Data Visualization Excellence: Showing Airbnb Insights

SHITAL PRASHANT JADHAV, INDER MUKHOPADHYAY

Contents :

Introduction :

Airbnb, Inc. is an American company that operates the online marketplace for lodging, vacation rentals, and tourism activities that you described. It provides a platform for hosts to offer short-term accommodations and tourism experiences to guests.

New York City is indeed one of the most diverse and populous cities in the United States. It consists of five boroughs: Manhattan, Brooklyn, Queens, the Bronx, and Staten Island. These boroughs were consolidated into a single city, and each has its own unique characteristics and attractions.

New York City is renowned as a global centre for the financial services industry, hub for media, entertainment , telecommunications, law, and advertising. Additionally, New York City is a popular destination for tourists from around the world, with a wide range of cultural attractions, museums, art galleries, shopping districts, and diverse culinary experiences.

Problem Statement :

- For the past few months, Airbnb has seen a major decline in revenue.
- Now that the restrictions have started lifting and people have started to travel more, Airbnb wants to make sure that it is fully prepared for this change.
- So, analysis has been done on a dataset consisting of various Airbnb listings in New York.

Objective :

- To conduct a through analysis of New York Airbnb Dataset.
- Ask effective questions that can lead to data insights.
- To process data ,analyze and share findings by data visualization and statistical techniques.

Key Insights :

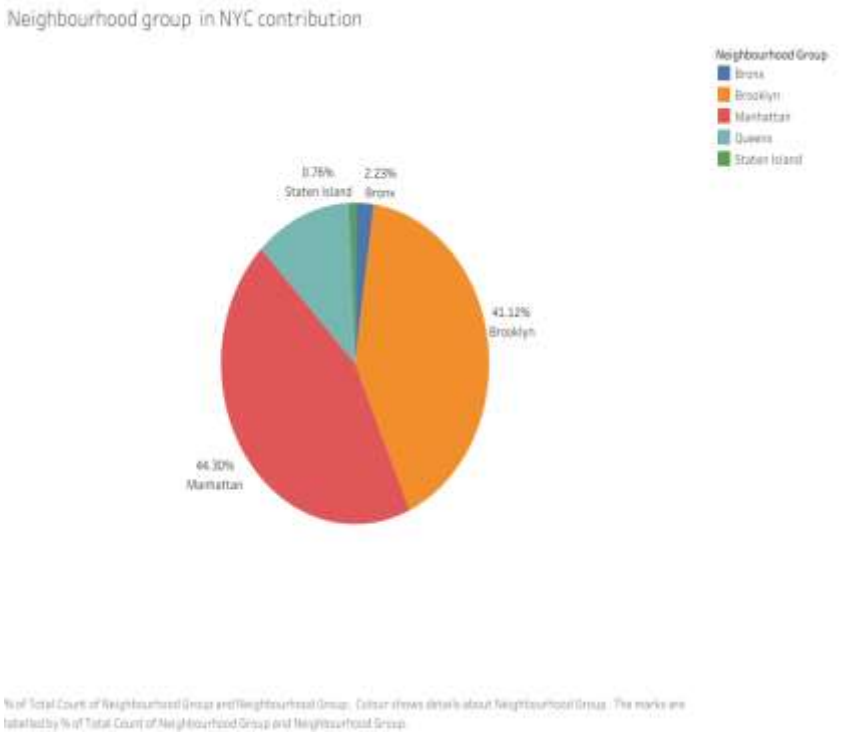
To understand some important insights we have explored the following questions:

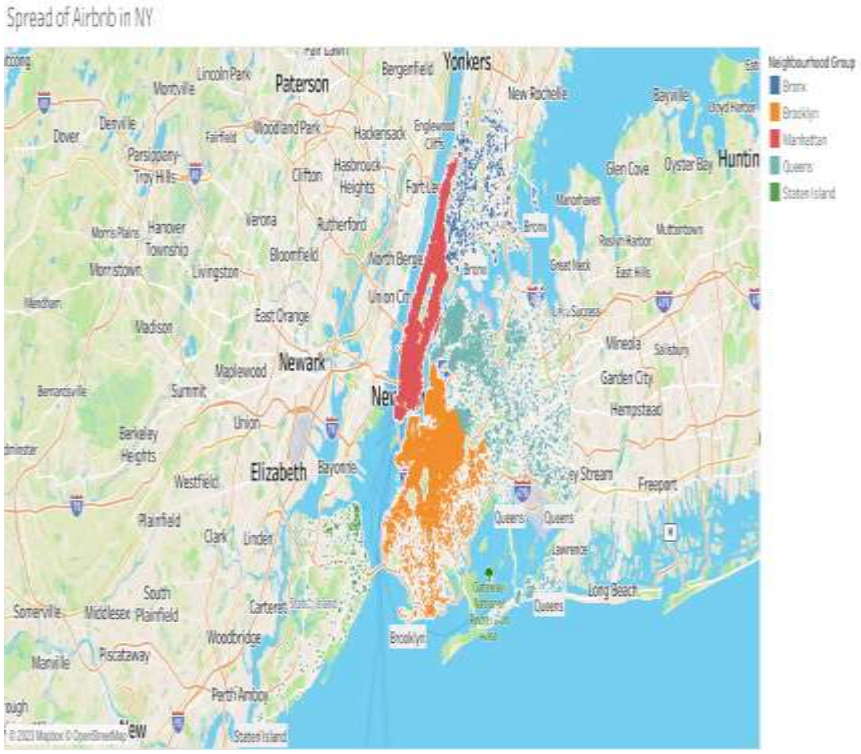
- How are the Airbnb listings spread out in NYC and most contributing neighbourhood.
- What type of rooms do customers prefer
- What could be the ideal number of minimum nights to increase customer bookings .

Based on customer review:

- Most preferred neighbourhood .
- Most preferred room type.
- Who are the Hosts who have the highest listings w.r.t. Neighbourhood.

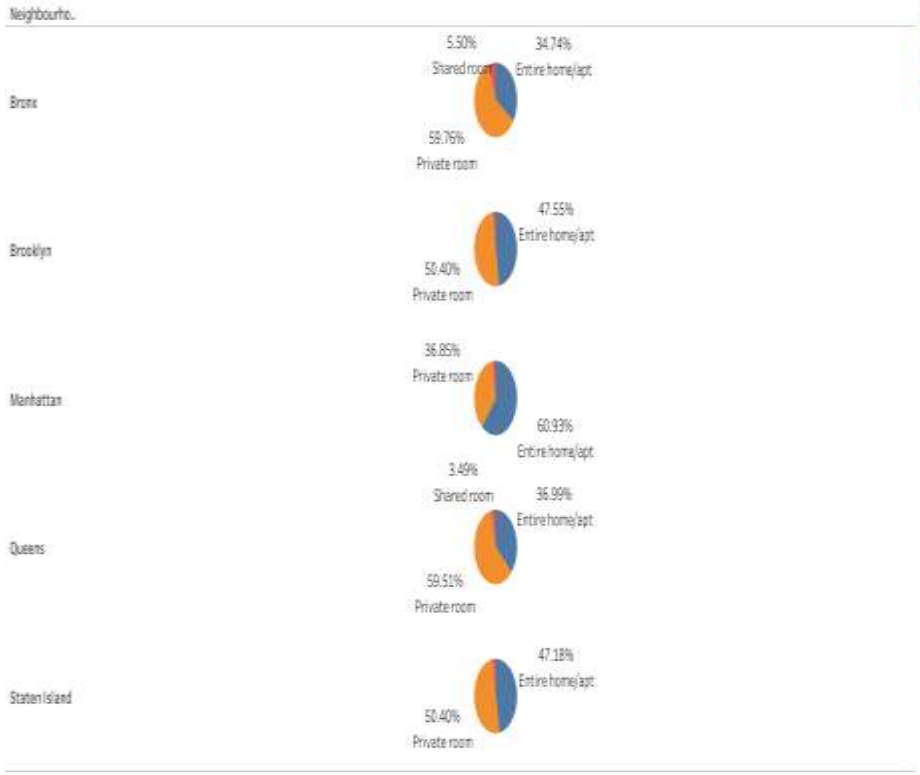
Most Contributing Neighborhoods

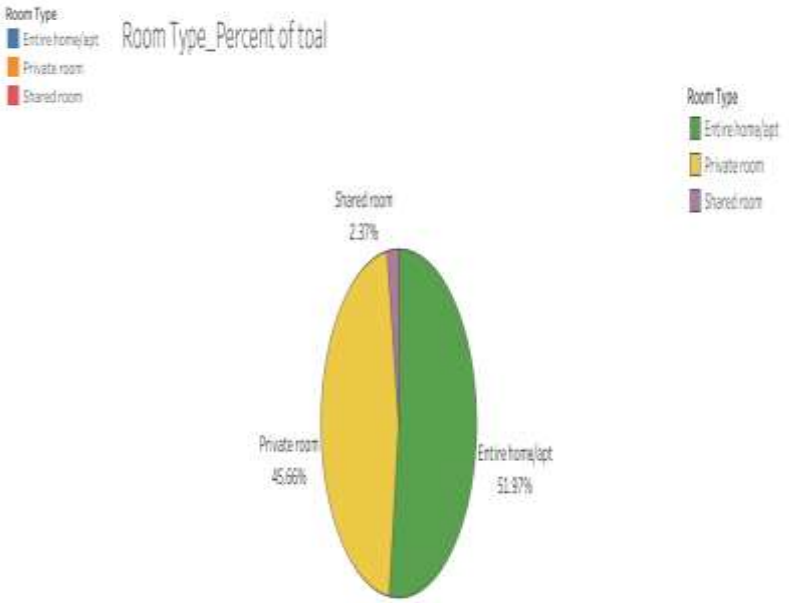




- Manhattan and Brooklyn are the prime hubs for Airbnb listings in New York City, with a significant presence in both neighbourhood groups.
- Listings are maximum in Manhattan (44%) & Brooklyn (41%) neighbourhood group.
- Staten Island has the smallest proportion of listings, representing only around 1% of the total.

Room type with respect to Neighbourhood group

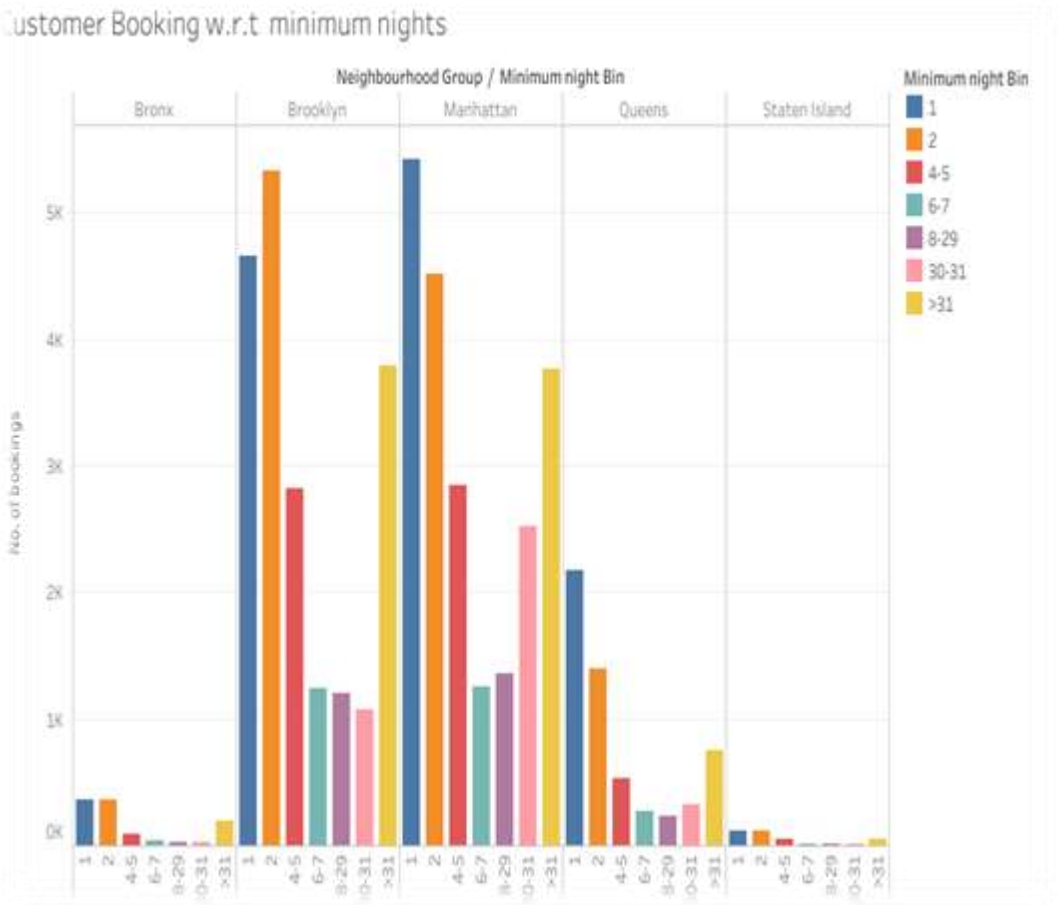




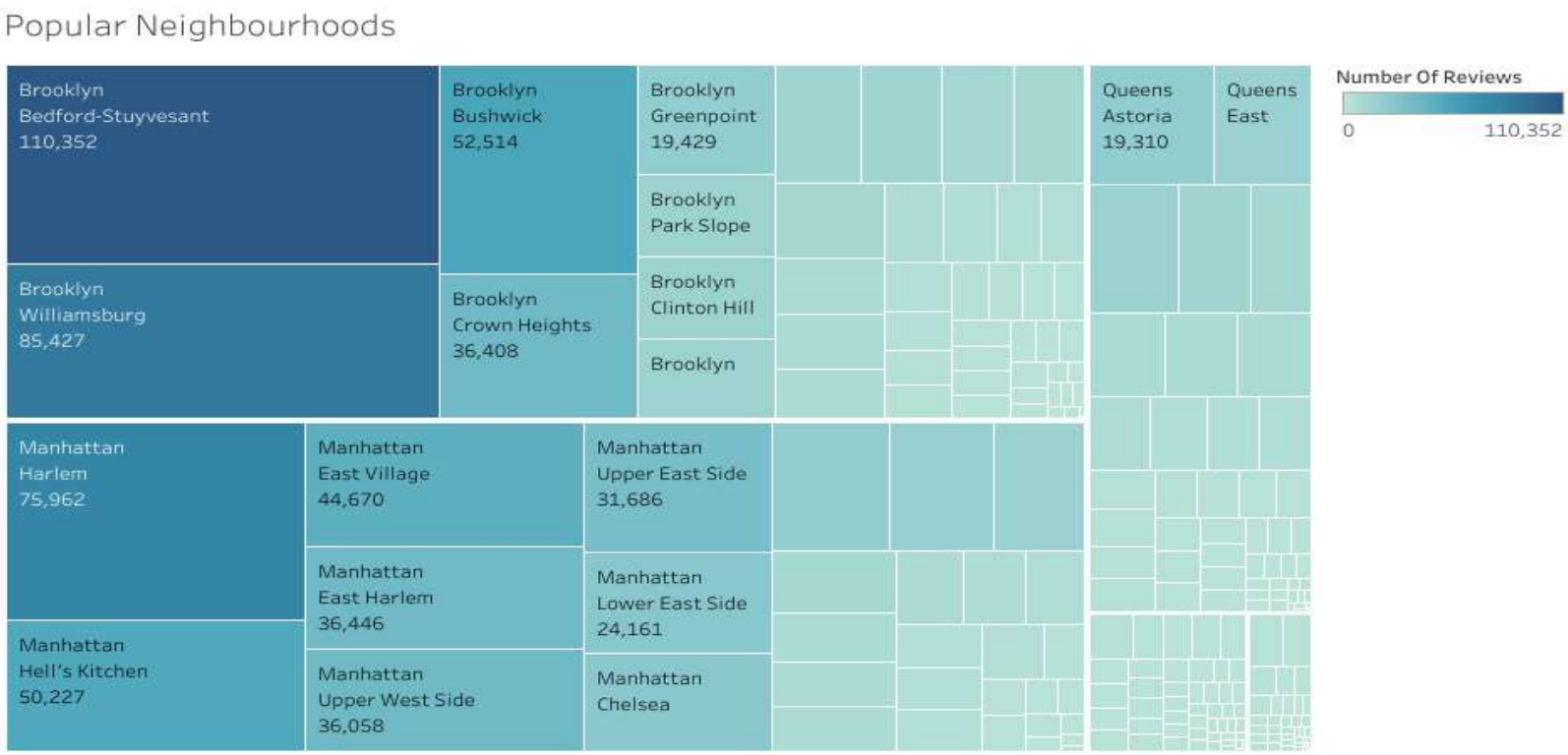
- Airbnb offers three distinct types of accommodations: entire homes/apartments, private rooms, and shared rooms.
- Private rooms are favoured by 45% of guests, while entire homes/apartments are preferred by 52%, leaving shared rooms at a lower 2.4% preference rate.
- To encourage more bookings and attract a wider range of guests, Airbnb can focus on promoting shared rooms by offering discounts and incentives. Additionally, efforts can be made to acquire more private listings to meet the high demand for this accommodation type.
- Queens & Bronx contribute 60% each to private rooms, more than the combined ratio of 45%. Whereas, Manhattan has a higher contribution in entire home (61%), compared to the combined ratio of 52%.

Customer Booking with respect to minimum nights

- The majority of bookings on Airbnb are made for listings with a minimum stay of 1-5 nights, indicating their popularity among guests.
- Notably, there is a significant increase in bookings for 30-day stays, which can be attributed to customers renting accommodations on a monthly basis.
- In terms of specific boroughs, Manhattan and Brooklyn stand out with a higher number of 30-day bookings compared to other areas. The reason could be either tourists booking long stays or mid-level employees who opt for budget bookings due company visits.

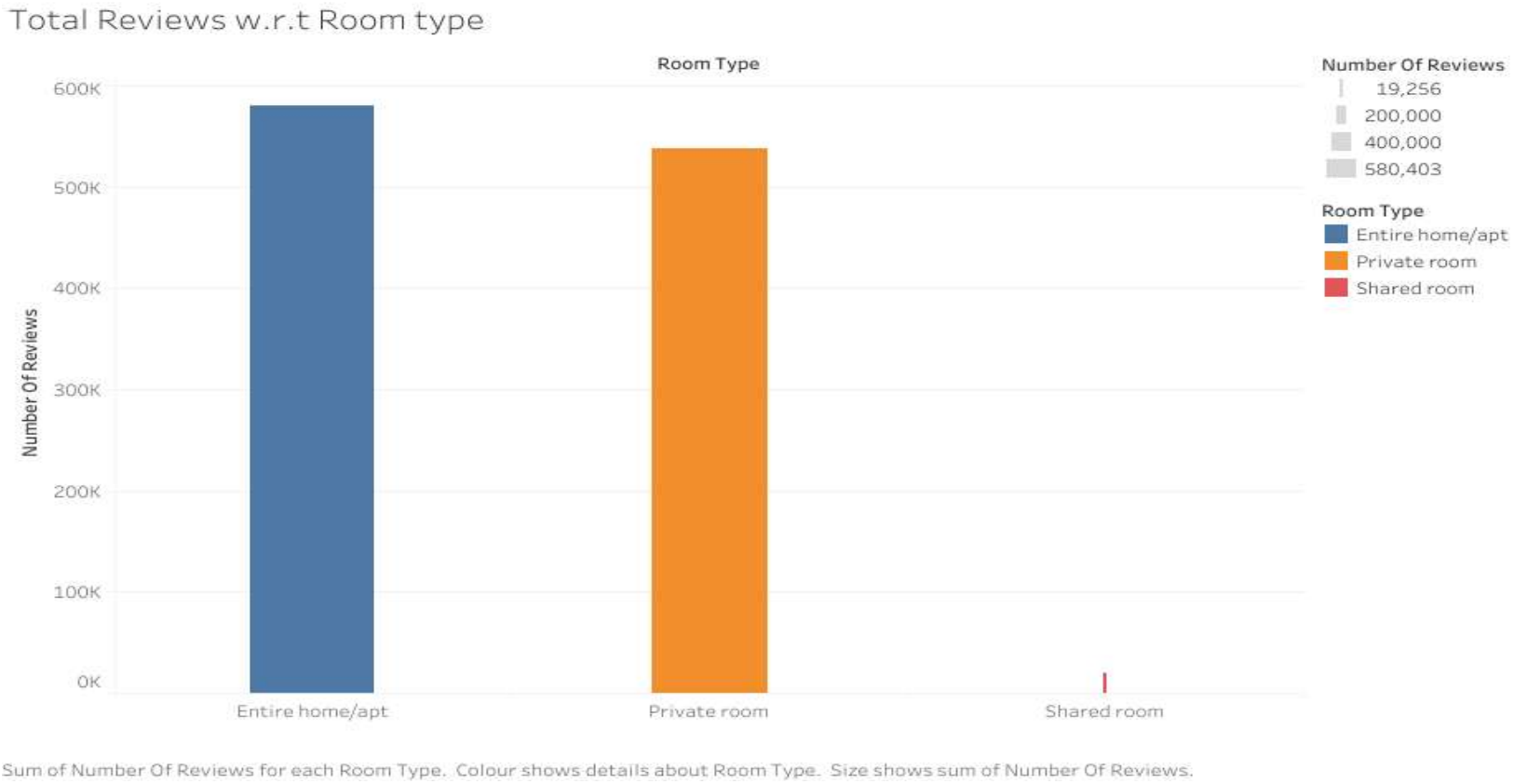


Number of Reviews w.r.t Neighbourhoods from Different Neighbourhood groups



- Customers exploring the neighbourhoods of Brooklyn and Manhattan are inclined to offer their feedback.
- Encouraging customers who visit the vibrant neighbourhoods of Brooklyn and Manhattan to share their feedback is essential in order to facilitate continuous improvement and share valuable insights with other listings.

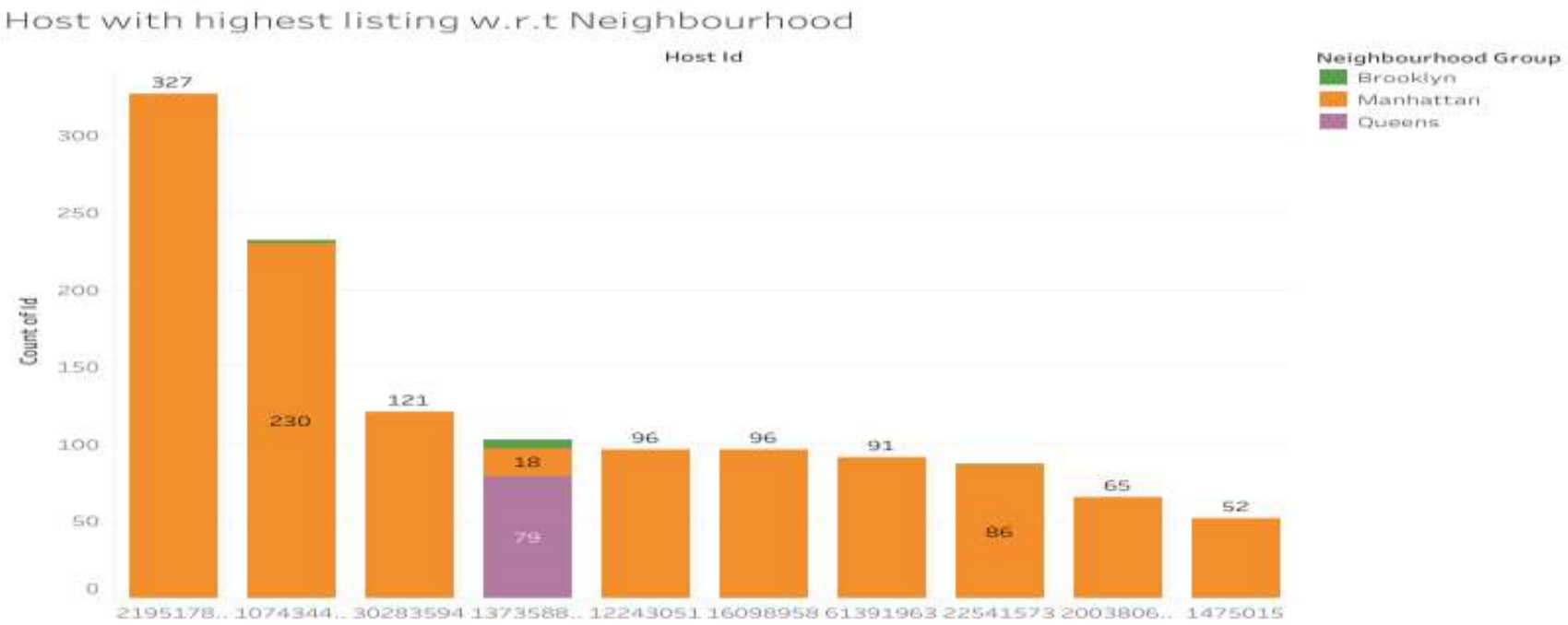
Total Reviews w.r.t Room Type



Sum of Number Of Reviews for each Room Type. Colour shows details about Room Type. Size shows sum of Number Of Reviews.

- Based on the maximum number of reviews received, it can be inferred that customers tend to favor the 'Entire home/apt' and 'Private rooms' options over 'Shared rooms'.

Host with highest listing w.r.t Neighbourhood



- An interesting observation is the presence of a single host managing multiple listings, particularly in the Manhattan area. This trend can be attributed to the fact that Manhattan attracts a significant number of tourists.
- Overall, the strategic decision of experienced hosts to focus on the Manhattan area stems from the high volume of tourists and financial enthusiasts it attracts, making it a profitable and sought-after location for short-term rentals.

Appendix- Data sources

The dataset includes columns that are clearly labelled and their meanings can be easily understood. Please refer to the accompanying diagram provided below.

Column	Description
id	listing ID
name	name of the listing
host_id	host ID
host_name	name of the host
neighbourhood_group	location
neighbourhood	area
latitude	latitude coordinates
longitude	longitude coordinates
room_type	listing space type
price	
minimum_nights	amount of nights minimum
number_of_reviews	number of reviews
last_review	latest review
reviews_per_month	number of reviews per month
calculated_host_listings_count	amount of listing per host
availability_365	number of days when listing is available for booking

Appendix – Data Methodology

- Cleaned the data set using python.
- The dataset was analyzed using both univariate and bivariate analysis techniques. Tableau was utilized to conduct the analysis and create visualizations, considering various parameters. The key parameters considered for analysis include geography-based bookings, room type-based bookings, number of reviews, and minimum number of nights. Inferences were drawn based on these parameters to gain insights from the data.
- Deriving features Applied statistical methods and created visualizations using Tableau.

Appendix – Data model Assumptions:

```
Categorical Variables:
    - room_type
    - neighbourhood_group
    - neighbourhood

Continous Variables(Numerical):
    - Price
    - minimum_nights
    - number_of_reviews
    - reviews_per_month
    - calculated_host_listings_count
    - availability_365
- Continous Variables could be binned in to groups too

Location Varibles:
    - latitude
    - longitude

Time Varibale:
    - last_review
```