# Capstone Project
Airbnb Bookings Analysis
By - Amrit Keshri

Airbnb Background

Airbnb was established in 2008 when two Hosts welcomed three guests to their San Francisco home, and has since grown to 4 million Hosts who have welcomed more than 900 million guest arrivals in almost every country across the globe. Every day, Hosts offer one-of-a-kind stays and unique Experiences that make it possible for guests to experience the world in a more authentic, connected way.

PEOPLE + PLACES + LOVE + AIRBNB =
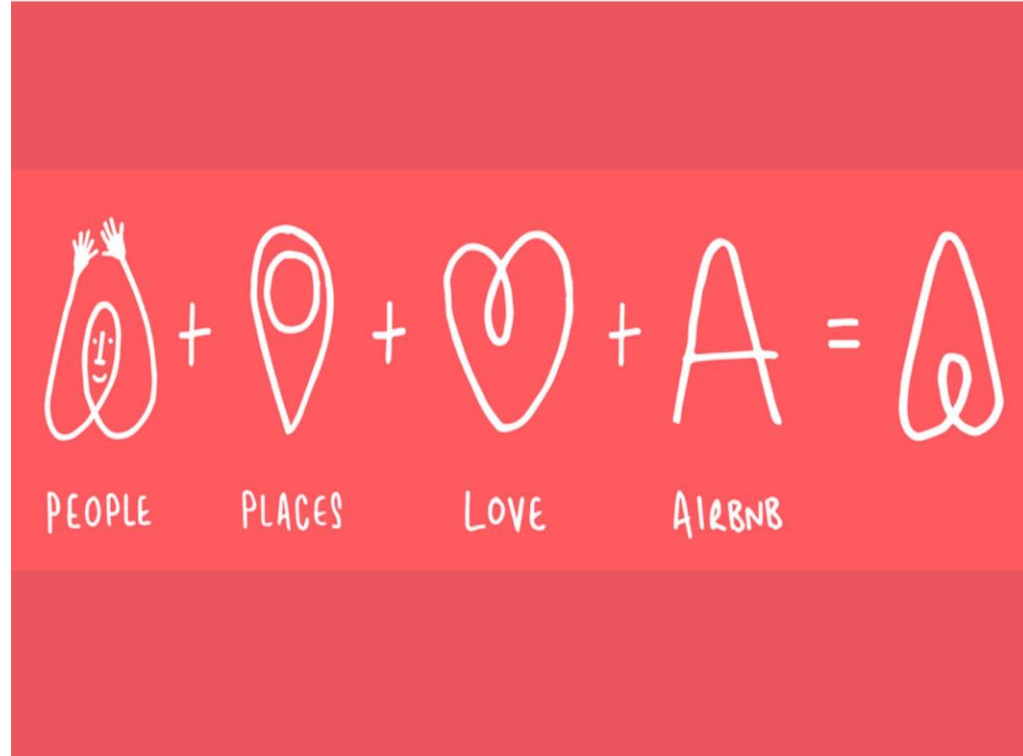
Objective

Understanding the Problem **:- This is the most important part as to understand each and every variable and their correlation with others.**

Data cleaning **:- In this step I've removed unnecessary data. Since there were many NAN values present in our data.**

Fixing the outliers **:- Examined the data for unusual observations and replaced it meaningful values for better analysis.**

EDA **:- In this part I've done Exploratory Data Analysis and tried to find conclusion , correlation and trends from the data.**

Data Summary

The dataset comprises of :-

- **Host Name**
- **Neighbourhood and Neighbourhood Group**
- **Longitude and Latitude**
- **Room Type**
- **Price**
- **Reviews and Availability**

A quick glance at the data shows that there are:

- **47905 unique listing in NYC in total.**
- **Over 1 million reviews have been written by guests.**
- **The price for a listing ranges from $1 0 per night to $1 0,000 per night. Listing with $1 0,000 price tag are in Brooklyn and Manhattan.**
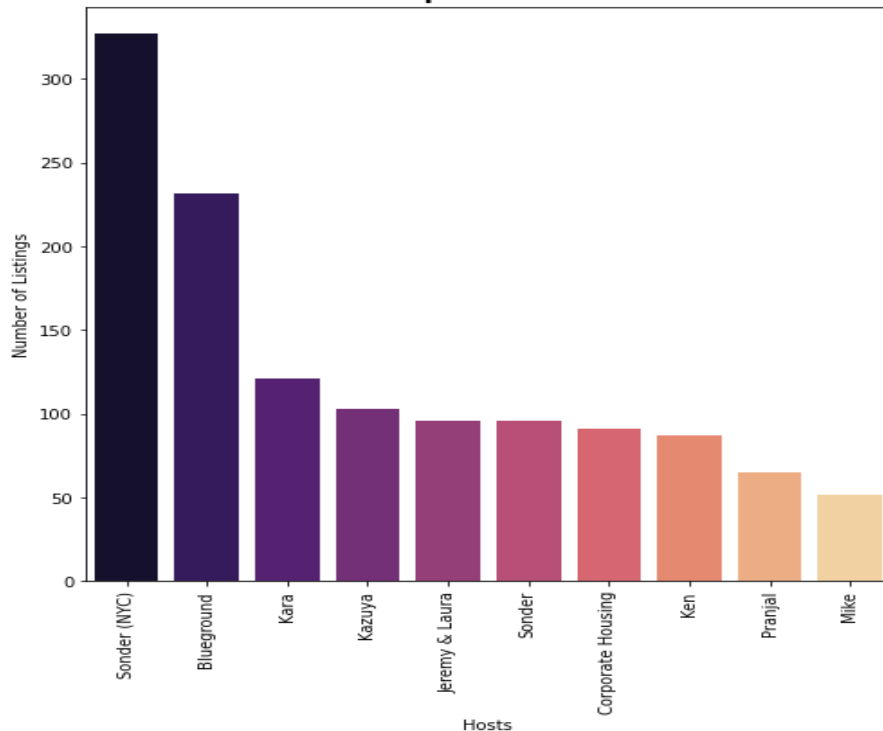
AI

EDA (Posting Analysis)

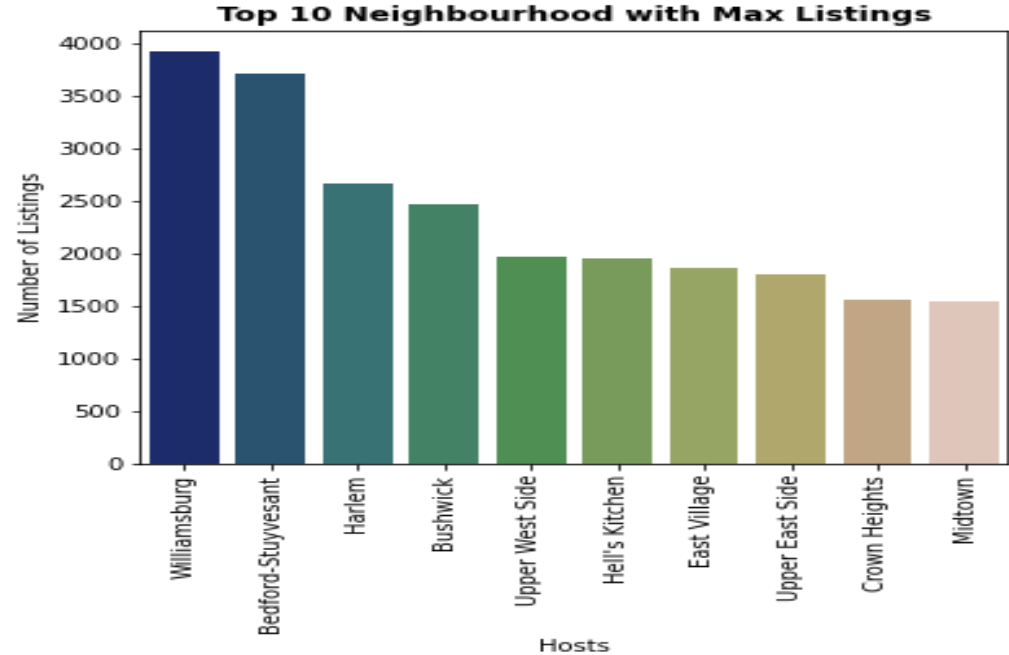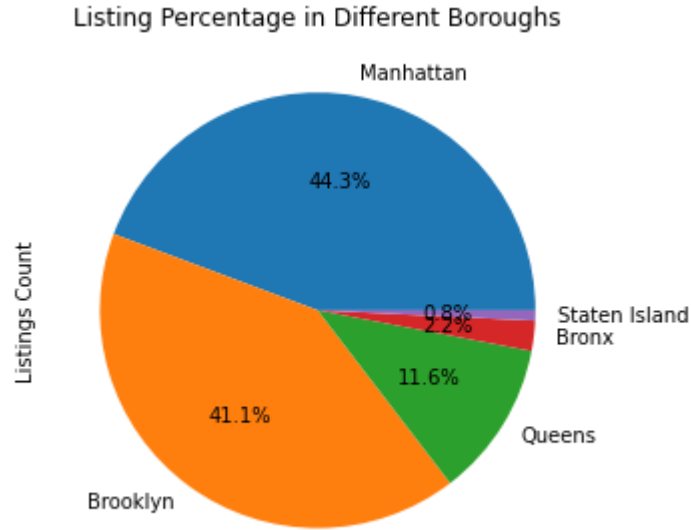**AI**

Number of postings and user distribution
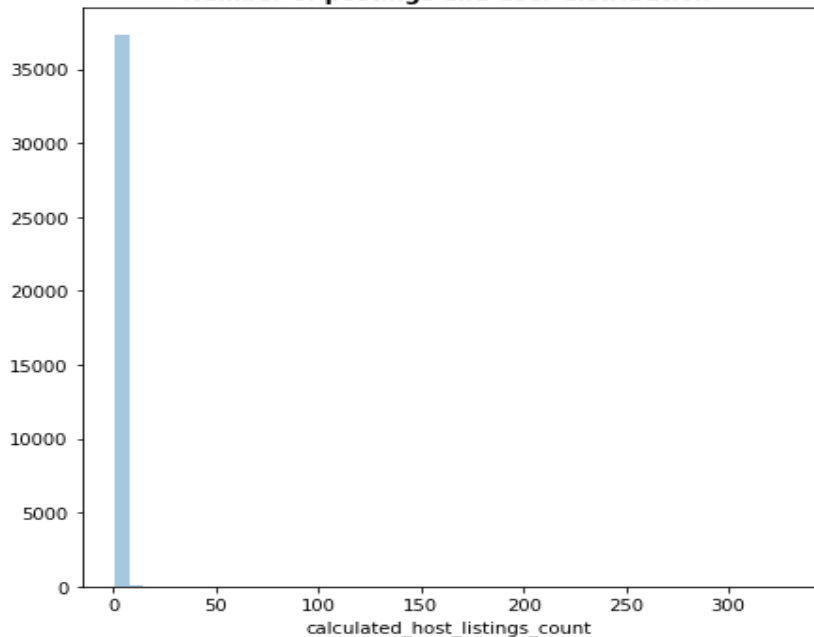
Top 10 Hosts

Most of the hosts have only one listing.

Sonder(NYC) and Blueground are the top hosts with most Number of the listings.

EDA

Listing Percentage in Different Boroughs



Top 10 Neighbourhood with Max Listings

Manhattan , Brooklyn and their Neighbourhood has the most number of listings which is obvious as these cities are the busiest among all others
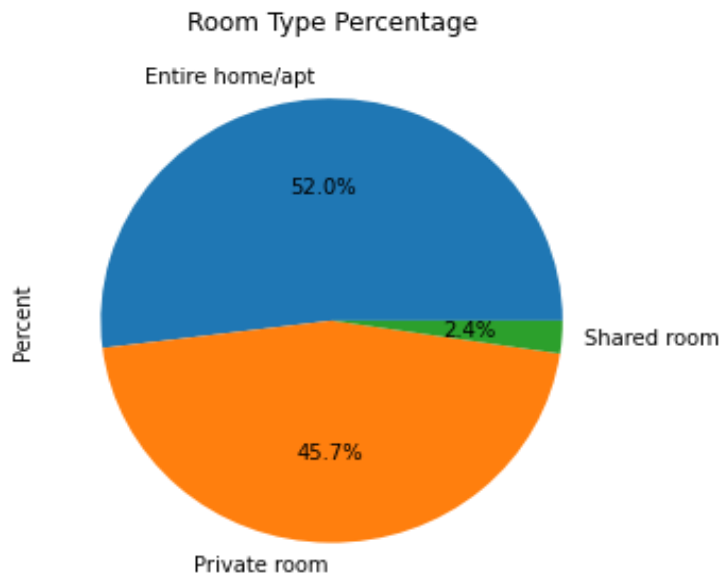
EDA

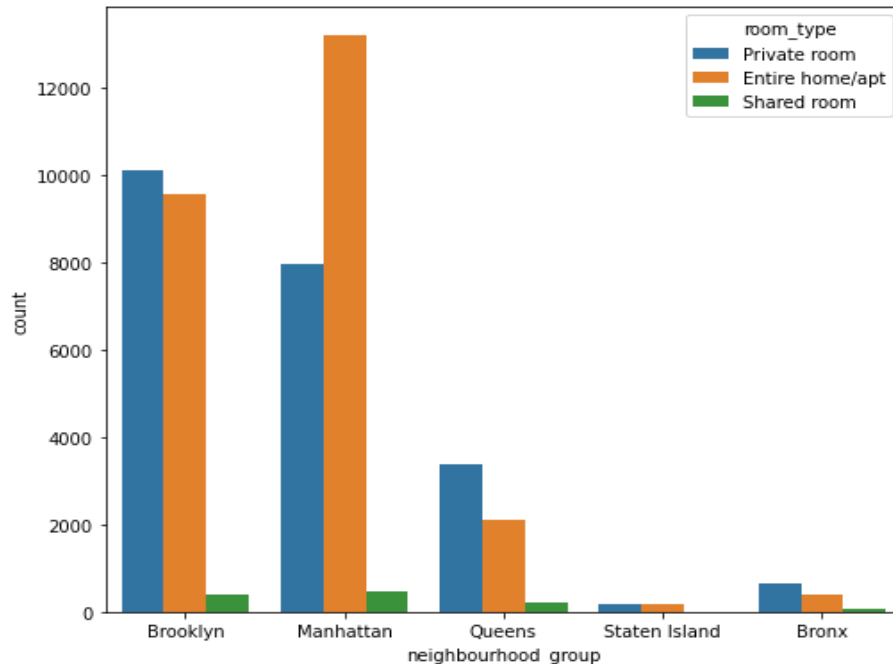**Top hosts have their listing clustered in the most traveled destinations.**

The more rooms a user has, the more they tend to have clustered properties, maybe because they buy or rent the apartments / rooms in bulk.

EDA

Room Type Percentage
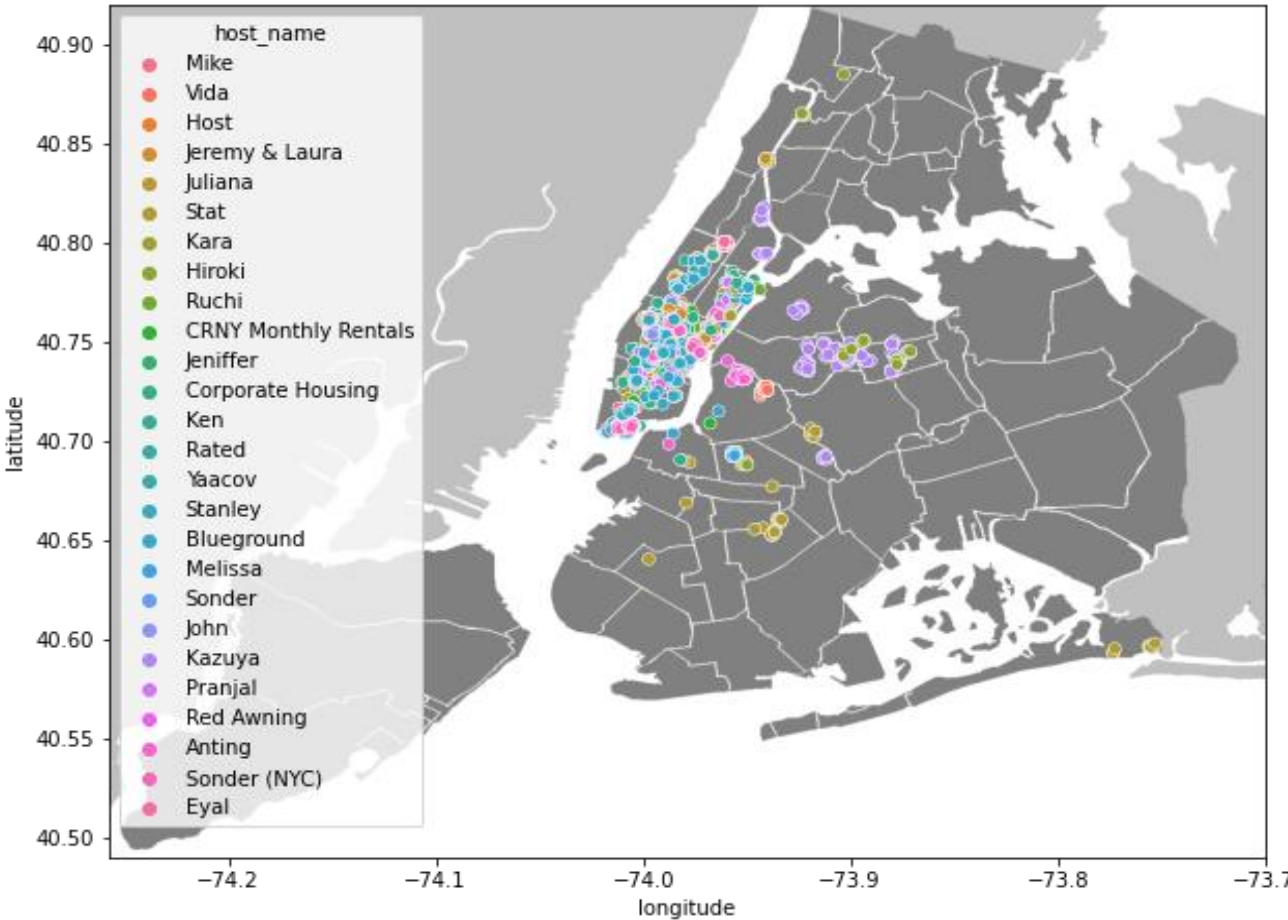
Entire Home /Apt and Private Room has the most contribution among all Room Types.

Brooklyn has the most Number of Private Rooms and Manhattan has the most number of Entire home / Apt. Shared Rooms Distribution is almost same for Brooklyn and Manhattan.

EDA

AI

Even though David , Michael and Alex have comparatively less listing count they are the top Busiest Hosts.
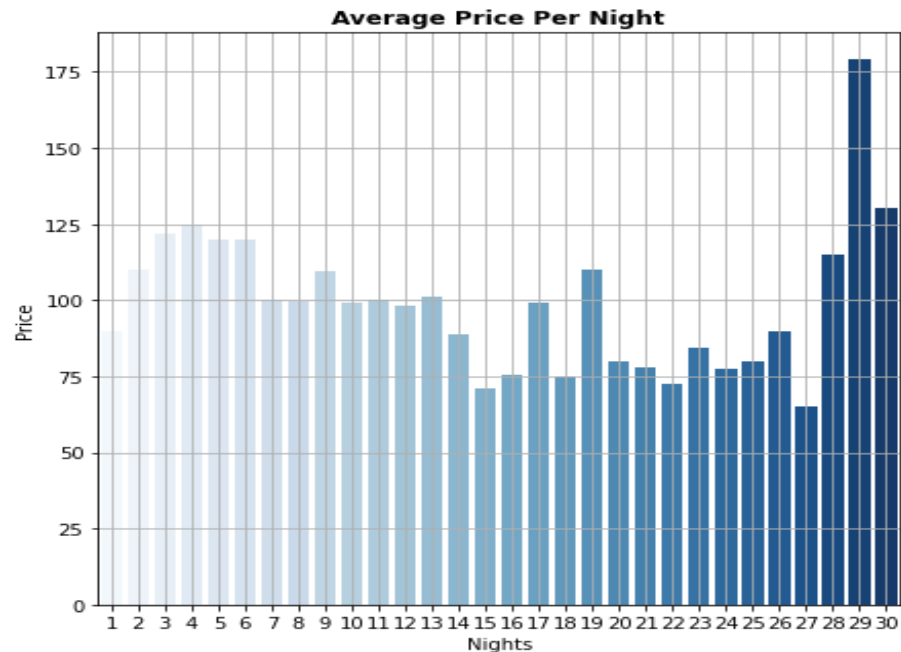
This can be because the listings under these hosts have a great value for money.

Top 20 Busiest Hosts

EDA (Pricing Analysis)
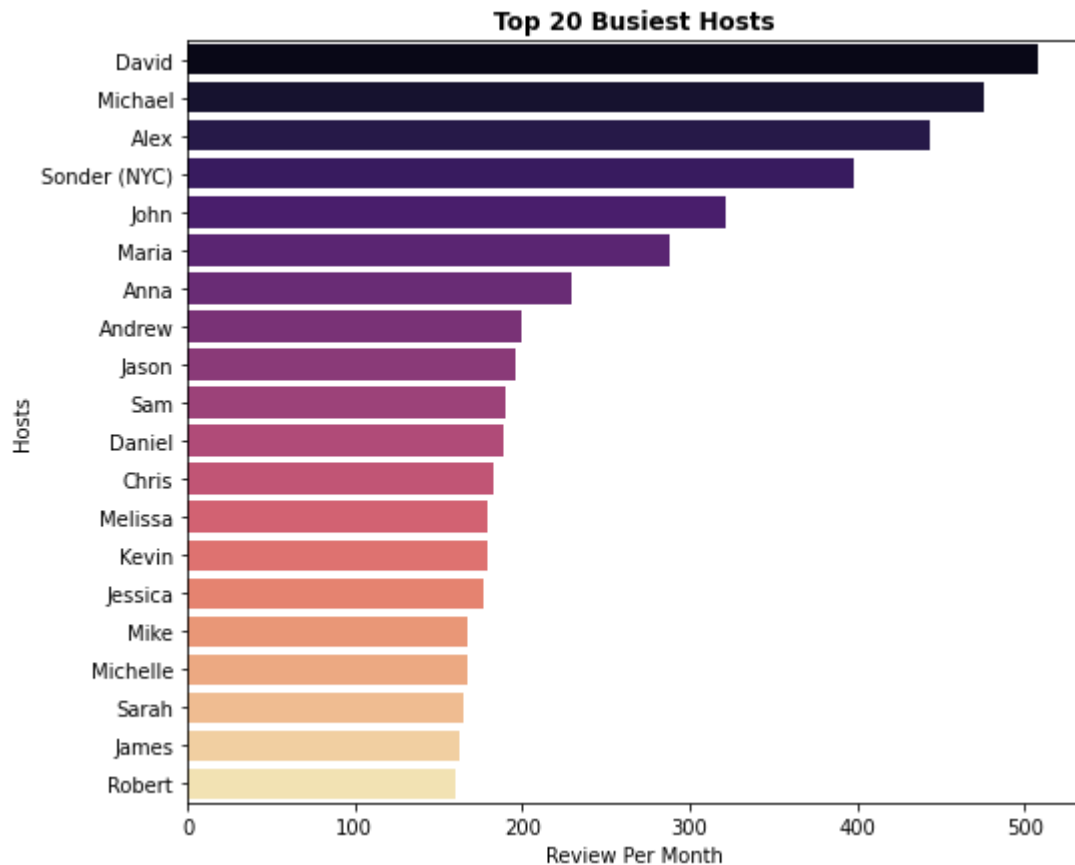
Average Price of Boroughs
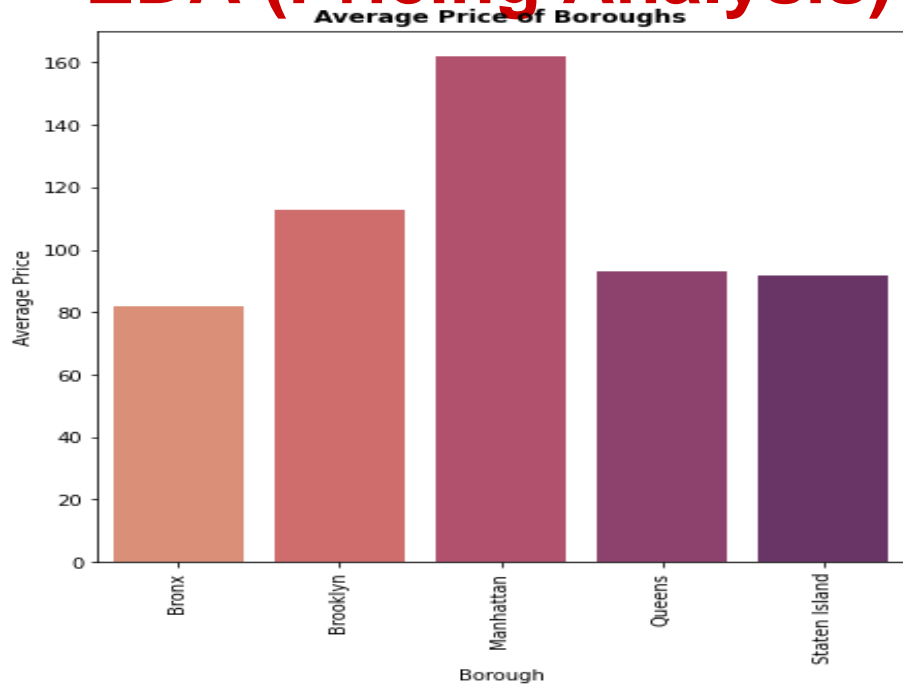
Manhattan has the highest average price for the listings follwed by Brooklyn which is obvious as standard of living is higher in these areas.

Average Price Per Night

It's generally cheaper to stay in rooms between 14 and 28 nights.

EDA

Using latitude and longitude points we are able to visualize all NYC listings with its price. However, it is important to note that we had to replace some extremely high values as they are treated as outliers for our analysis.
we can see that the most expensive rooms tend to be located in the Manhattan zone.
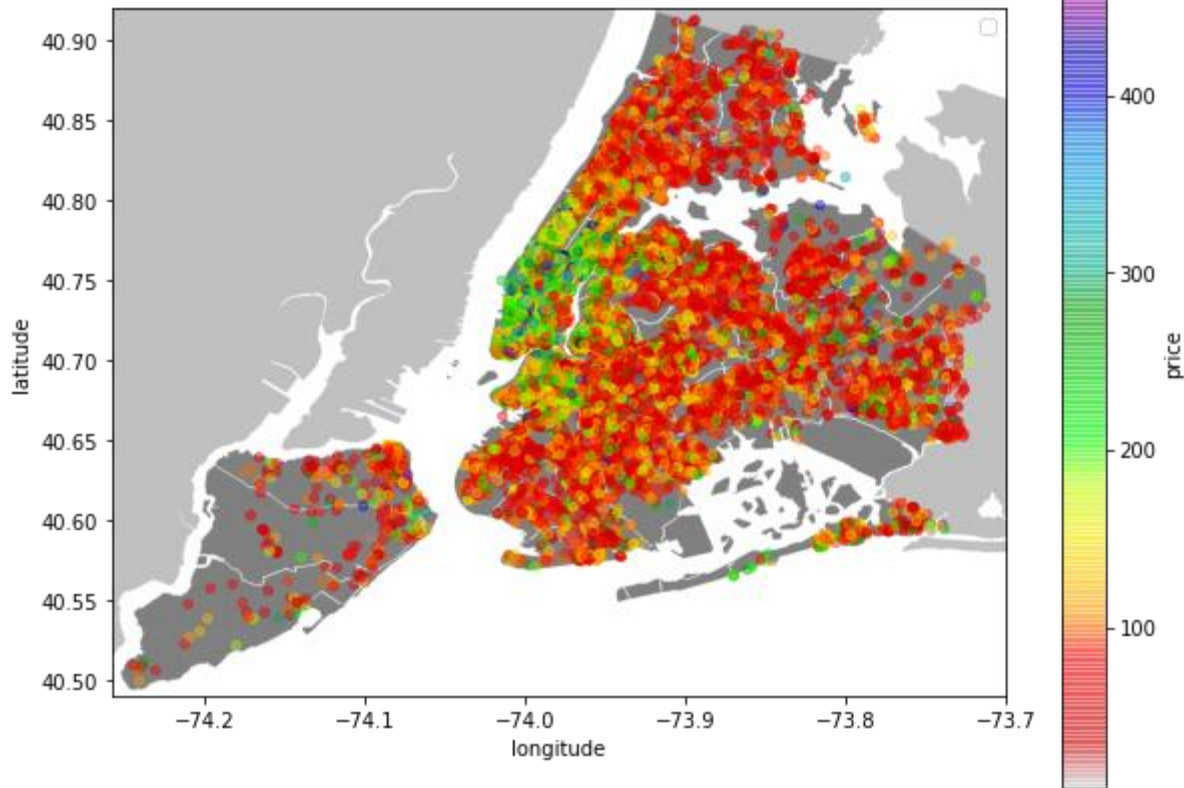
EDA

Manhattan has the highest range of prices for the listings followed by Brooklyn. Queens and Staten Island appear to have very similar distributions, Bronx is the cheapest of them all. This distribution and density of prices were completely expected , as it is no secret that Manhattan is one of the most expensive places in the world to live in, where Bronx on other hand appears to have lower standards of living.

Price distribution of each room type in each NYC zone
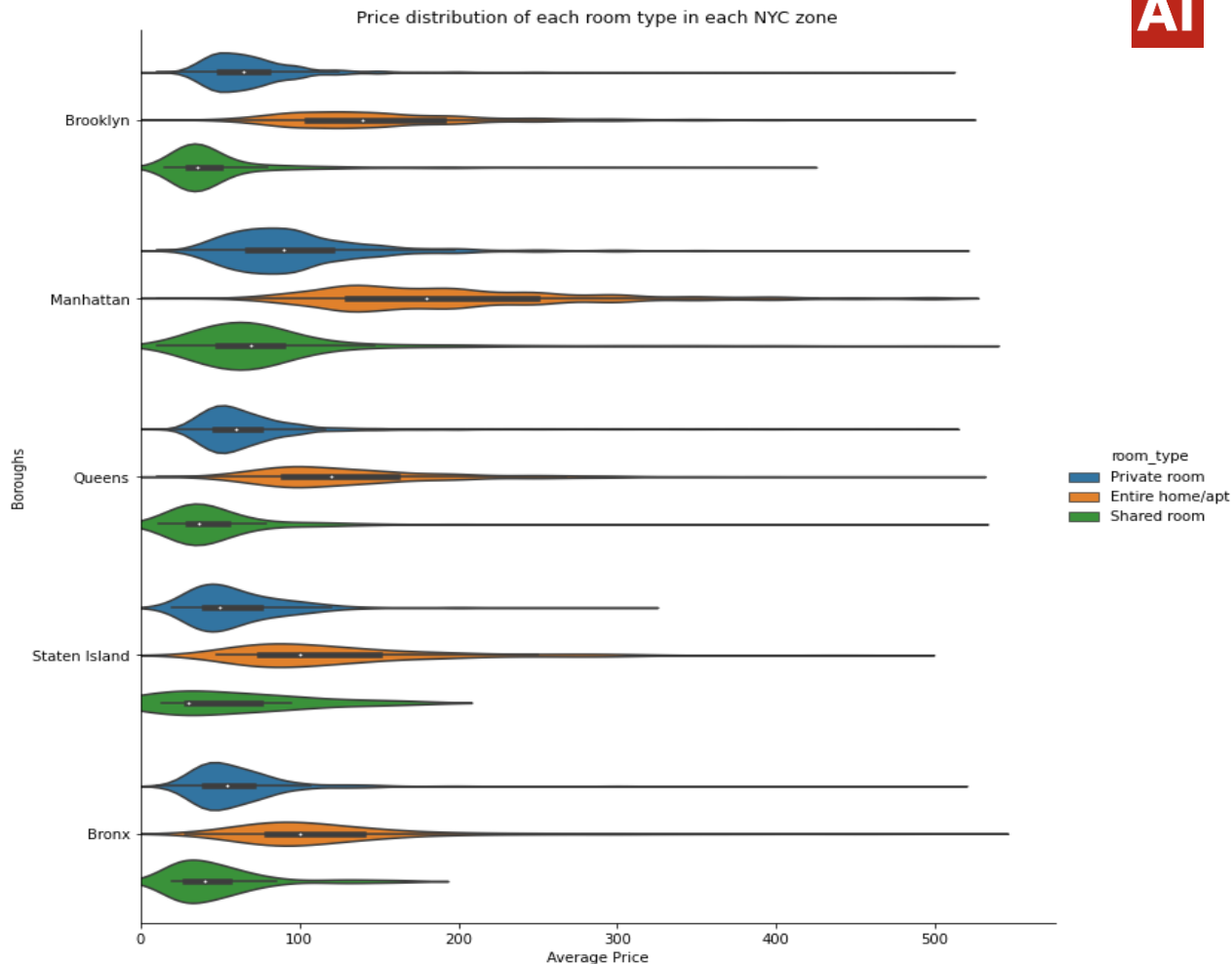
EDA

When looking at the Popularity/Price plot, we can see that rooms with more reviews tend to be cheaper (although the Entire appartment class is more noisy than the others). This can be explained as: the more expensive is an apartment, the less people stays there and therefore, the less reviews.

Now from the slide 9 we can actually conclude that the busiest hosts are the ones which have cheaper listings comapred with the others as these are affordable by major population.

EDA (Room Description)

Top 25 Describing words for Listings

Hosts use very simple terms describing the space and the area where the listing is. This technique was somewhat expected as dealing with multilingual customers can be tricky and you definitely want to describe your space in a concise and understood form as much as possible.

Thankyou