Storytelling Case Study: Airbnb, NYC
By Shubham and Priya.

Problem Statememnt

Background

During the COVID Phase, 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. And can make necessary changes to increase the revenue.

Objective

To prepare for the next best steps that Airbnb needs to take as a business by analysing a dataset consisting of various Airbnb listings in New York. And based on this analysis the management can take decisions.

# Average preferred price by customers according to the location

Room Type	Entire home/apt	Private room	Shared room
Neighbourhood_group			
Bronx	127.51	66.78	59.80
Brooklyn	178.32	76.50	50.52
Manhattan	249.23	116.77	88.97
Queens	147.05	71.76	69.02
Staten Island	173.84	62.29	57.44
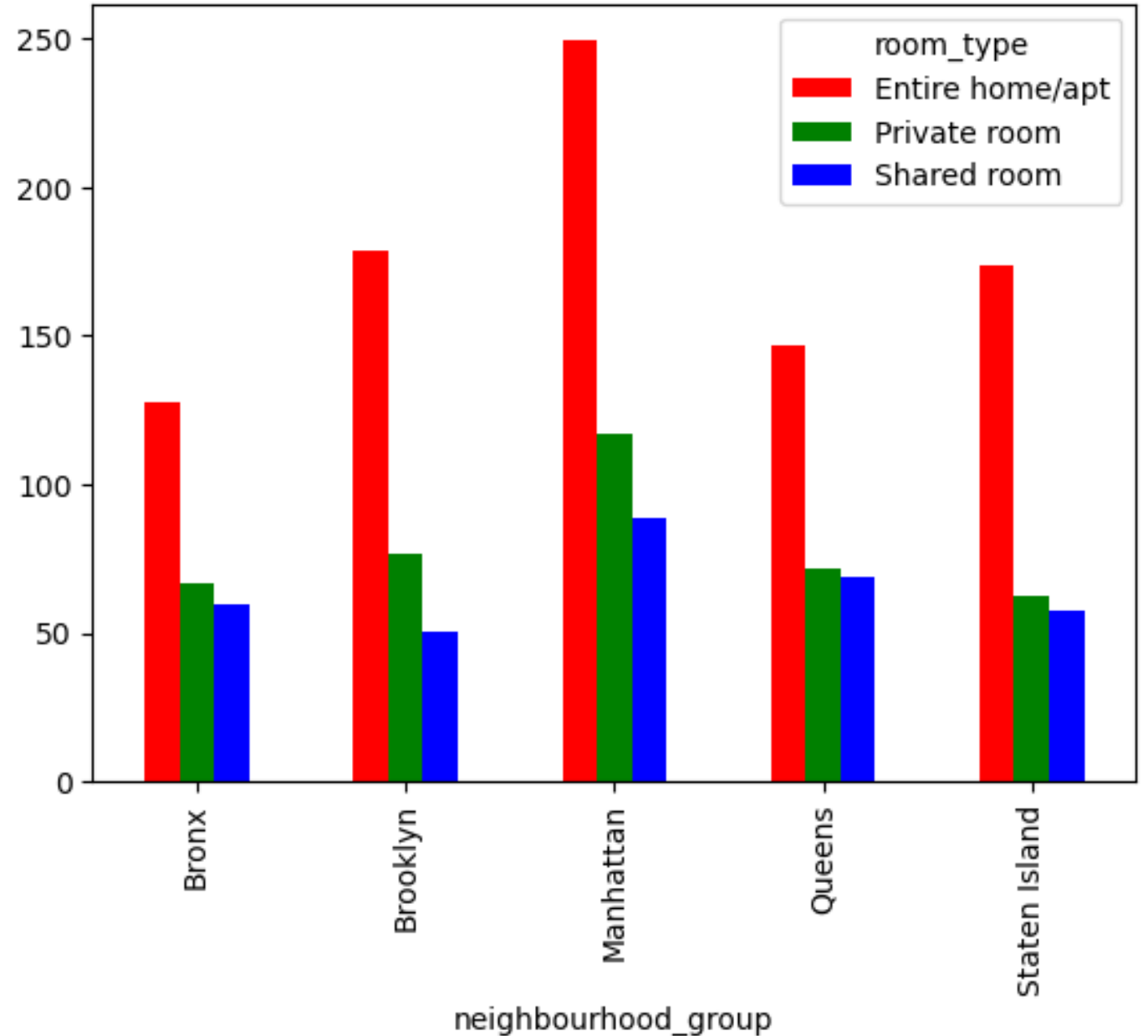
The number of active hosts in each location (Where most hosts prioritize property ownership)

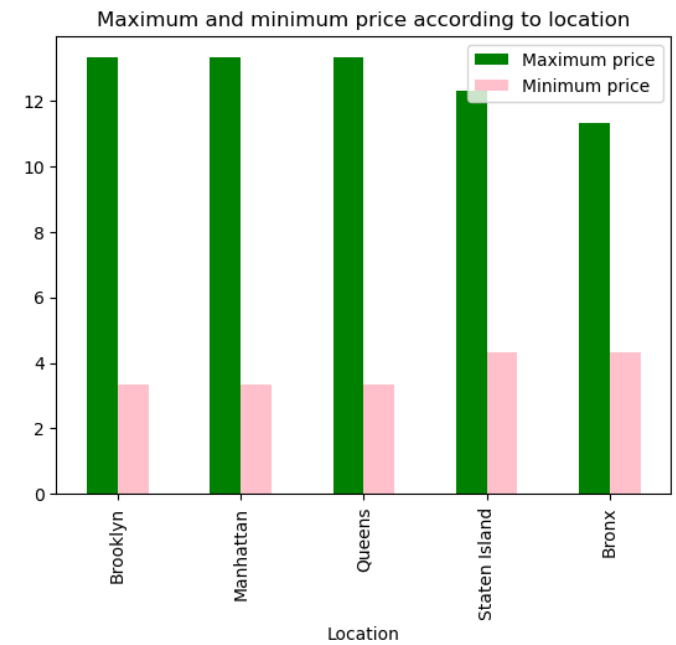| Location	Host
Manhattan	21661
Brooklyn	20104
Queens	5666
Bronx	1091
Staten Island	373
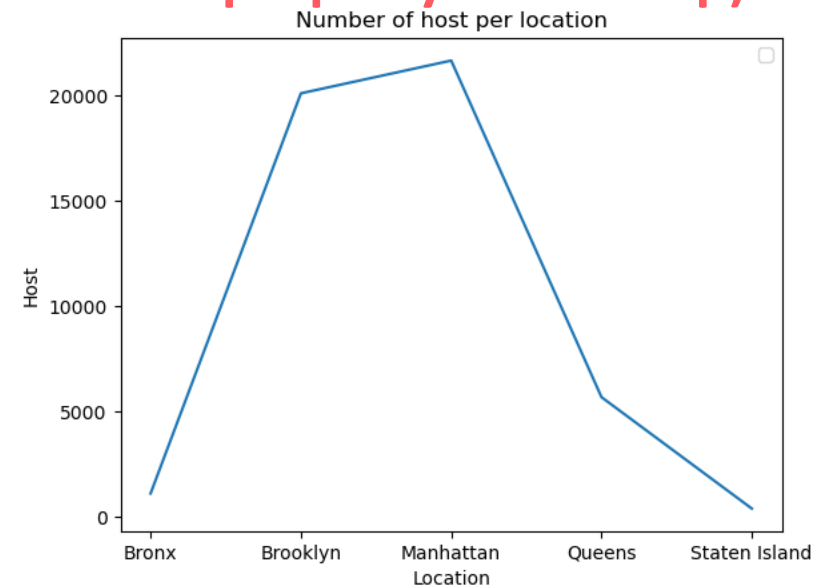
*Manhattan is the primary hub where hosts predominantly conduct their business

locations having highest and lowest tariff paid by customers

| Location	Maximum price	Minimum price
Brooklyn	13.28	3.32
Manhattan	13.28	3.32
Queens	13.28	3.32
Staten Island	12.28	4.32
Bronx	11.28	4.32

Which Airbnb hosts are the most popular in New York basis no of reviews and availability in a year.

host_id	host_name	Number of reviews	Availability in a year
22959695	Gurpreet Singh	1157	0.0
99392252	Michael	732	0.0
121391142	Deloris	693	0.0
792159	Wanda	480	0.0
37818581	Sofia	479	0.0
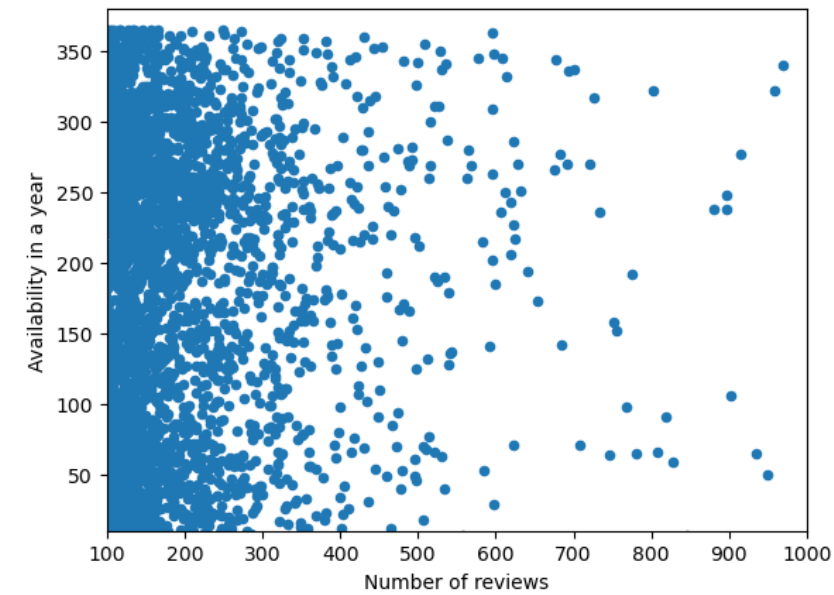
Count and Proportion (%) of each room type
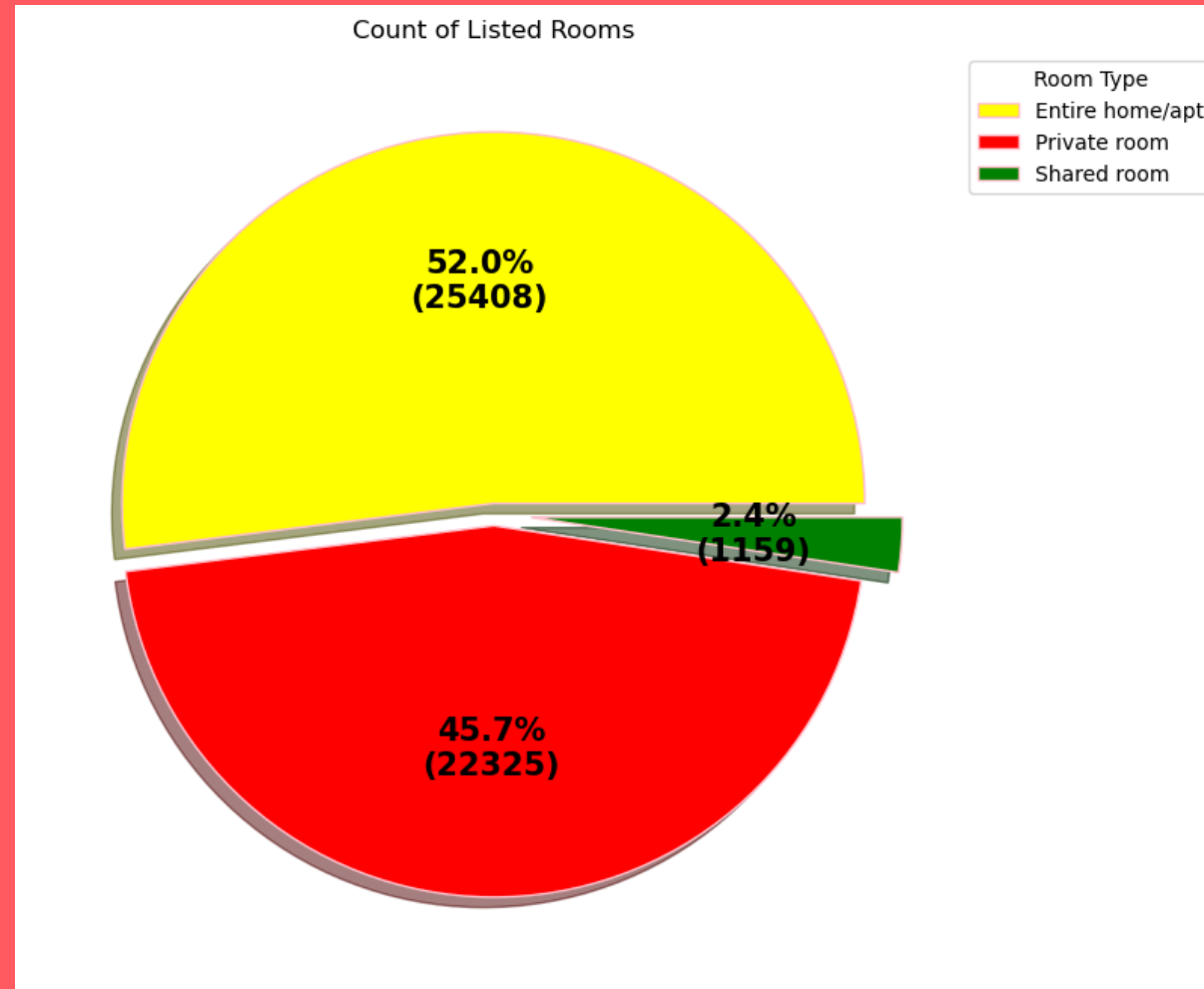
Top 25 most used words in listing names

Words	Counts
In	16725
room	9379
bedroom	7231
private	6978
apartment	6112
cozy	4627
studio	3772
brooklyn	3629
apt	3571
spacious	3387
1	3357
with	3092
2	3079
east	2967
and	2869
manhattan	2853
&	2820
park	2632
sunny	2536
beautiful	2320
near	2295
williamsburg	2293
-	2255
village	2055
heart	2044

Top 25 Used Words for Listing Names
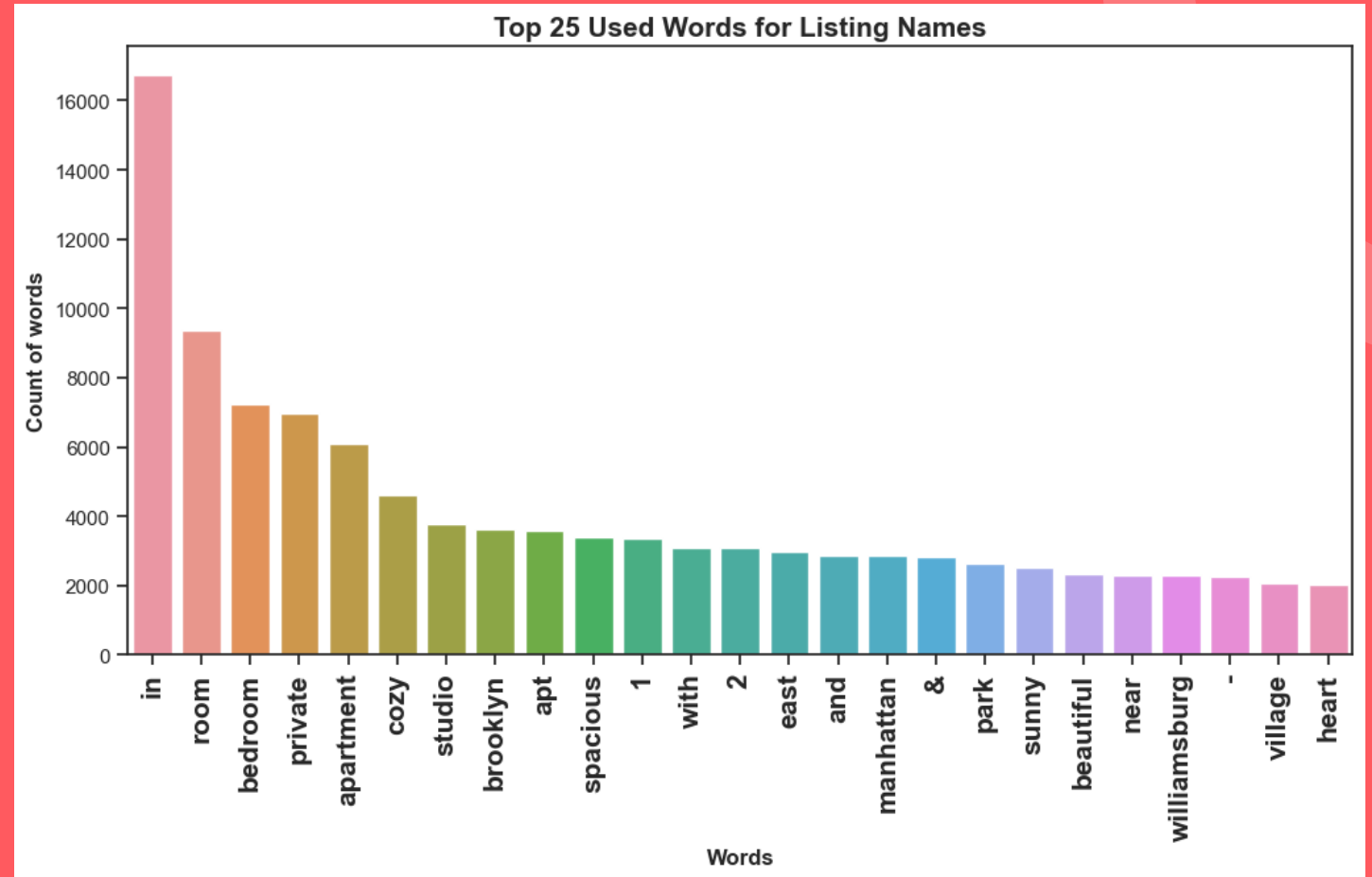
Top 10 hosts with most listings

Host_Id	Counts
2787	327
2845	232
4632	121
4869	103
7192	96
7322	96
7356	91
8967	87
7490	65
7549	52

Hosts with most listings in New York
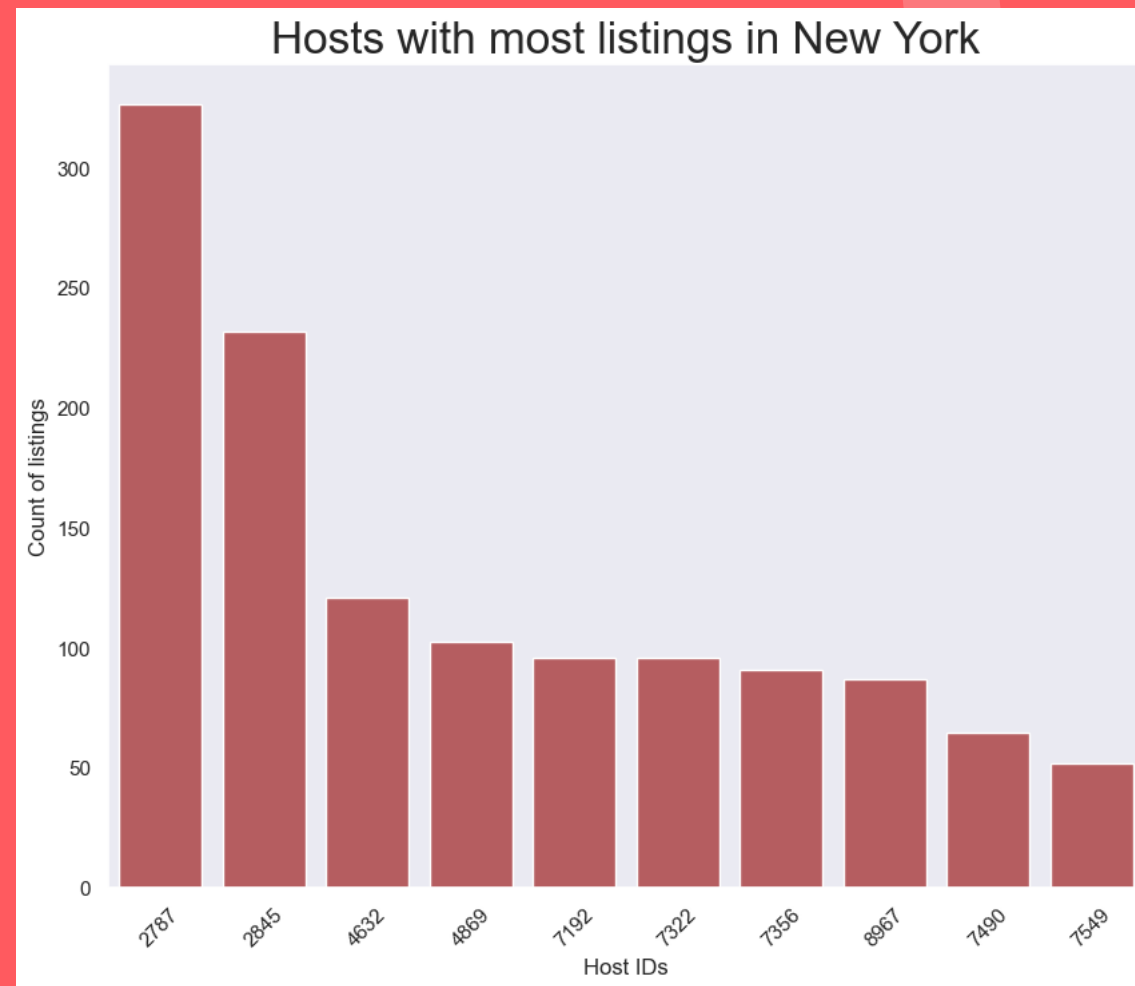
Top 3 hosts ranked by their revenue

Host Name	Host ID	Total Revenue
Sonder (NYC)	219517861	82795
Blueground	107434423	70331
Sally	156158778	37097

The highest revenue is generated by Sonders (NYC) followed by Blueground and Sally

three hosts based on their turnover
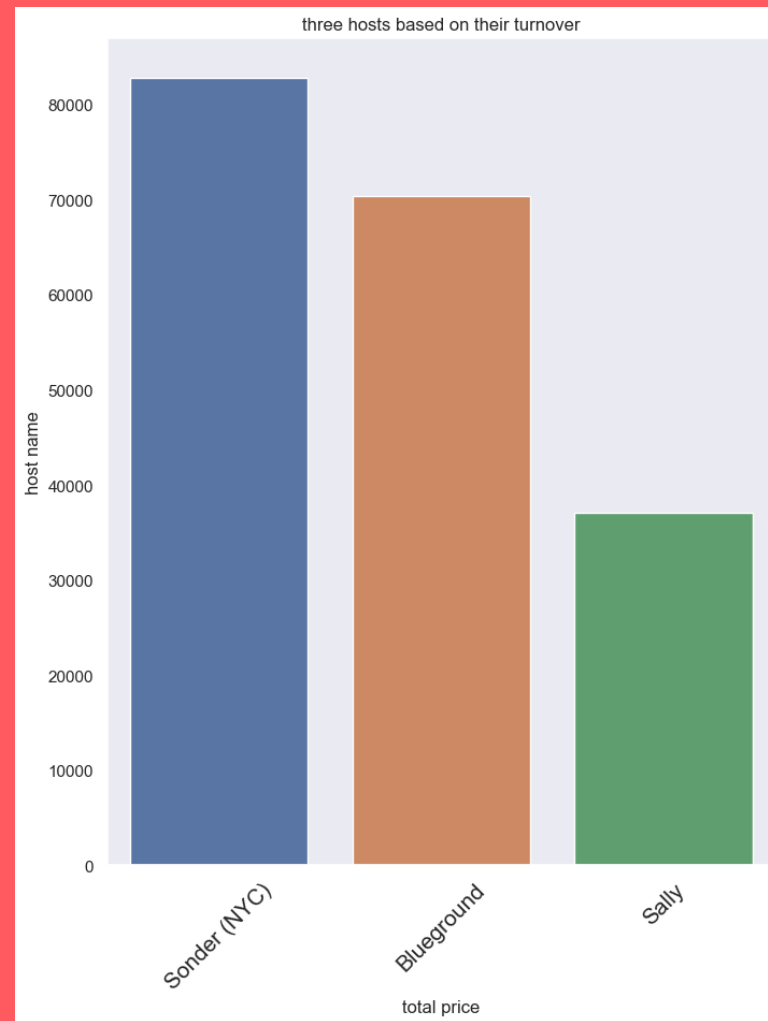
Neighbourhoods have the 10 highest listing count

Neighbourhood	listing_value
Williamsburg	3917
Bedford-Stuyvesant	3713
Harlem	2655
Bushwick	2462
Upper West Side	1969
Hell's Kitchen	1954
East Village	1852
Upper East Side	1797
Crown Heights	1563
Midtown	1545

find total no. nights spend per location
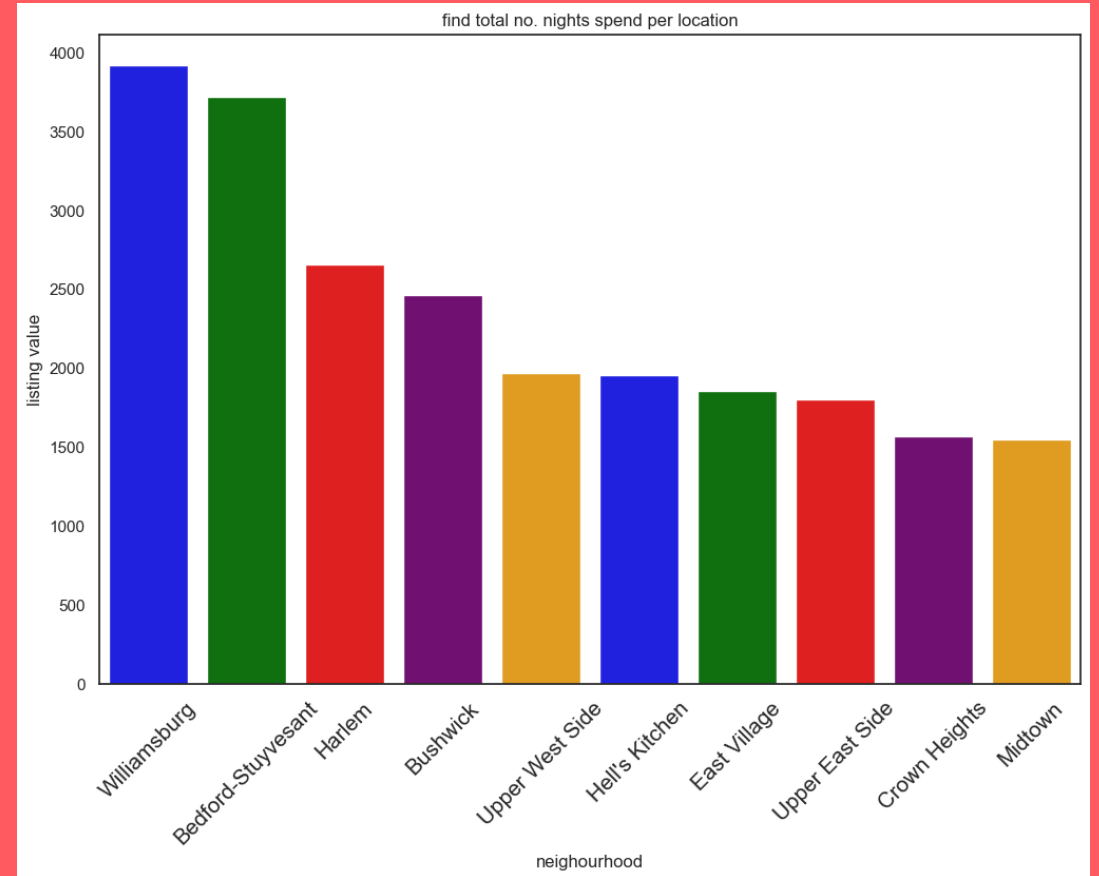
The highest revenue is generated by <u>Sonders (NYC)</u> followed by <u>Blueground</u> and <u>Sally</u>

Conclusion

- Manhattan emerges as the focal point for hosts conducting their business in New York. It serves as the epicenter of hosting activities, likely due to its central and attractive location

- Brooklyn, Queens, and Manhattan stand out as the regions where customers are willing to pay the highest nightly rates, reaching up to 10,000,while there are listings available for as low as 10

- When considering room types, there are clear pricing distinctions. Entire homes and apartments command the highest average price at approximately $157,followed by private rooms at around $75, and shared rooms at roughly $60.

- Notably, 'Entire home/apt' listings dominate the market with a substantial 52% share, while 'Shared Room' listings constitute a mere 2.4% of the total.

- Private rooms in Brooklyn and Manhattan attract guests who prefer extended stays, indicating a preference for these areas when seeking longer-term accommodations.

- Analyzing listing descriptions, we find that words like 'bedroom,' 'cozy,' 'private,' 'apartment,' and 'spacious' are recurrent, reflecting what guests value in their accommodations. In contrast, terms like 'park,' 'near,' 'village,' and 'heart' are comparatively less emphasized.

- Furthermore, the dataset reveals that the top 10 hosts collectively account for almost 2.5% of all listings, totaling 1,270 properties. Among these hosts, Sonder (nyc), Blueground, and Sally have the highest turnovers, with Sonder (nyc) standing out as the best-performing host.

- Manhattan proves to be the preferred choice for the majority of customers, with 63.2% of them opting for stays in this vibrant borough. Only a small fraction 1.6% chosen shared rooms, indicating a strong preference for private and entire home accommodations.

Methodology used for Analysis

1.Introduction and Library Imports:
   a.   Introduction and Importing Required Libraries
2.Data Loading:
   a.   Loading the Dataset
3.Data Preprocessing:
   a.   Data Cleaning and Preprocessing
   b.   Handling Duplicates, Cleaning Columns, and Dealing with Null Values
   c.   Data Transformation
4.Exploratory Data Analysis (EDA):
   a.   Exploratory Data Analysis and Visualization
   b.   Analyzing Relationships Between Features through Visualizations
5.Average Preferred Price by Location:
   a.   Determining the Average Preferred Price by Location
6.Number of Active Hosts by Location:
   a.   Identifying Locations with the Highest Number of Active Hosts
7.Rent by Location:
   a.   Analyzing Rent Variations by Location

8.Most Popular Airbnb Host in New York:
   a.   Finding the Most Popular/Demanded Airbnb Host in New York
9.Total Count of Each Room Type:
   a.   Counting the Total Number of Each Room Type
10.Room Types and Availability by Neighborhood Groups:
   a.   Exploring the Relationship Between Room Types and Availability in Different Neighborhood Groups
11.Top 25 Most Used Words in Listing Names:
   a.   Discovering the Top 25 Most Frequent Words in Listing Names
12.Top 10 Hosts with Most Listings:
   a.   Finding the Top 10 Hosts with the Most Listings
13.Top Three Hosts by Turnover:
   a.   Identifying the Top Three Hosts Based on Their Turnover
14.Total Nights Spent by Location:
   a.   Calculating the Total Number of Nights Spent per Location
15.Total Nights Spent by Room Types:
   a.   Analyzing the Total Number of Nights Spent by Room Types
16.Top 10 Highest Listing Neighborhoods:
   a.   Identifying the Top 10 Neighborhoods with the Highest Number of Listings