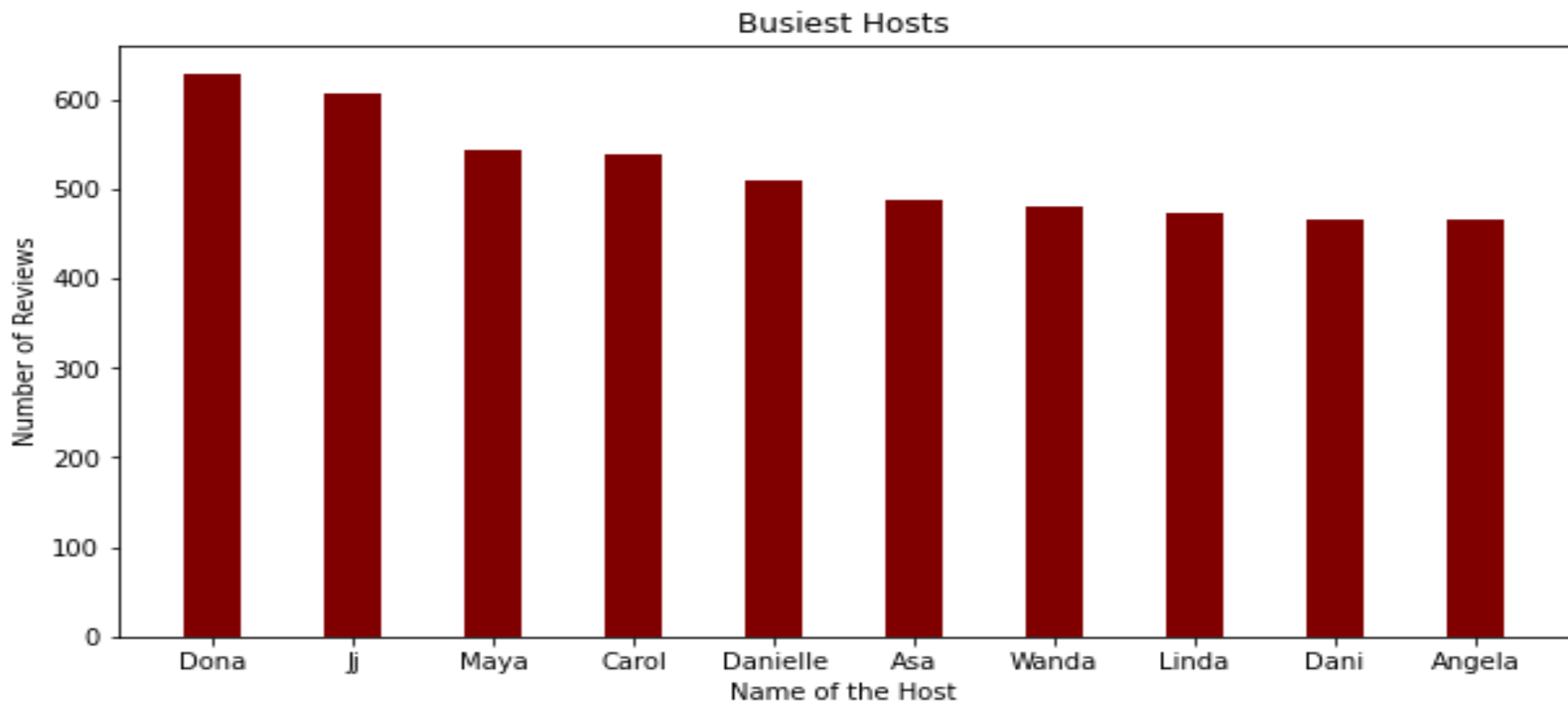




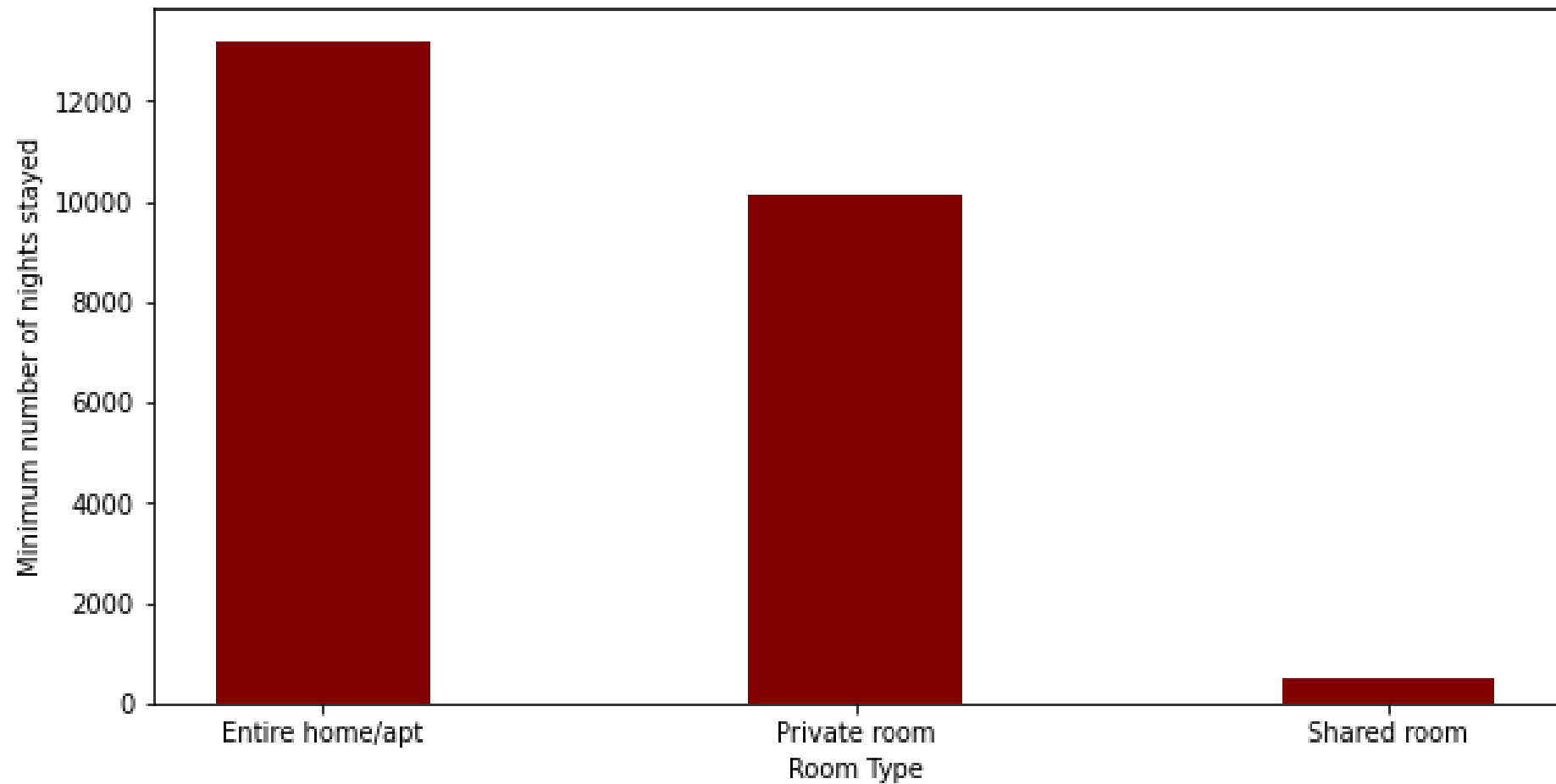
# Top busiest Hosts

- ❑ There are a total of 10 Hosts.
- ❑ Hosts listed room type as Entire home and Private room which is preferred by most number of people.
- ❑ From the data we can conclude that top 5 hosts are :-
  - ❑ 1. Dona
  - ❑ 2. Ji
  - ❑ 3. Maya
  - ❑ 4. Carol
  - ❑ 5. Danielle

# Top Busiest Hosts Continued:-



Traffic Areas



## Difference of traffic among different areas

- ❑ We can conclude from above graph that are preferring Entire home/apt
- ❑ People are less lean towards shared room.
- ❑ We can Say that People are preferring Entire home/apt or Private room which are present in Manhattan, Brooklyn, Queens.
- ❑ People are preferring listings which are less in price.

# Data Visualization through Tableau and Ms- Excel

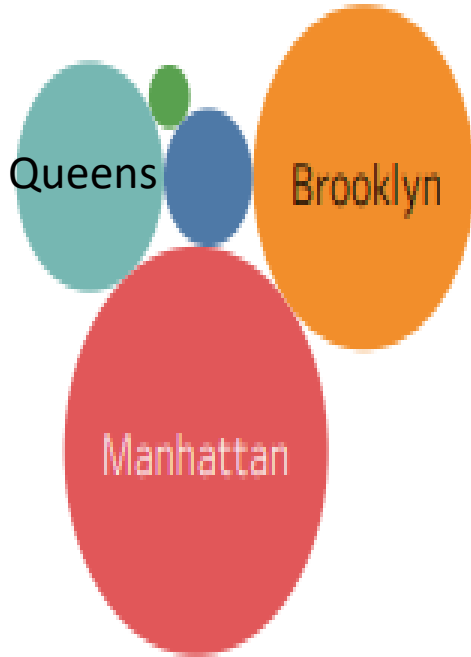
- ❑ With the help of Tableau and MS-Excel we will infer the data about different host and areas related to them.
- ❑ With these tools we will answers these question :-
- ❑ Different host listings in neighbourhood groups.
- ❑ Number of reviews given to a particular room type.
- ❑ Category of room and it's price with respect to different neighbourhood.
- ❑ Maximum how many nights does a host stayed in a particular room type.

Hosts listings in different neighbourhood groups :-

❑ Blue colour indicates **Staten Island**

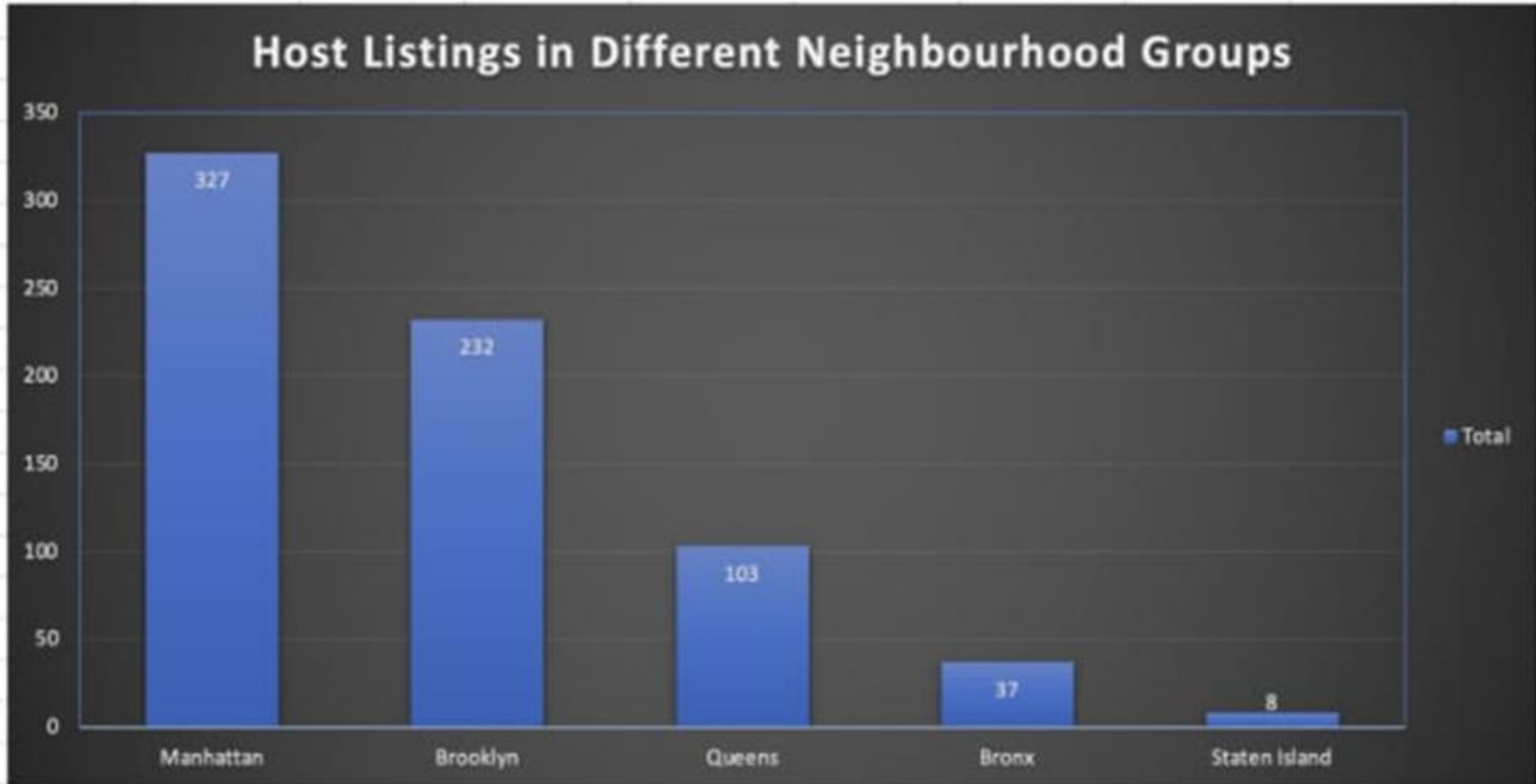
❑ Green colour indicates **Bronx**

Through Tableau



**Different Host Listings** shows that hosts are majorly concentrated towards Manhattan followed by Brooklyn then Queens then Staten Island then Bronx.

# Data through Excel



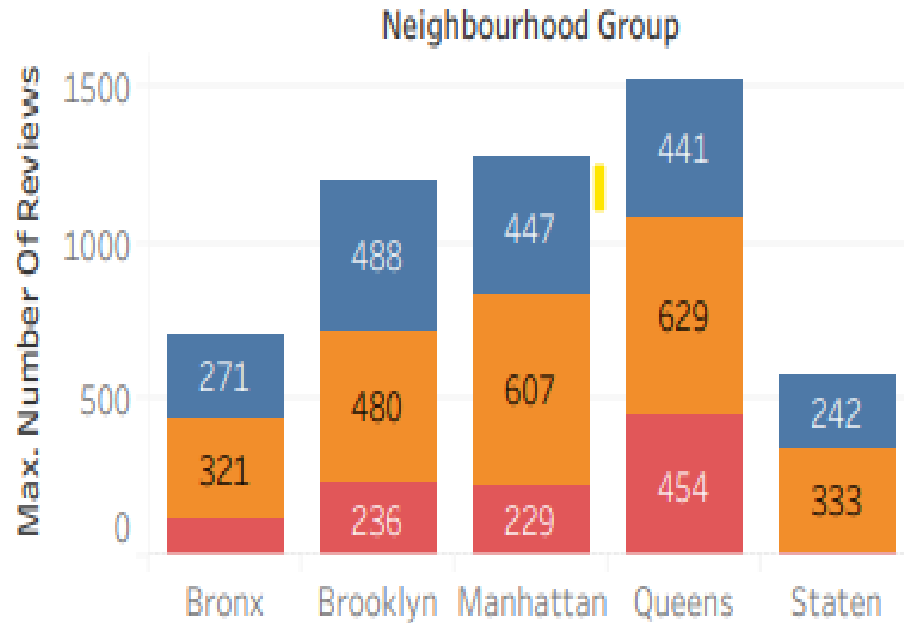
# Number of reviews with respect to room type

## Through tableau

Saffron colour Show **Private Room**

Blue colour shows **Entire home/Apt**

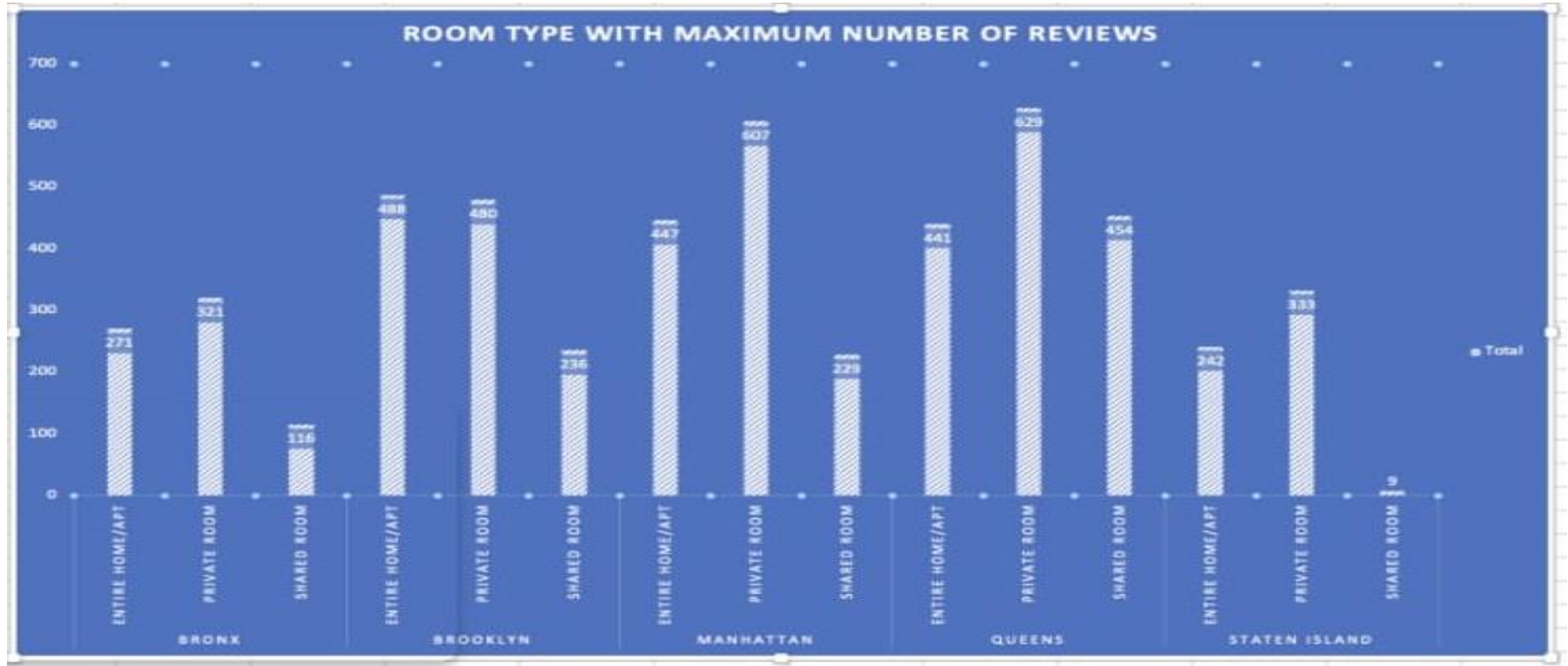
Pink colour Shows **Shared Rooms**



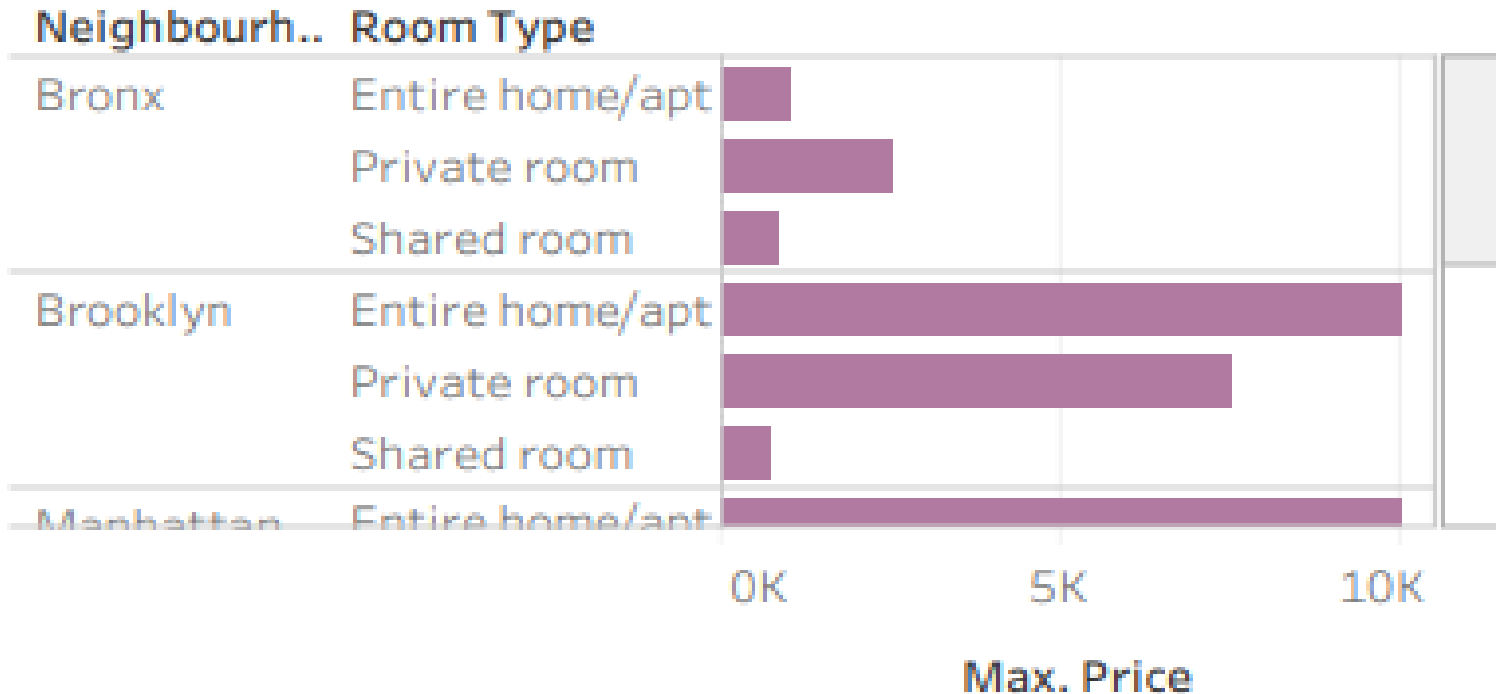
If we compare different Room Category with different cities we will come to know that Private rooms has the most number of reviews followed by Entire Home/Apt and then Shared rooms which means the most popular category was Private room among Hosts.



# Data through Excel



# Room type with respect to price in different neighbourhood

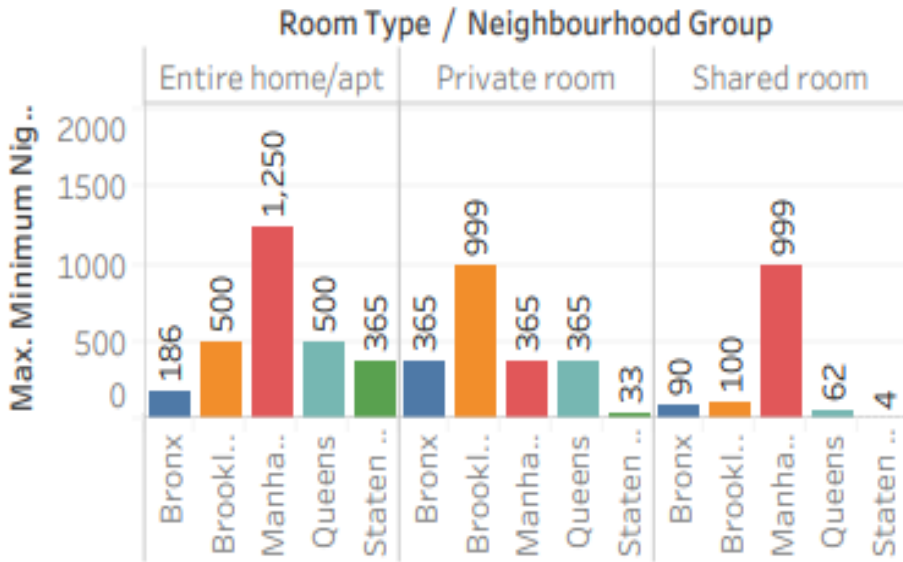


Graph shows that Private room was most expensive in Bronx, but in Brooklyn Entire Home was the most expensive and same was the case in Manhattan.

# Maximum number of nights stayed in which room type

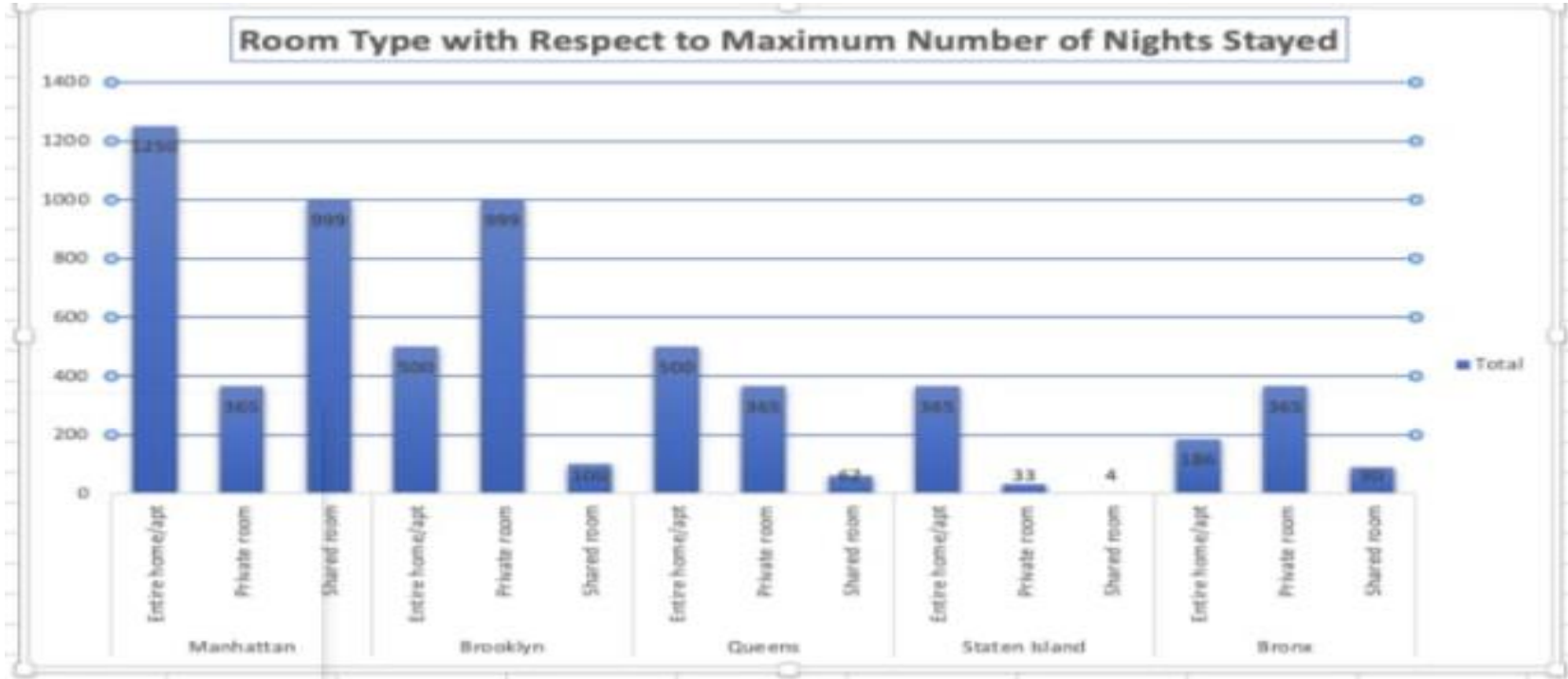


Through Tableau



From this graph we can infer that People had spent most of the time in Entire home/Apt than other two categories of the rooms.

# Data through Excel



# Room type with respect to maximum Price



# Analysis Summary

- ☐ The people who prefer to stay in Entire home or Apartment they are going to stay bit longer in that particular Neighborhood 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 Neighborhood group that means that place is a tourist place.
- ☐ If people are not staying more then one night means they are travelers.

# Thank You

