EDA Capstone Project- 1

EDA ON AIRBNB NYC

Team Artistic Algorithm:

Chandan Baraliya
Bhupendra Singh

Highlighting Airbnb Business

- What Is Airbnb? Airbnb, as in "Air Bed and Breakfast," is a service that lets property owners rent out their spaces to travelers looking for a place to stay. Travelers can rent a space for multiple people to share, a shared space with private rooms, or the entire property for themselves.

- Since 2008, guests and hosts have used Airbnb to expand on traveling possibilities and present a more unique, personalized way of experiencing the world.

- Today, Airbnb became one of a kind service that is used and recognized by the whole world.

- Airbnb is based on a peer-to-peer business model. This makes it simple, easy to use, and tends to be more profitable for both parties. The model also gives you the opportunity to customize and personalize your guests' experience the way you want

Problem Statement

- For this project we have analyze Airbnb's New York City(NYC) data of 2019. NYC is not only the most famous city in the world but also top global destination for visitors drawn to its museums, entertainment, restaurants and commerce.

- Our main objective is to find out the key metrics that influence the listing of properties on the platform. For this, we will explore and visualize the dataset from Airbnb in NYC using basic exploratory data analysis (EDA) techniques.

- Data analysis on thousands of listings provided through Airbnb is a crucial factor for the company.

- We will be finding out the distribution of every Airbnb listing based on their location in NYC, including their price range, room type, listing name, and other related factors.

Understanding of Data given

- **Id-** These gives us the listing id.
- **Name-** listing name
- **Host_id-** host id
- **Host_name-** host name
- **Neighbourhood_group-** NYC borough
- **Neighbourhood-** NYC neighborhood
- **Latitude-** listing latitude
- **Longitude-** listing longitude
- **Room_type-** type of rooms (Entire home/apt, Private room, Shared room)
- **Price -** listing price
- **Minimum_nights-** minimum nights people stay
- **Number_of_reviews-** Total number of reviews
- **Last_review-** Date of last review
- **Reviews_per_month-** Average number of reviews per month
- **Calculated_host_listing_count-**total number of listings for this host
- **availabilty_365-**number of days listing is available.
- **Number of Data set were nearly 49000.**

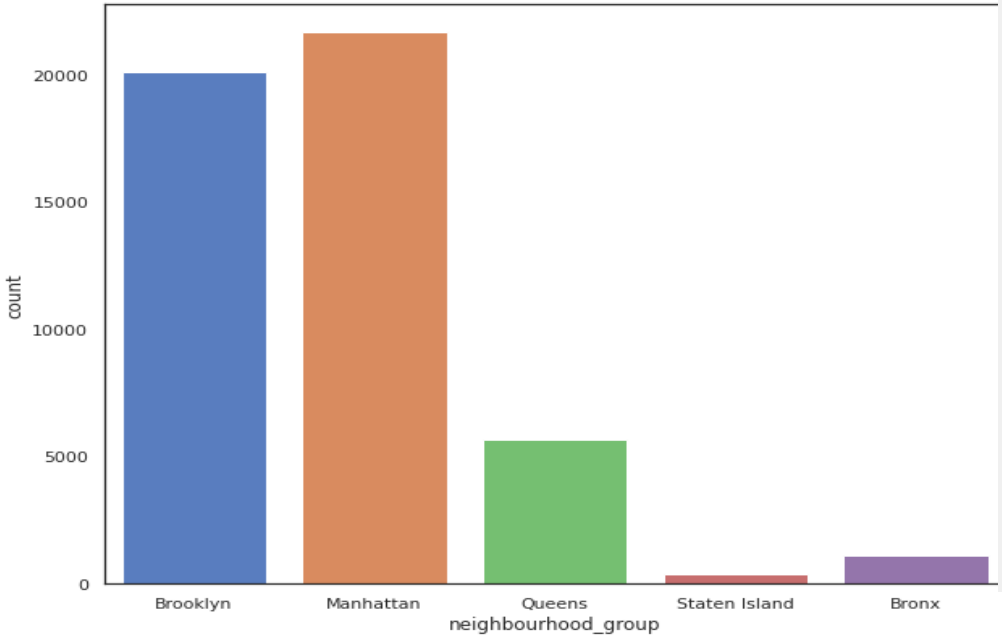
Agenda

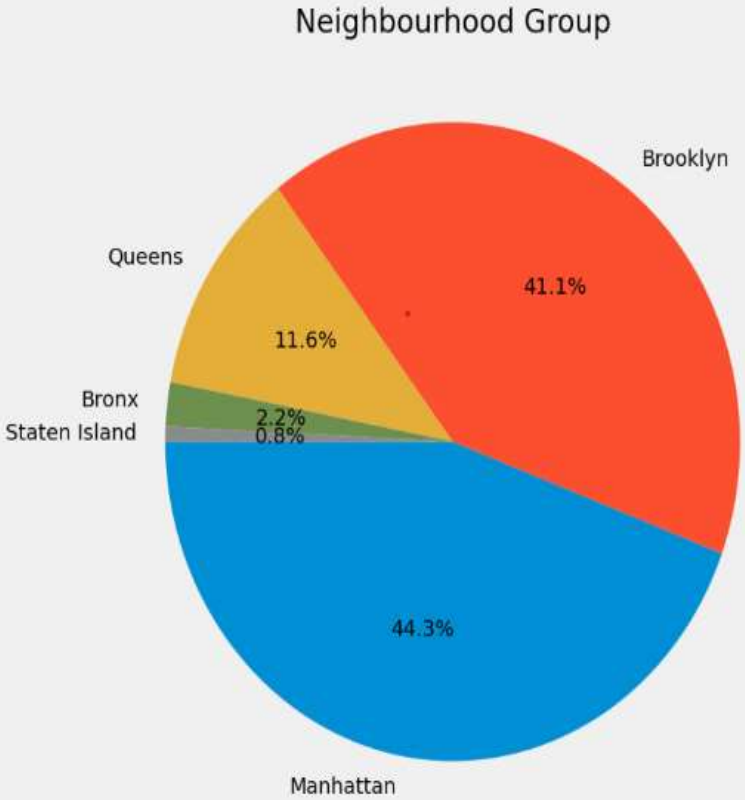
➔ We try to answer following questions for Airbnb:

- What is the average preferred price of neighbourhood group to the location?
- What is the share of room types in neighbourhood group
- Where the customer pays the highest and lowest rent according to location?
- Most preferred room types by customer.
- Find the total count of each room type
- Room types and their relationship with availability in different neighbourhood groups?
- Find top 10 hosts with most listings
- Find top 10 hosts with most reviews

Various Neighbourhood group and its listing insights

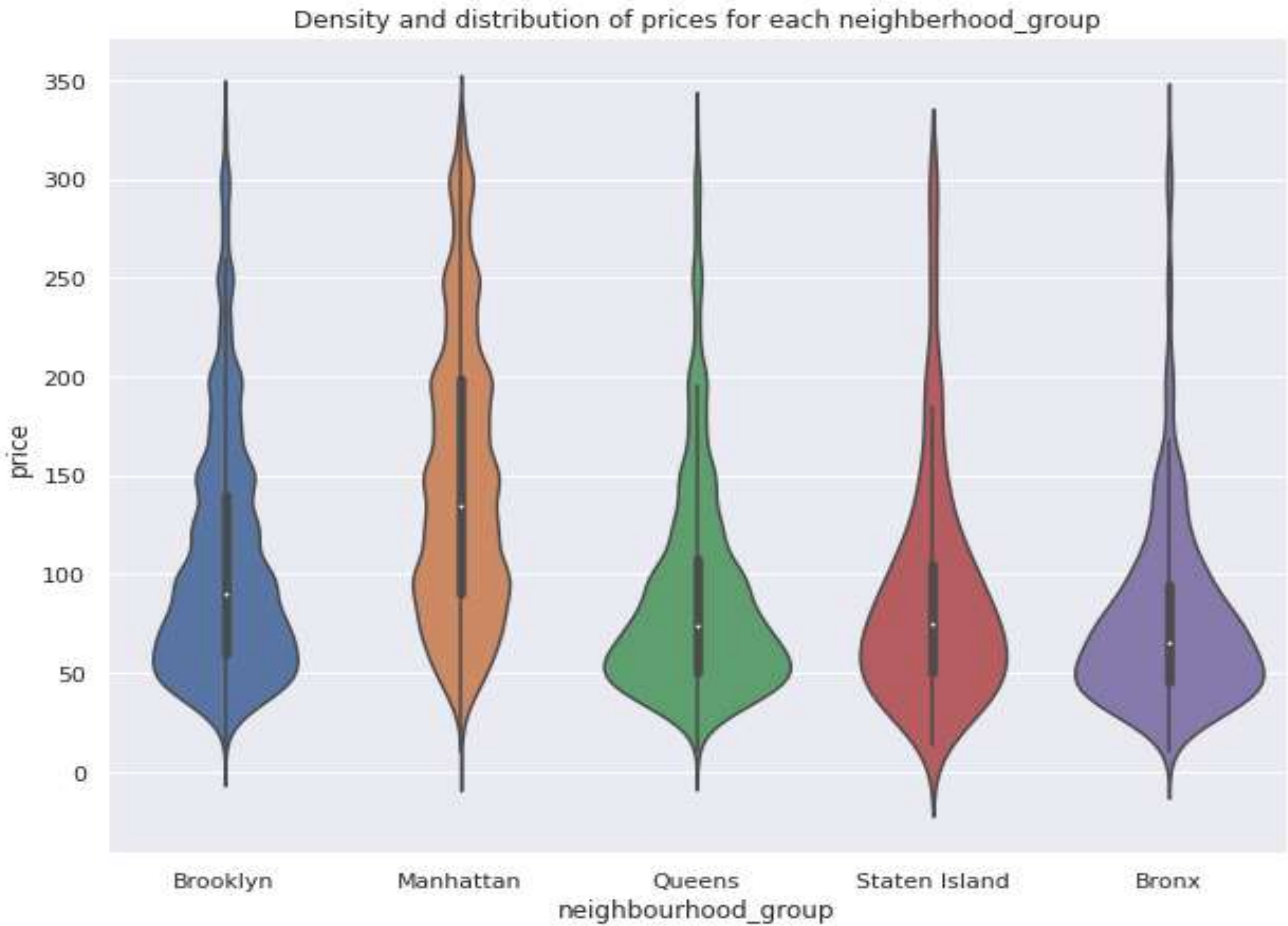
- **Manhattan has the highest number of listing of about 44.3 % followed by Brooklyn of 41%.**
- **State Island stands the least number of listing less the 1 %.**





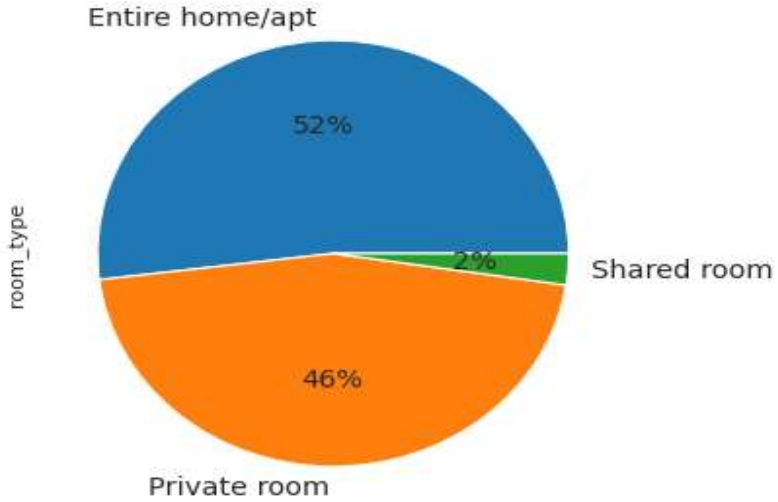
What is the average price for each neighborhood group

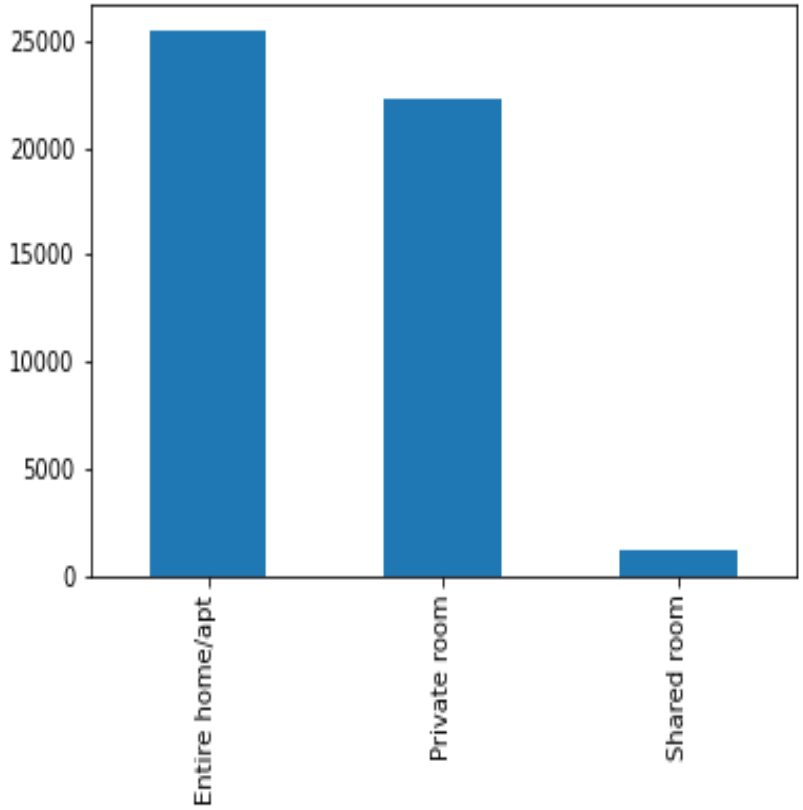
- As usual Manhattan being the costliest place to live in NYC, having average price more than 140 USD followed by Brooklyn with around 80 USD on an average for the listings.
- Queens and Staten Island has nearly the same of 75 USD.
- The highest price range could go just above 360 USD.



What is the share of various room types

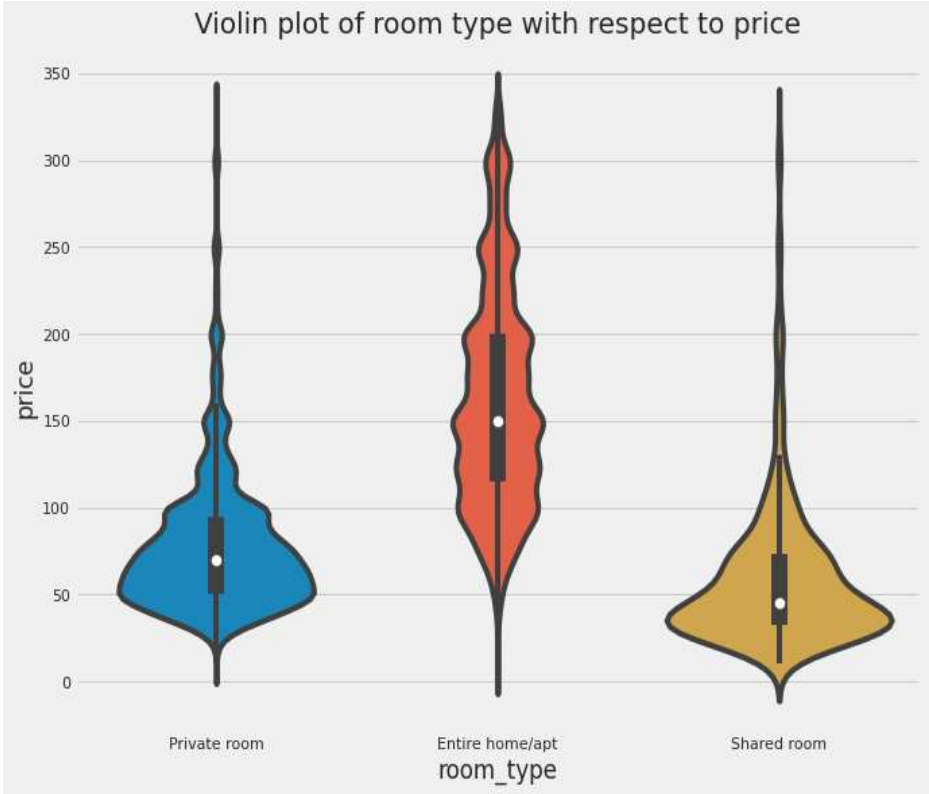
- **The entire home/apt room are the highest in NYC followed by private rooms.**
- **Shared room holds are only 2 % in Airbnb NYC.**



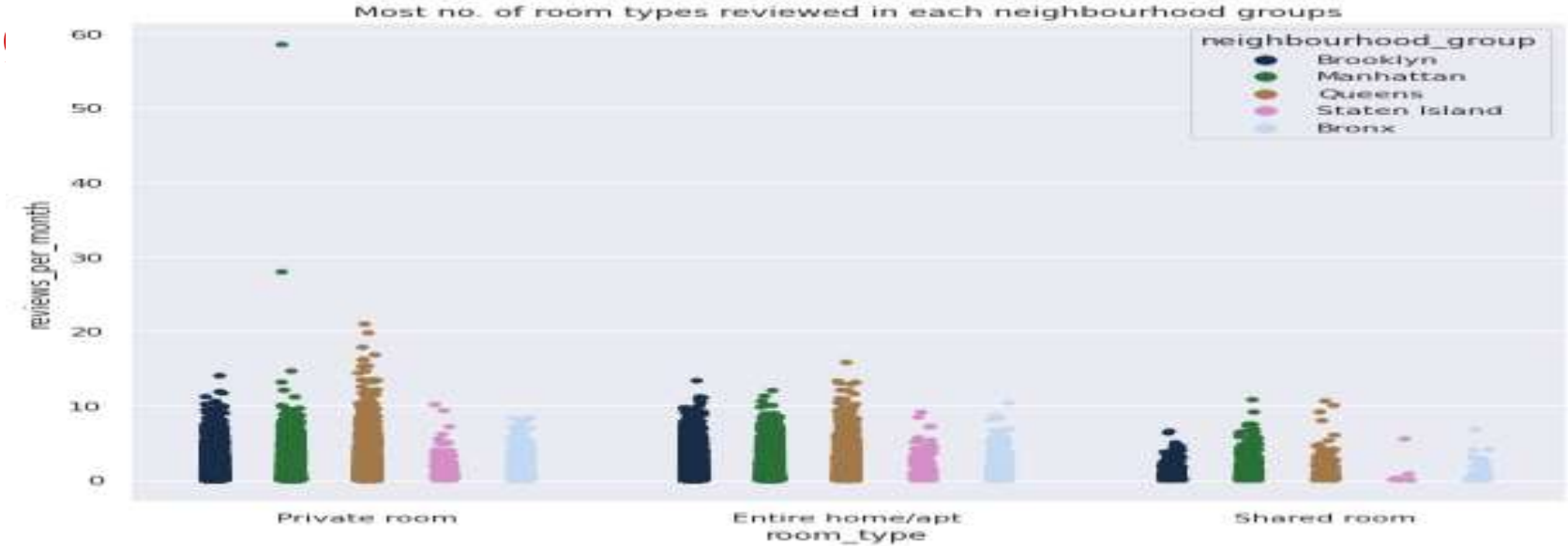


What is average cost of stay in different room types

- **Entire Home may averagely cost 150 USD .**
- **Private room average cost is of around 65 USD.**
- **Shared rooms cost the least price of less then 50 USD.**
- **So, lets check what people prefer to live in with review counts.**

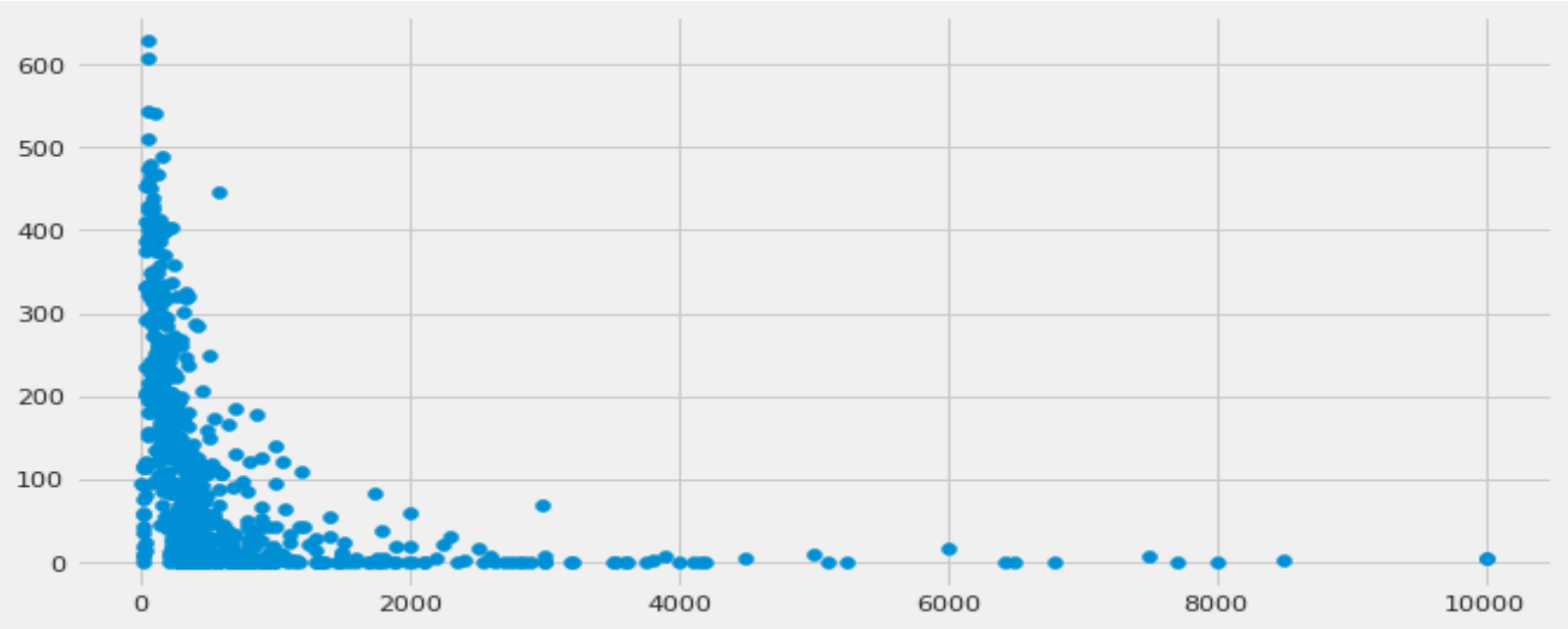


What room types are prefered in each neighbourhood



- **Here we can conclude that Private rooms in Manhattan Neighborhood group had the highest number of reviews going 50+ reviews per month followed by the chase.**
- **After Manhattan, Queens had the highest number of reviews in all the room types.**
- **There were less reviews received from shared rooms as compared to other room types and it was from Staten Island followed by Bronx.**

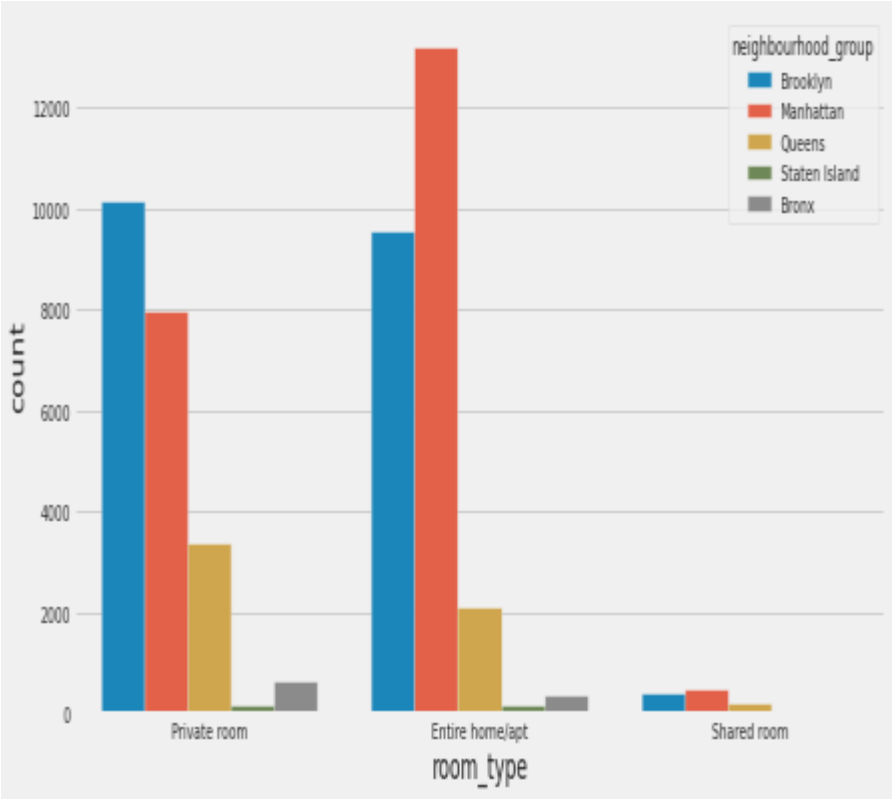
Price To review Analysis



Price on X axis and Reviews on Y axis.
It Says us that people like more to stay where price is less.

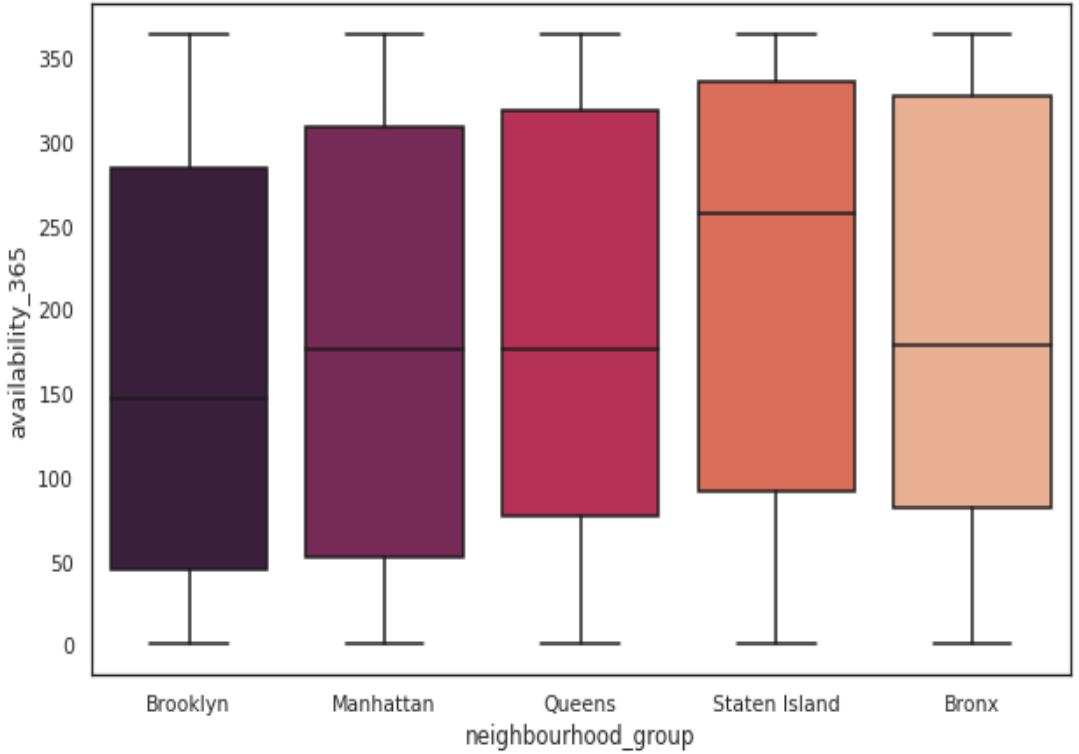
How much room types are in neighbourhood group

- **Manhattan has the highest number of Entire home/apt.**
- **Brooklyn holds the top share of private rooms.**
- **State island has an equal percentage of Private rooms and Entire home.**
- **Also, we see that State island has a smaller number of room types as compared to another neighbourhood. Which matches our neighbourhood group to listing counts.**

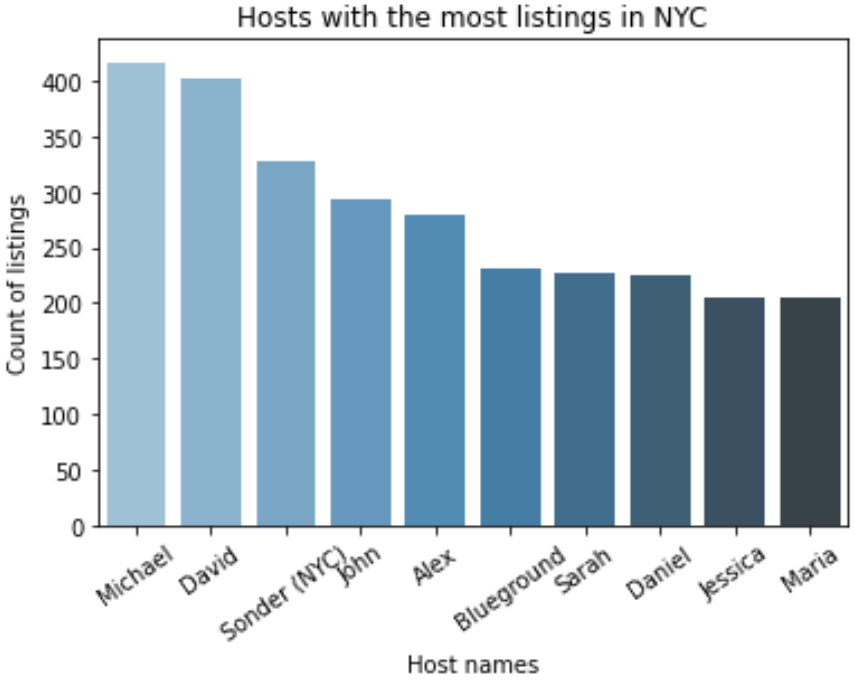


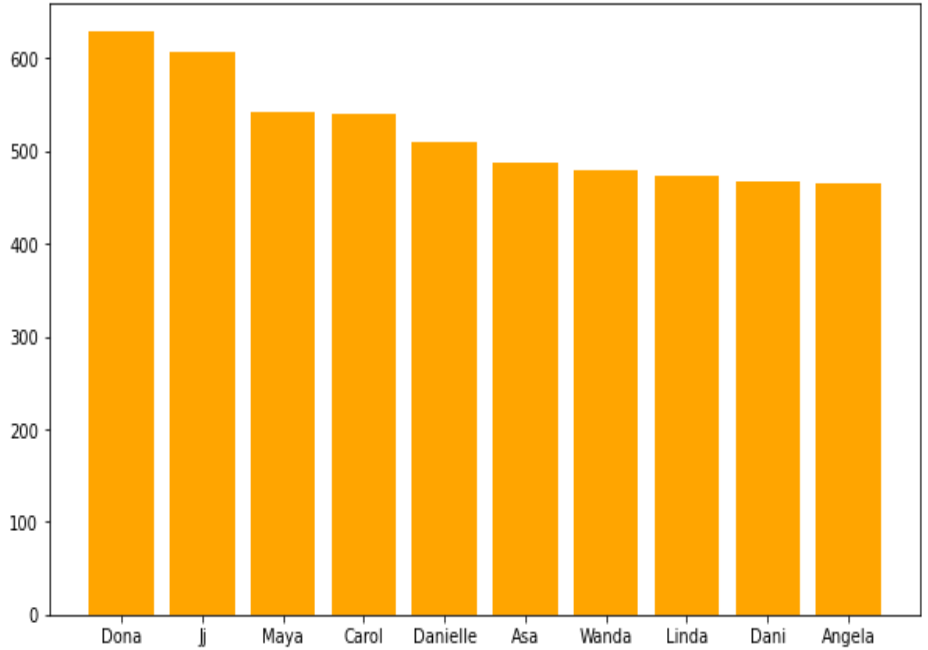
How much is avability of rooms at neighbourhood group

- **Staten Island has the most avability of rooms all the year.**
- **Brooklyn having the least availability of rooms through a year with the second most count listing gives an opportunity to have an increase in number of rooms.**
- **Manhattan, Bronx and Queens and nearly an equal avability of rooms throughout the year.**



Top 10 Hosts with listing counts and reviews



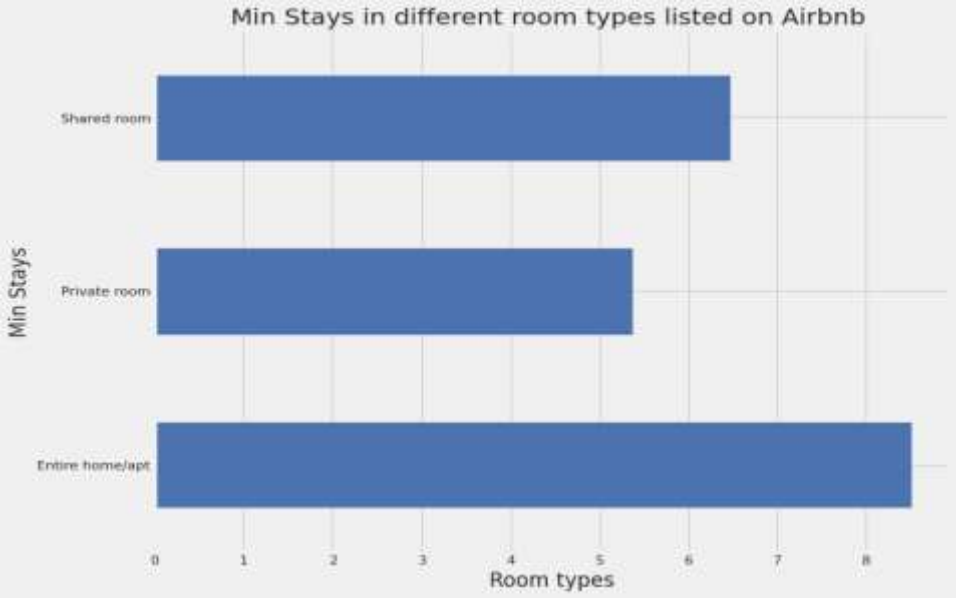


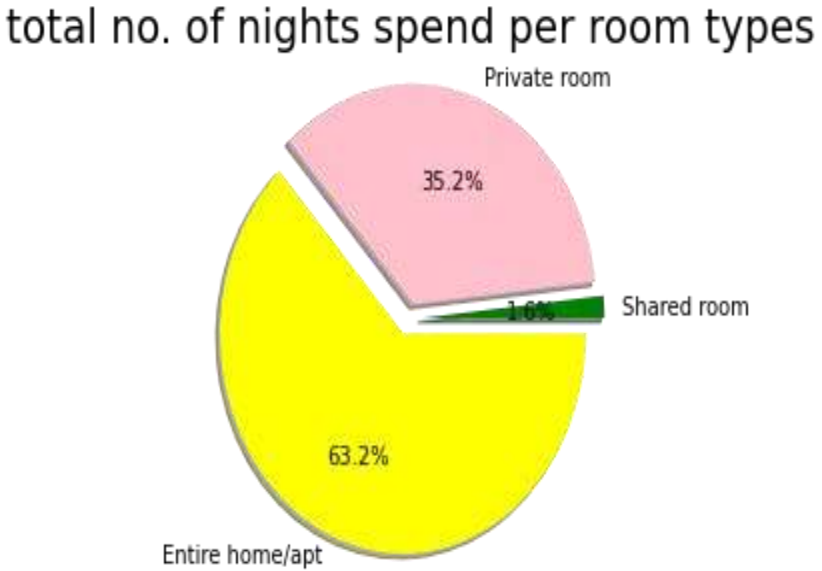
i) Dona, Ji, Maya, Carol, Danielle has the maximum number of reviews.

ii) While Micheal, David, Sonder(NYC), John has the highest number of listing counts.

The being the top faciltors and busiest host.

Total no. of nights spend per room types





- **Here we can state that in which room type customers prefer for night stay.**
- **From pie chart we can conclude that 63.2% customers loved to stay in apartment for maximum durations**
- **Only 1.6% customers spend night in shared room.**

Challenges Faced

- Reading the dataset and understanding of columns.
- For answering some of the questions we had to understand the business model of airbnb that how they work.
- Handling NaN values,null values and duplicates.
- Designing multiple visualizations to summarize the information in the dataset and successfully communicate the results and trends to the reader.
- Removing the outliers for some data set. Finding and sorting few impossible dataset.

Analysis Summary:

We defined some points which can help airbnb in their business:

Manhattan is the most focused place in New York for hosts to do their business

- **As people loved to stay when in need for maximum number of nights in Entire Apartment and Brooklyn stands second highest focused place by people. Also, as the average cost in Brooklyn is 80 USD which is less then Manhattan Airbnb can increase number of entire apartment in Brooklyn.**

Analysis Summary:

We defined some points which can help airbnb in their business:

- Top most place in nighbourhood is williamsburg and Bedford.
- Top rated host as per number of reviews are Dona and jj.
- The avability of rooms in a year is highest in Queens Island and lowest in Brooklyn.
- So it is a great opportunity to increase some Entire apartment in Brooklyn.

Team-Artistic Algorithm