**Feel Like Home  
(aka FLH)   
A home rental company**

**Data Analysis**

**Aim**

The aim of this analysis is to leverage data to identify strategies that maximize revenue for "Feel Like Home" (FLH) during the post-COVID travel surge. This includes targeting the right neighborhoods, pricing strategies, identifying successful property types, understanding customer preferences, optimizing host attributes, and boosting traction for less popular properties.

**Business Objectives**

Target High-Demand Neighborhoods:

* Objective: Identify and prioritize neighborhoods with high demand.
* Approach: Analyze neighbourhood groups and its neighbourhoods.

Optimize Pricing Strategies:

* Objective: Determine the optimal pricing ranges preferred by customers.
* Approach: Conduct pricing analysis.

Identify Successful Property Types:

* Objective: Understand which property types are most and least successful to focus on high-performing ones and improve the others.
* Approach: Analyze room types, number of reviews given by customers.

Understand Customer Preferences:

* Objective: Discover key factors that customers prioritize in their stays to enhance property listings and services.
* Approach: Analyze customer reviews, location, types of rooms.

Enhance Host Attributes:

* Objective: Identify host attributes that contribute to higher bookings and customer satisfaction.
* Approach: Examine host\_attributes such as number of listings, reviews.

Boost Traction for Unpopular Properties:

* Objective: Develop strategies to increase visibility and bookings for properties with lower demand.
* Approach: Implement targeted marketing campaigns, offer discounts to attract more customers.

**Data Description**

This dataset has 48895 rows and 16 columns.

The 16 features are:

1. id: Unique identifier for the listing.
2. name: Name of the listing.
3. host\_id: Unique identifier for the host.
4. host\_name: Name of the host.
5. neighbourhood\_group: Larger area the listing is located in.
6. neighbourhood: Specific area the listing is located in.
7. latitude: Latitude coordinate of the listing.
8. longitude: Longitude coordinate of the listing.
9. room\_type: Type of room being offered (e.g., Entire home/apt, Private room).
10. price: Price per night in the listing's currency.
11. minimum\_nights: Minimum number of nights required for booking.
12. number\_of\_reviews: Total number of reviews the listing has received.
13. last\_review: Date of the last review received.
14. reviews\_per\_month: Number of reviews the listing receives per month.
15. calculated\_host\_listings\_count: Number of listings the host has.
16. availability\_365: Number of days the listing is available for booking in a year.

**Notebook Contents :**

1. Dataset Information

Observations

2. Exploratory Data Analysis (EDA)

2.1. Automated EDA(Sweetviz)

      2.2. Manually performing EDA with Observations

Handling missing values

Univariate analysis

      1. id

      2. host\_id

      3. neighbourhood\_group

      4. neighbourhood

      5. room\_type

      6. number of reviews

      7. minimum nights

      8.calculated\_host\_listings

      9.availablity 365

      Bivariate analysis

      1. Grouping by neighbourhood\_group and neighbourhoods

      2. average of (prices per neighbourhood\_group and room types)

      3. Count of Room Types per Neighbourhood Group

      4. Columns wrt price

      5. number of reviews wrt neighbourhoods

host attributes and calculate the average price and number of reviews

      6.Correlation

Q1 What are the neighborhoods that FLH needs to target?

Q2 What are the pricing ranges preferred by FLH customers?

Q3 What are the types of properties that are most successful and least successful for FLH?

Q4 What are customers looking for most in their stays with FLH?

Q5 What attributes do the hosts have that FLH should target to add to their service to increase revenue?

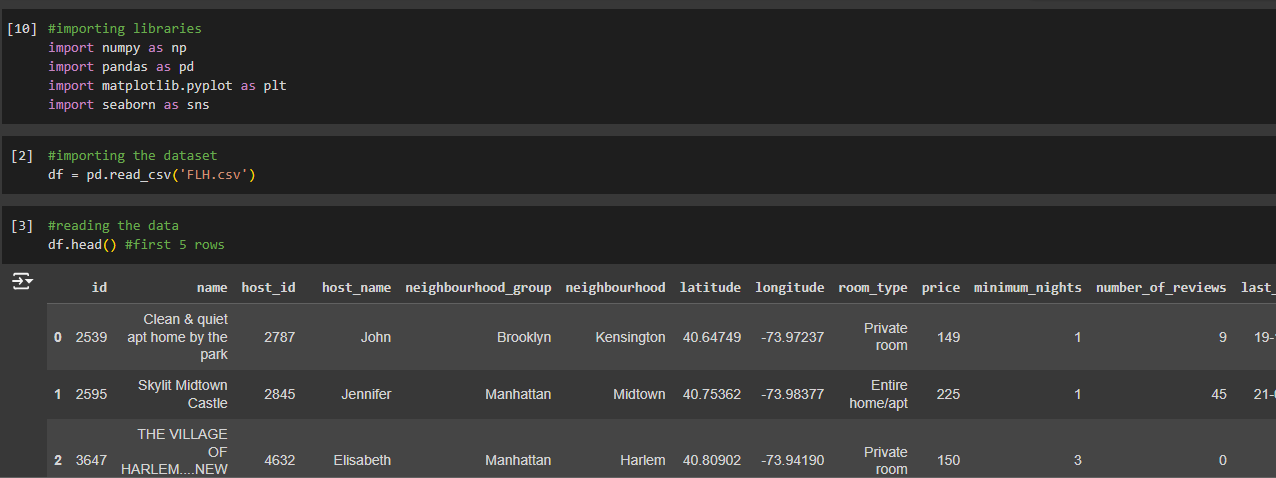
Q6 How to get unpopular properties more traction?

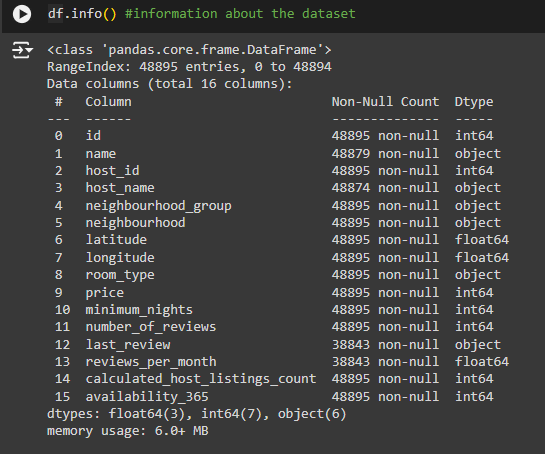
Conclusion

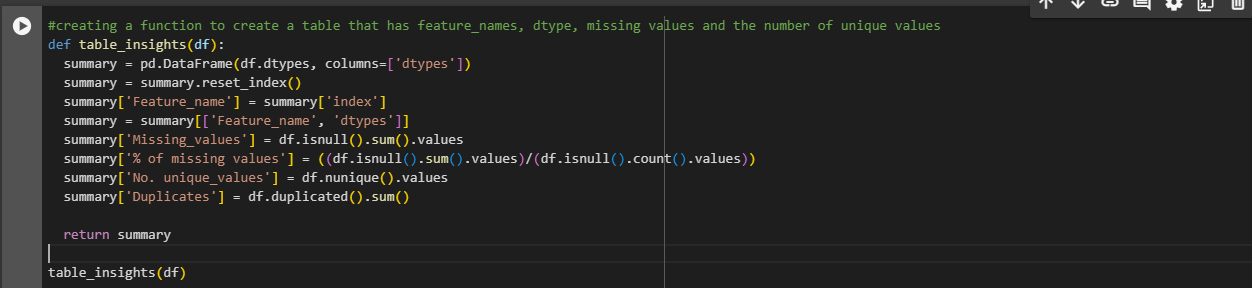
CONNECTING MYSQL

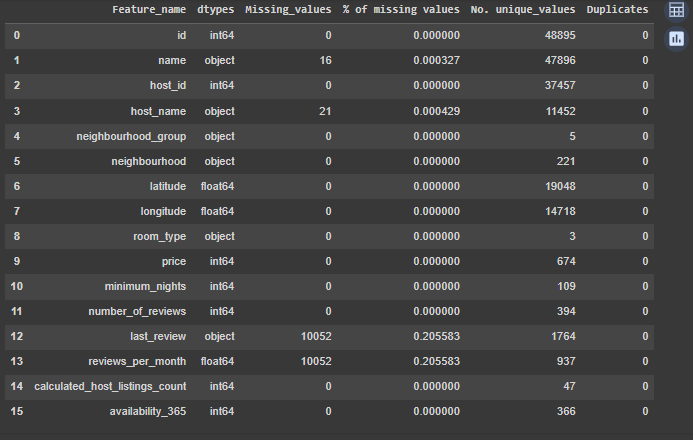
1. Dataset Information

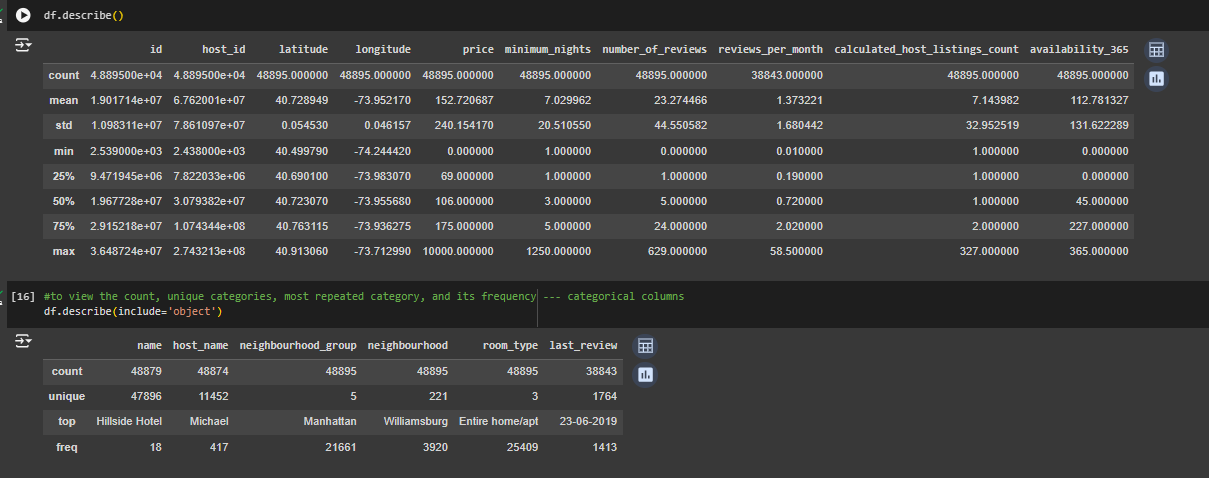
Importing libraries and reading the dataset



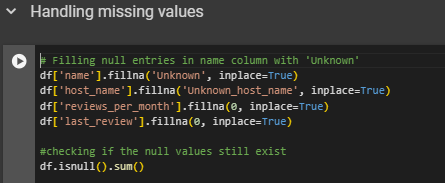






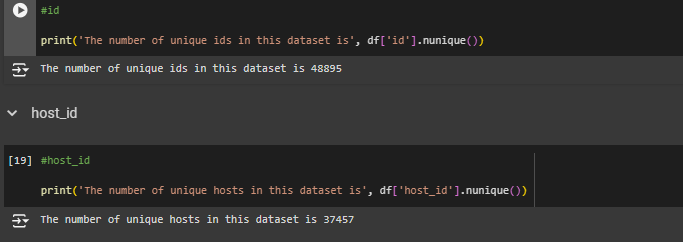


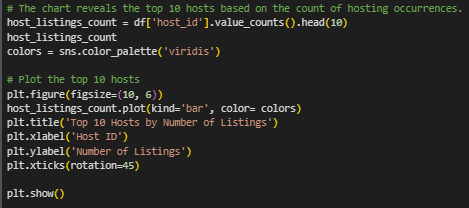
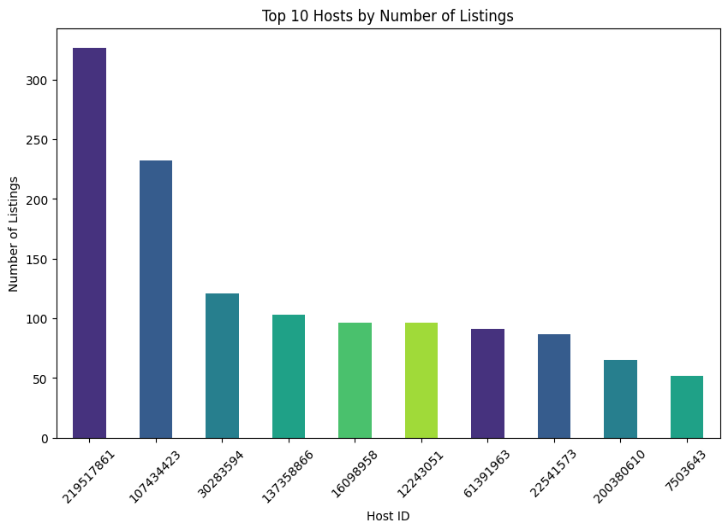
**Observations:**

* The dataset comprises of 48,895 rows with 16 columns.
* The columns with null values are 'name', 'host\_name', 'last\_review' and 'reviews\_per\_month', with the latter two having 10,052 null entries.
* There are no duplicate values in the dataset.
* The 'number\_of\_reviews','price','reviews\_per\_month', 'availability\_365' columns seem to have outiers, but as we are only analyzing the data, I think its okay to proceed without handling the outlier.

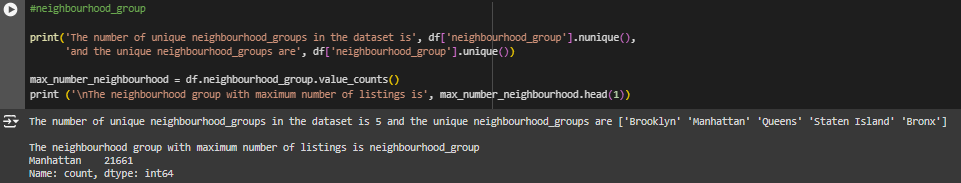
1. EDA

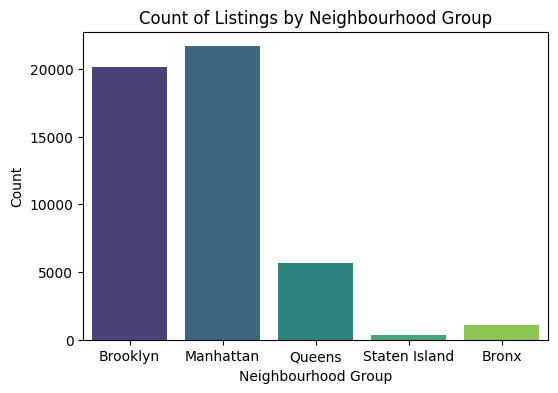
* 

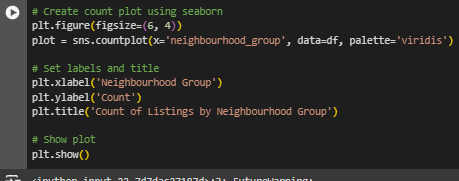


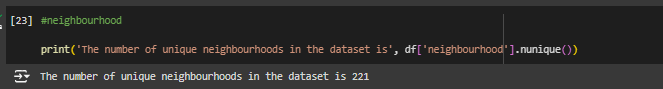
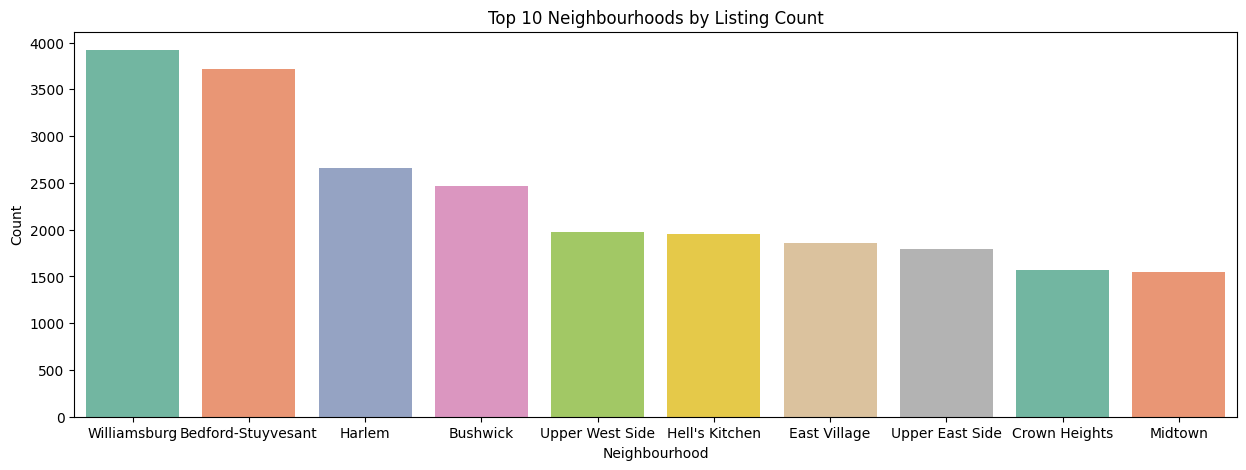
* From this chart we can see that this host '219517861' has the highest number of listings.
* Identifying the top hosts is important as it can lead to collaborations, recognition, and potential incentives for these hosts. Recognizing and rewarding top hosts could encourage them to continue providing exceptional service, contributing to positive customer experiences and loyalty.



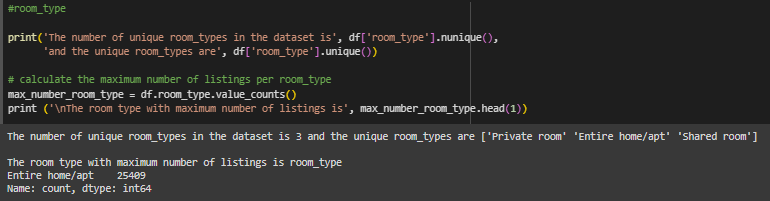
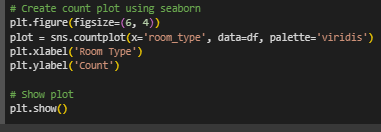


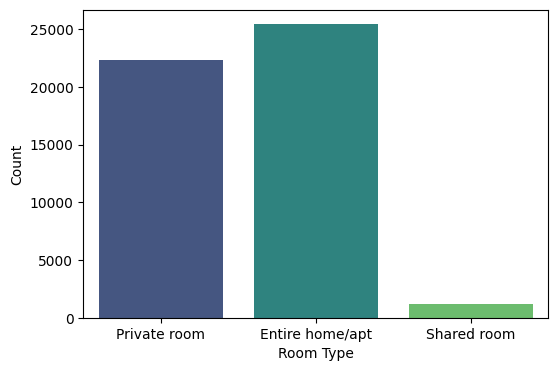


* This chart provides insights into the distribution of listings across different neighborhood groups.
* Here we can see that Manhattan has higher number of listing and Staten Island has lower number of listing.

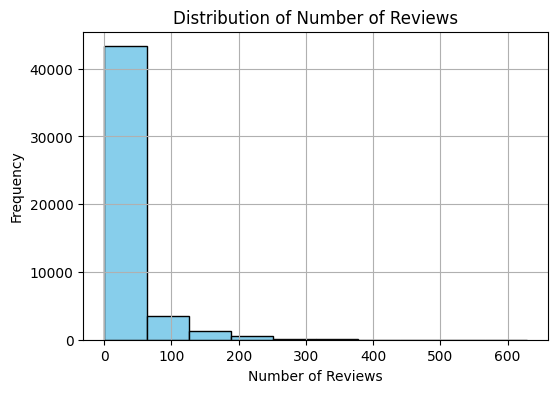
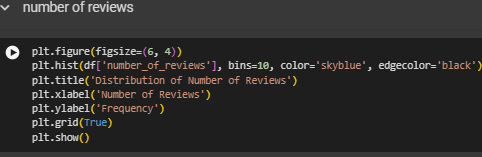


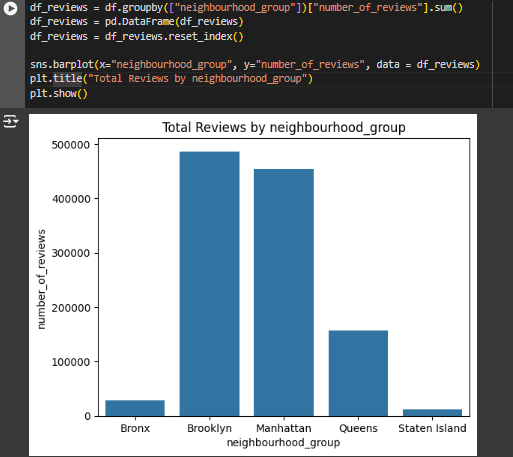
* The chart reveals the distribution of listings across different neighbourhoods.
* We see that Williamsburg has higher number of listings followed by Bedford-Stuyvesant.

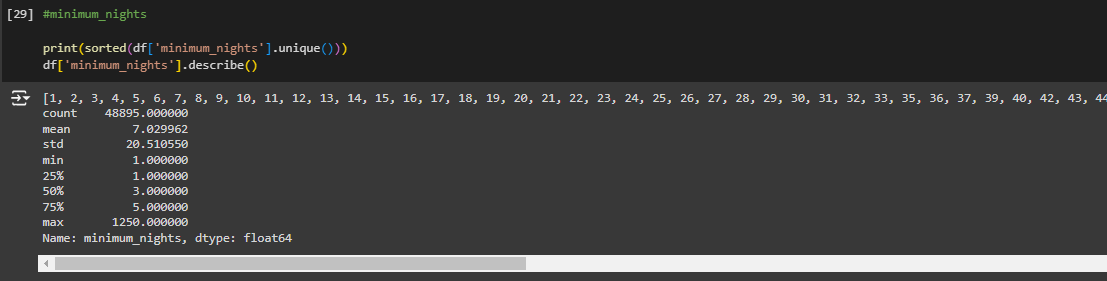


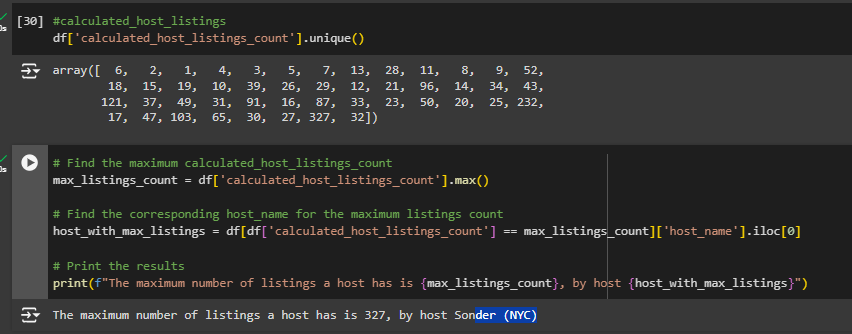


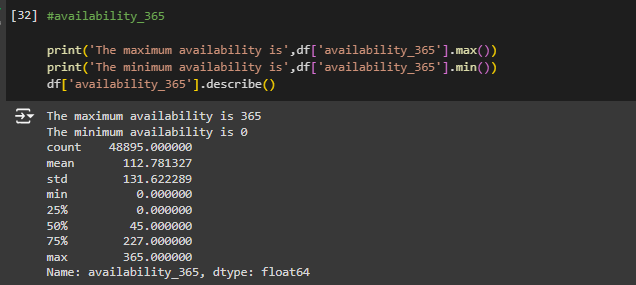
* The chart reveals the distribution of different room types in the dataset.
* Each bar represents a specific room type, and the height of the bar indicates the count (frequency) of the room type.
* Here we can see that 'Entire home/apt' Room type has higher number of frequency and 'Shared room' has lower frequency.
* Understanding distribution is benificial to highlight and promote the most popular room types.

  
 Most listings have between 0 and 50 reviews.

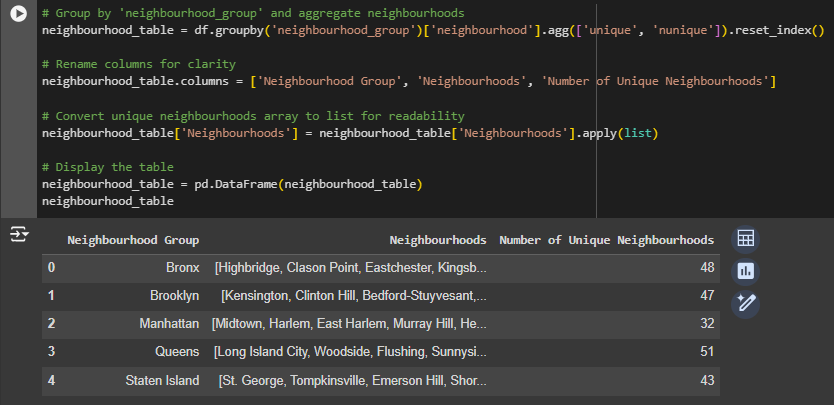


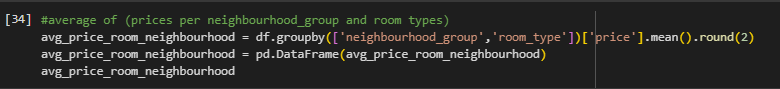


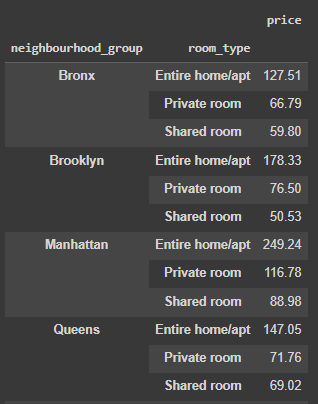
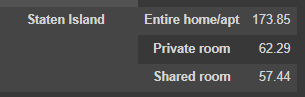


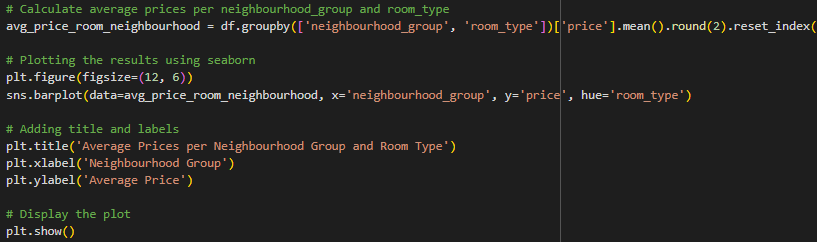


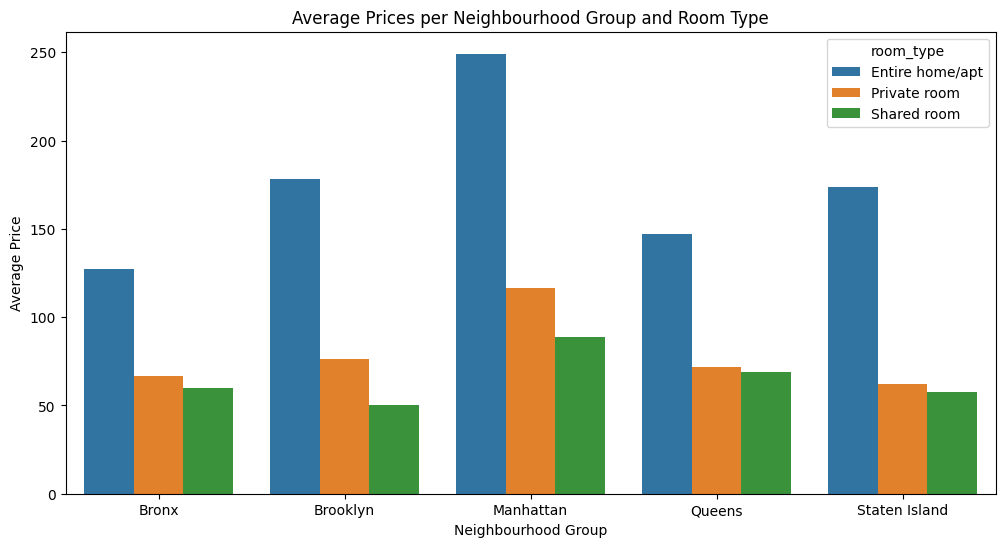
**Bivariate analysis:** This involves examining the relationship between two variables.

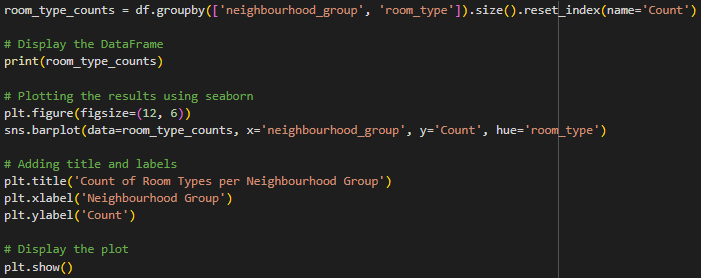
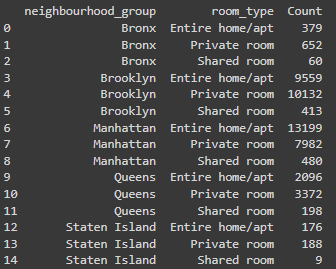
****

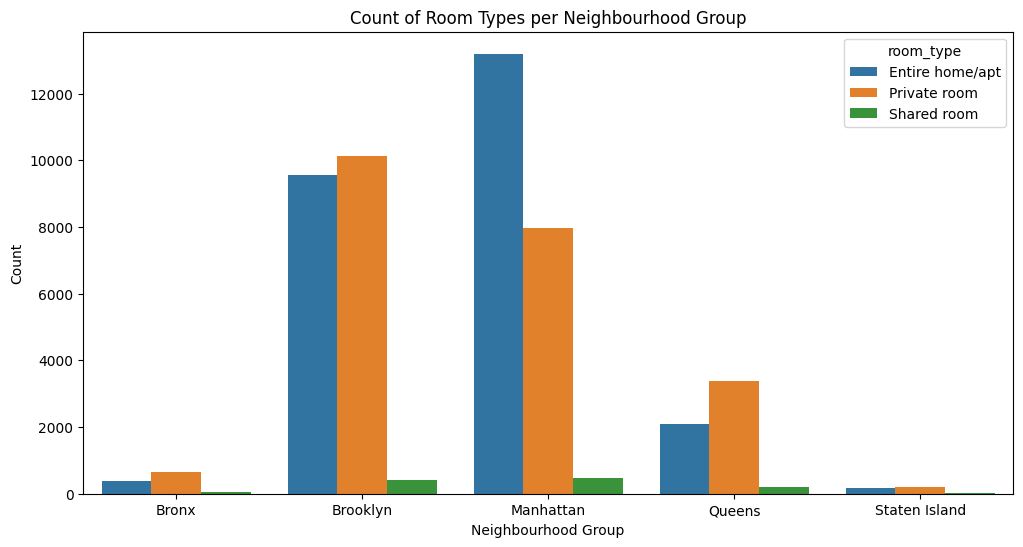
****

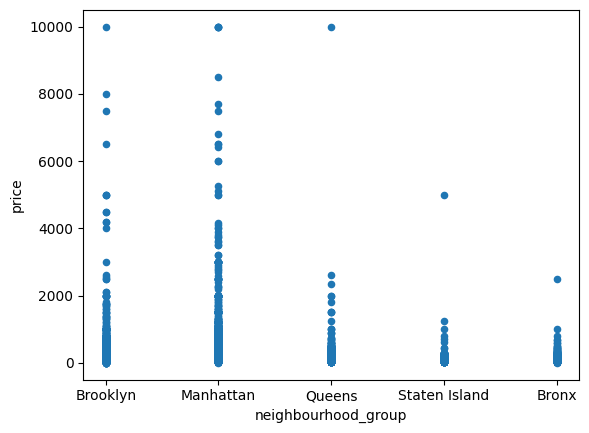
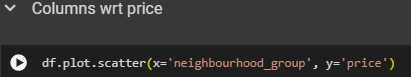
 

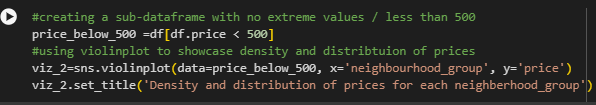


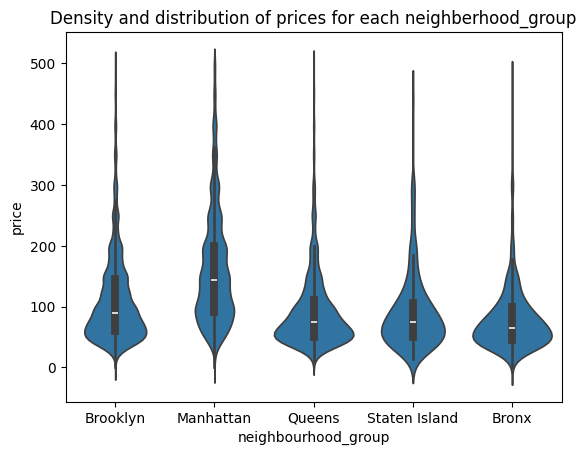






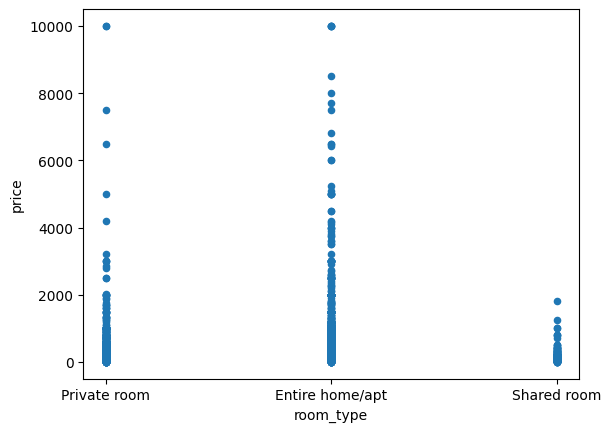




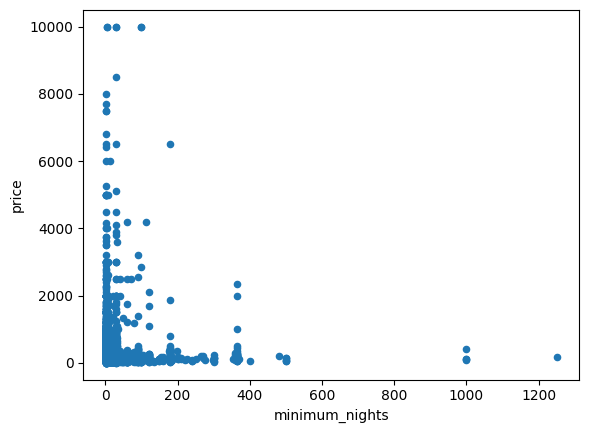


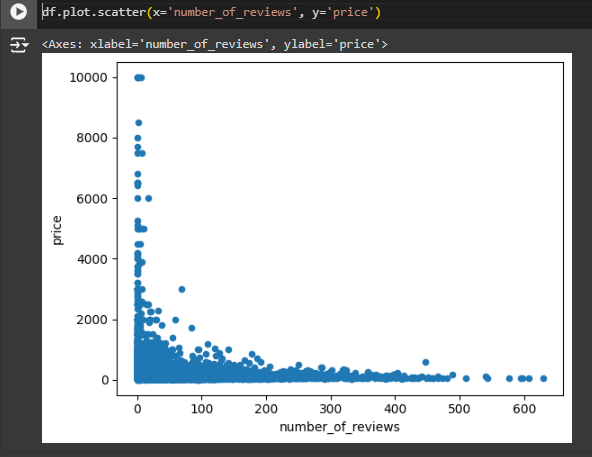
* we can state that Manhattan has the highest range of prices for the listings with $150 price as average observation, followed by Brooklyn with \$90 per night.
* Queens and Staten Island appear to have very similar distributions, Bronx is the cheapest of them all. This distribution and density of prices were completely expected; for example, Manhattan is one of the most expensive places in the US.



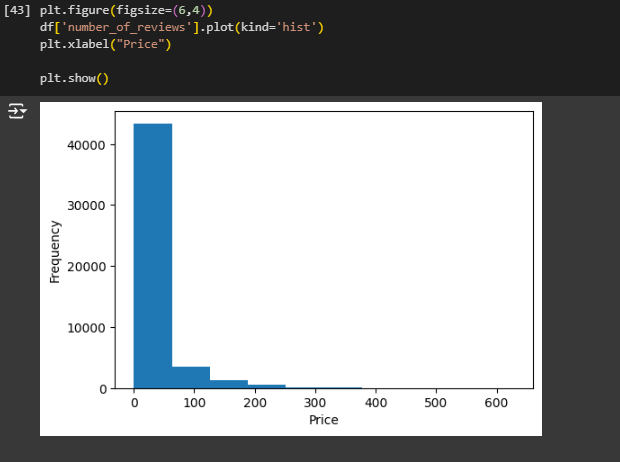




There is a negative relationship as when we increase the minimum number of nights the price decreases.

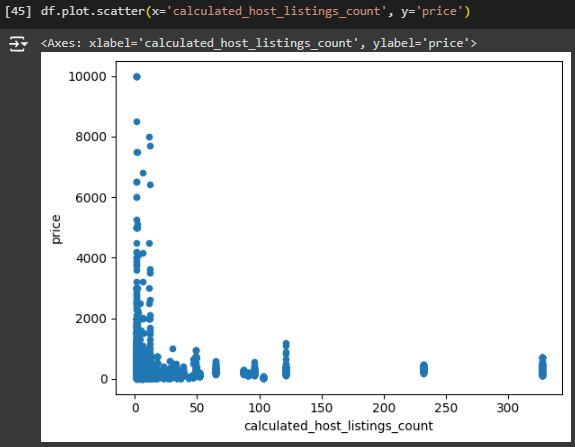


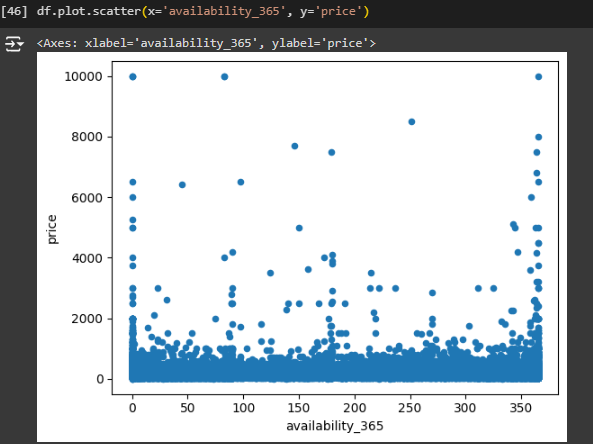
From visualization we can say that most number of people like to stay in less price and their reviews are higher in those areas.



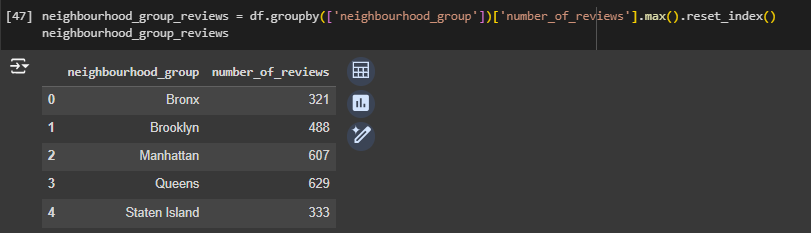
Low cost rooms or in range 0-50 $ have more reviews

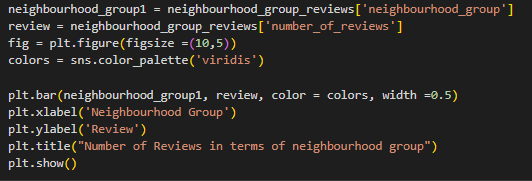
Lower the price, more the number of reviews

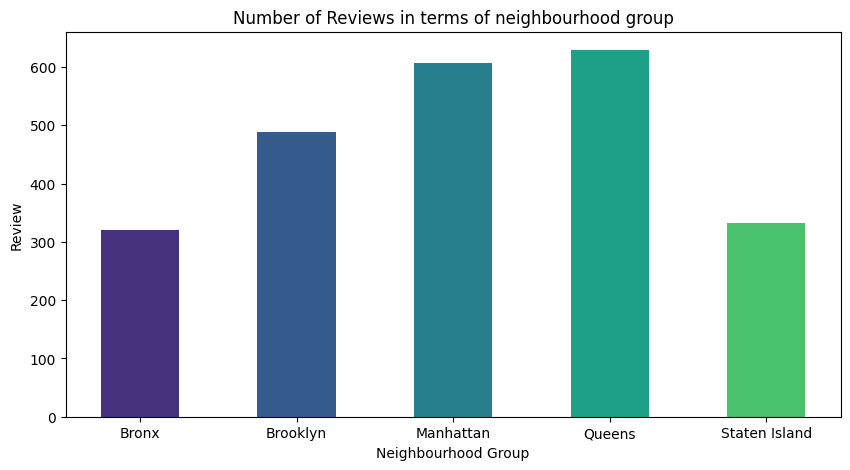


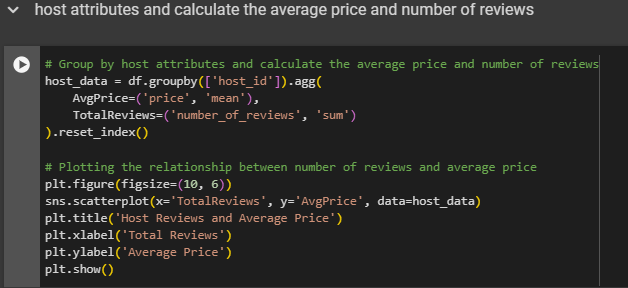


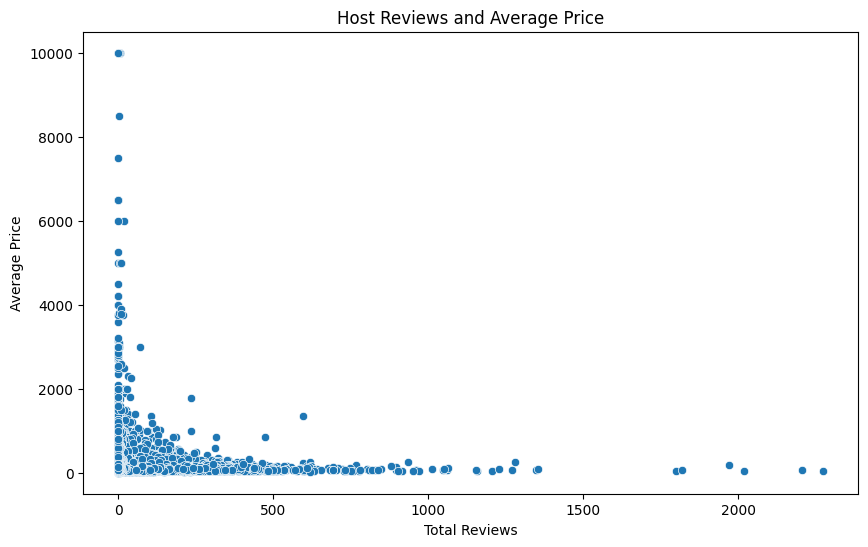
number of reviews wrt neighbourhoods

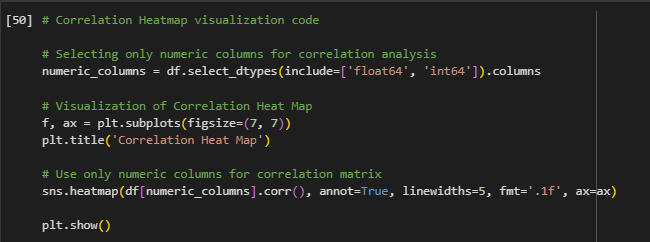
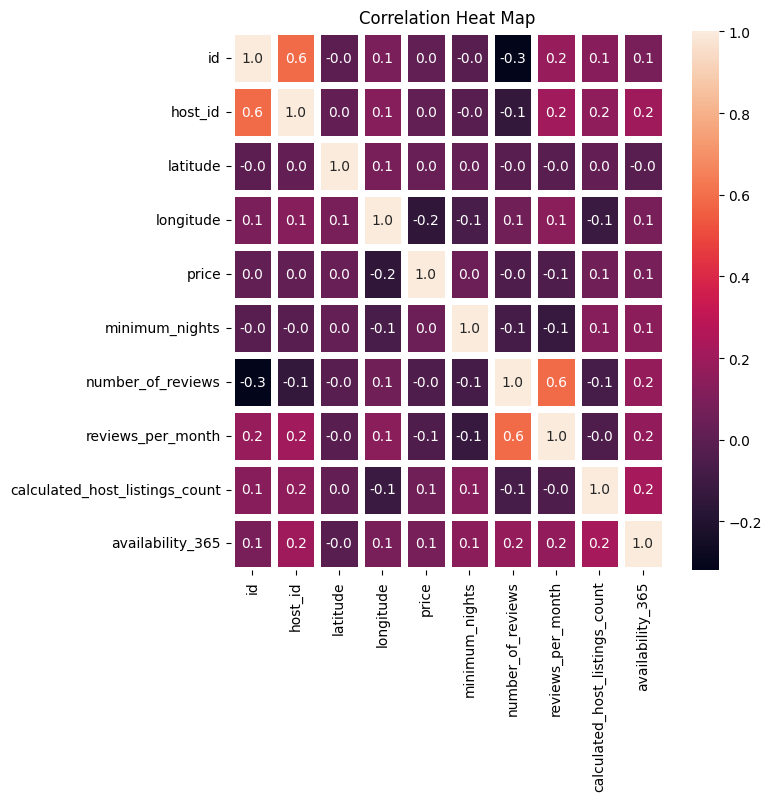




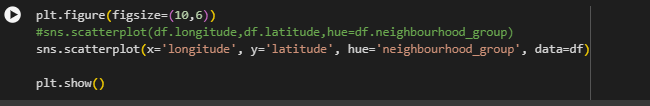


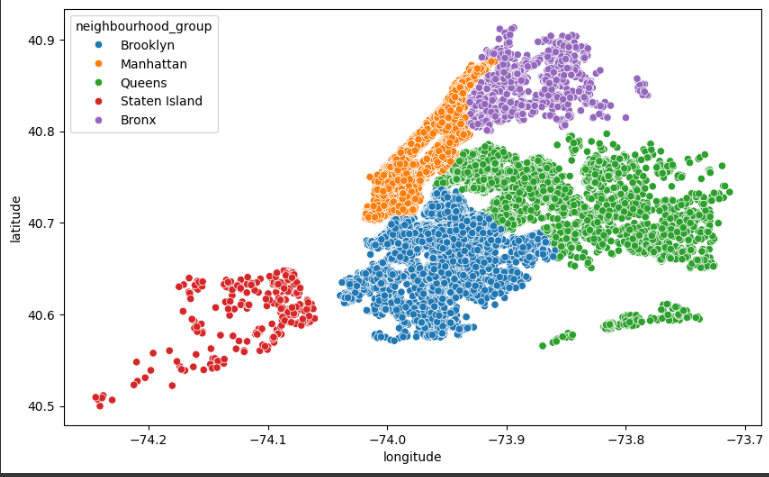


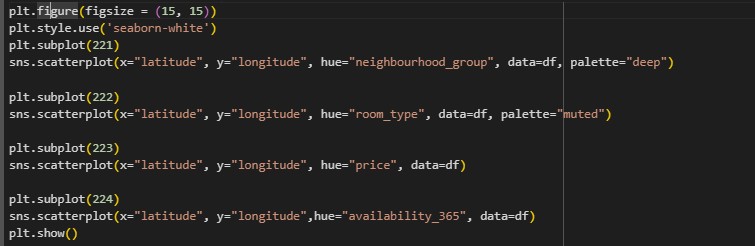


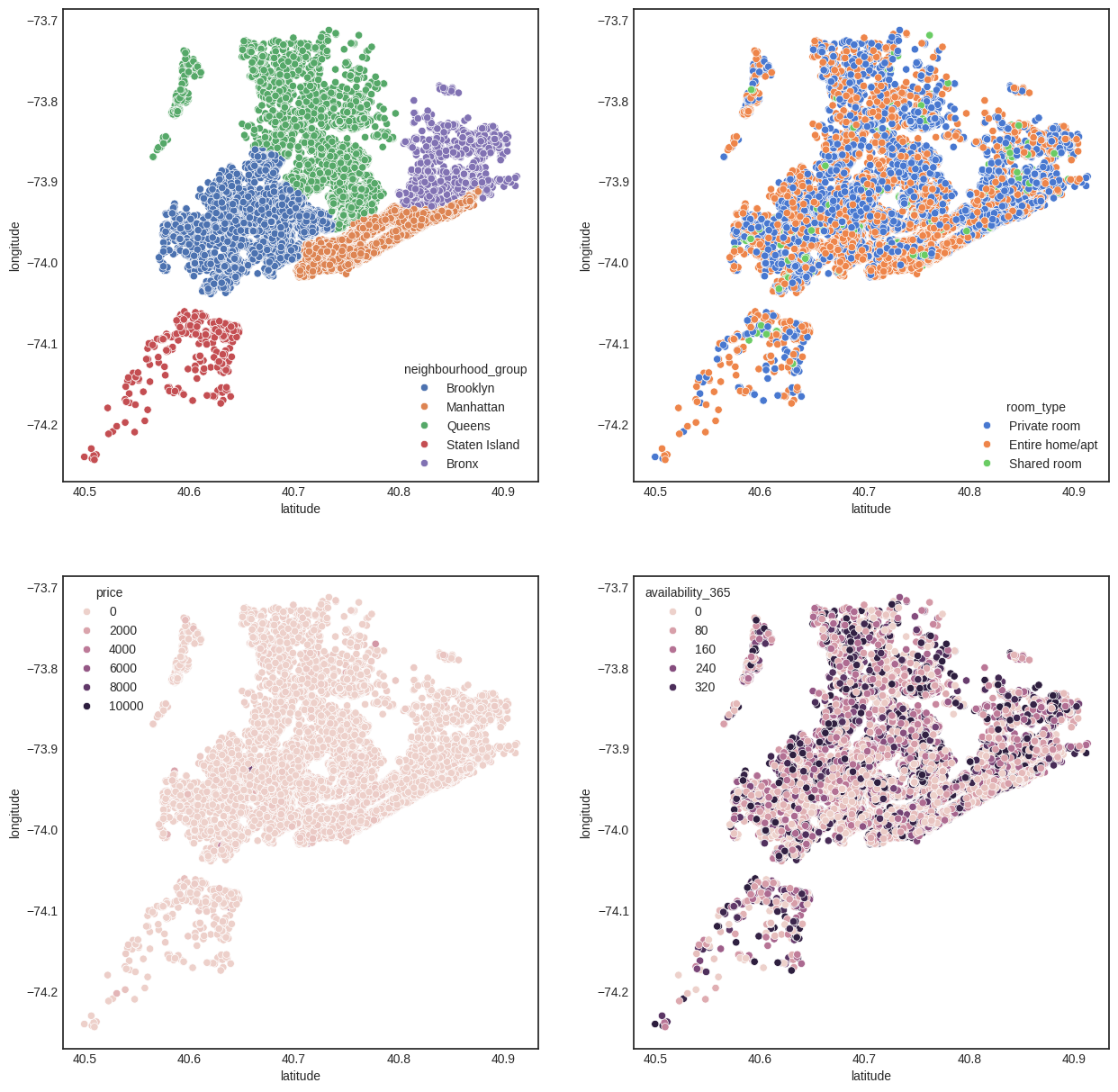


Number of reviews and reviews per month have strong positive correlation.







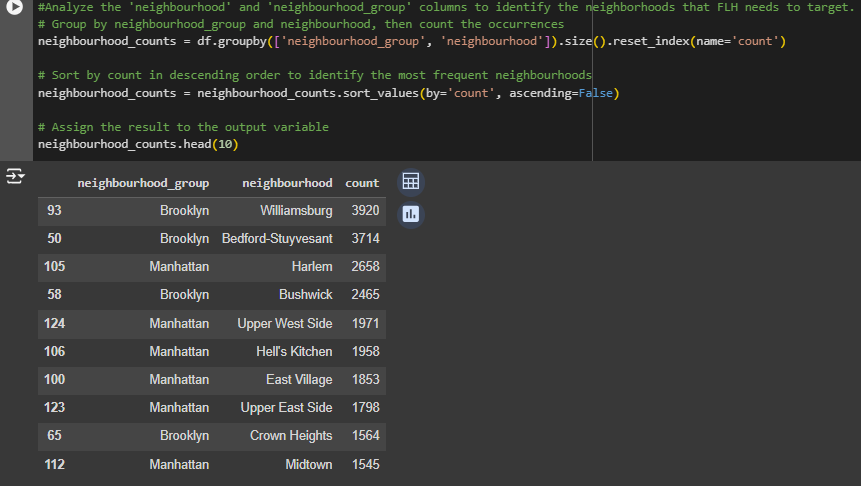


* There are very less number of people who prefers shared room,and in Staten Island many prefers Entire home/apartment.
* The price range in Manhattan are pretty high whereas the prices in Staten Island are low even most of them prefer Entire home/ apartment.

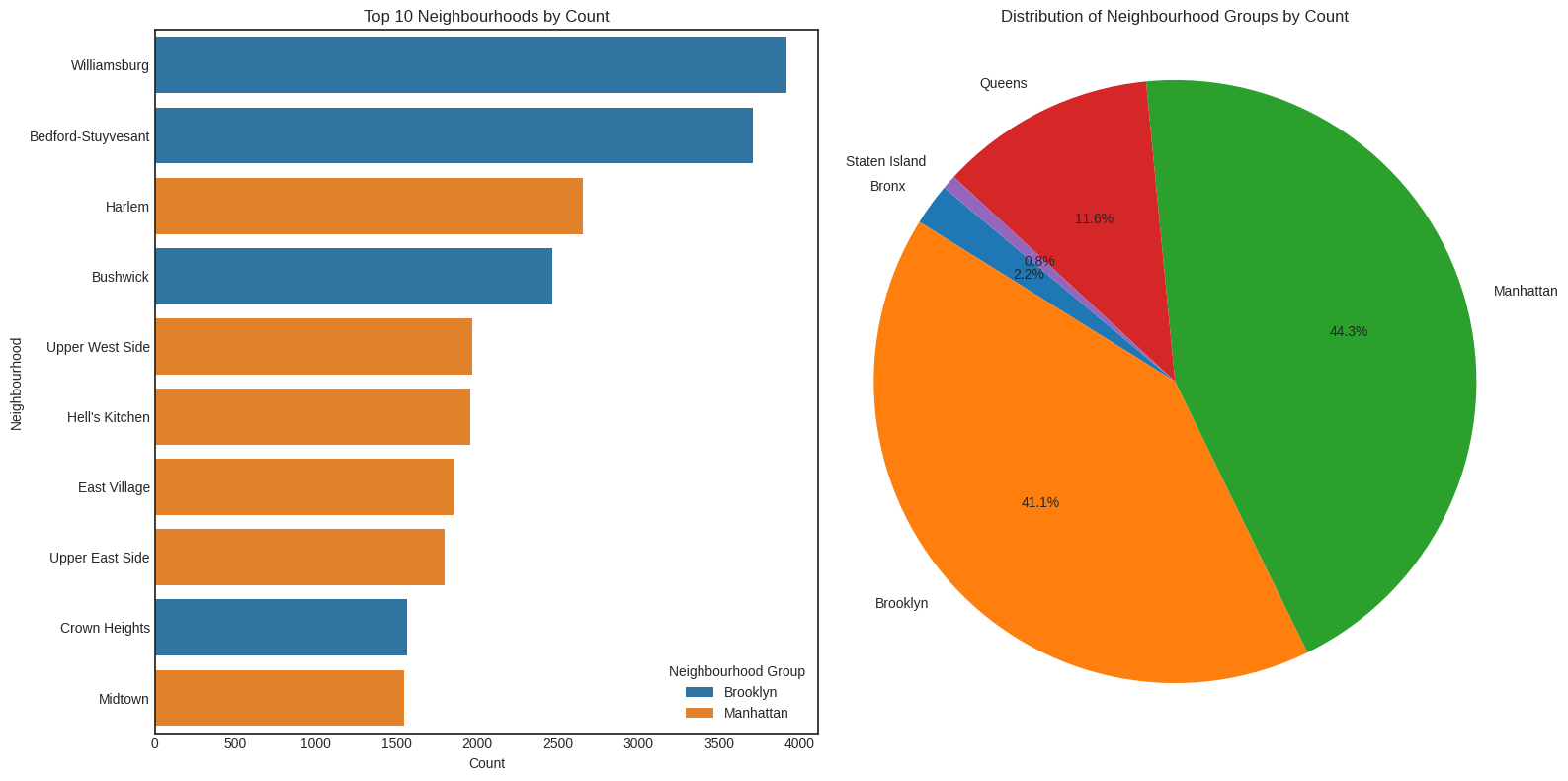
Q1 What are the neighborhoods that FLH needs to target?

Identifying neighborhoods with the highest demand can help FLH focus its marketing and expansion efforts in areas that are most likely to attract customers.

Targeting neighborhoods with the highest number of listings can help FLH focus marketing efforts where there is already a high demand. Investing in these areas with promotions and better services can further attract more customers, thereby increasing revenue.







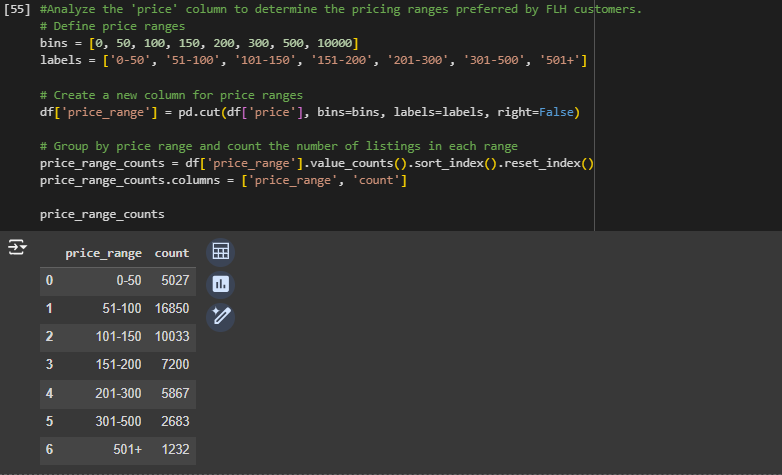
Target Neighborhoods for FLH

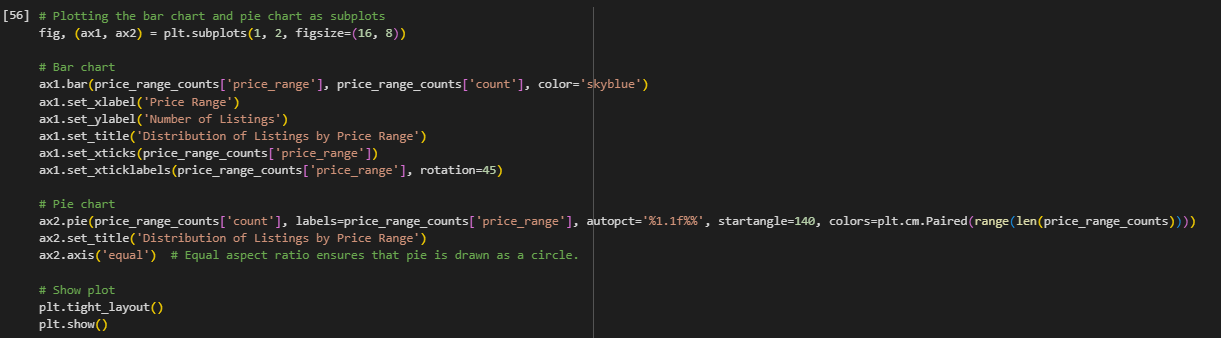
* High Priority Neighborhoods: Based on the data, FLH should focus on neighborhoods with the highest listing counts.
* 85 % of the listing are Manhattan and Brooklyn neighbourhood\_group
* Brooklyn: Williamsburg (3920), Bedford-Stuyvesant (3714), Bushwick (2465)
* Manhattan: Harlem (2658), Upper West Side (1971)

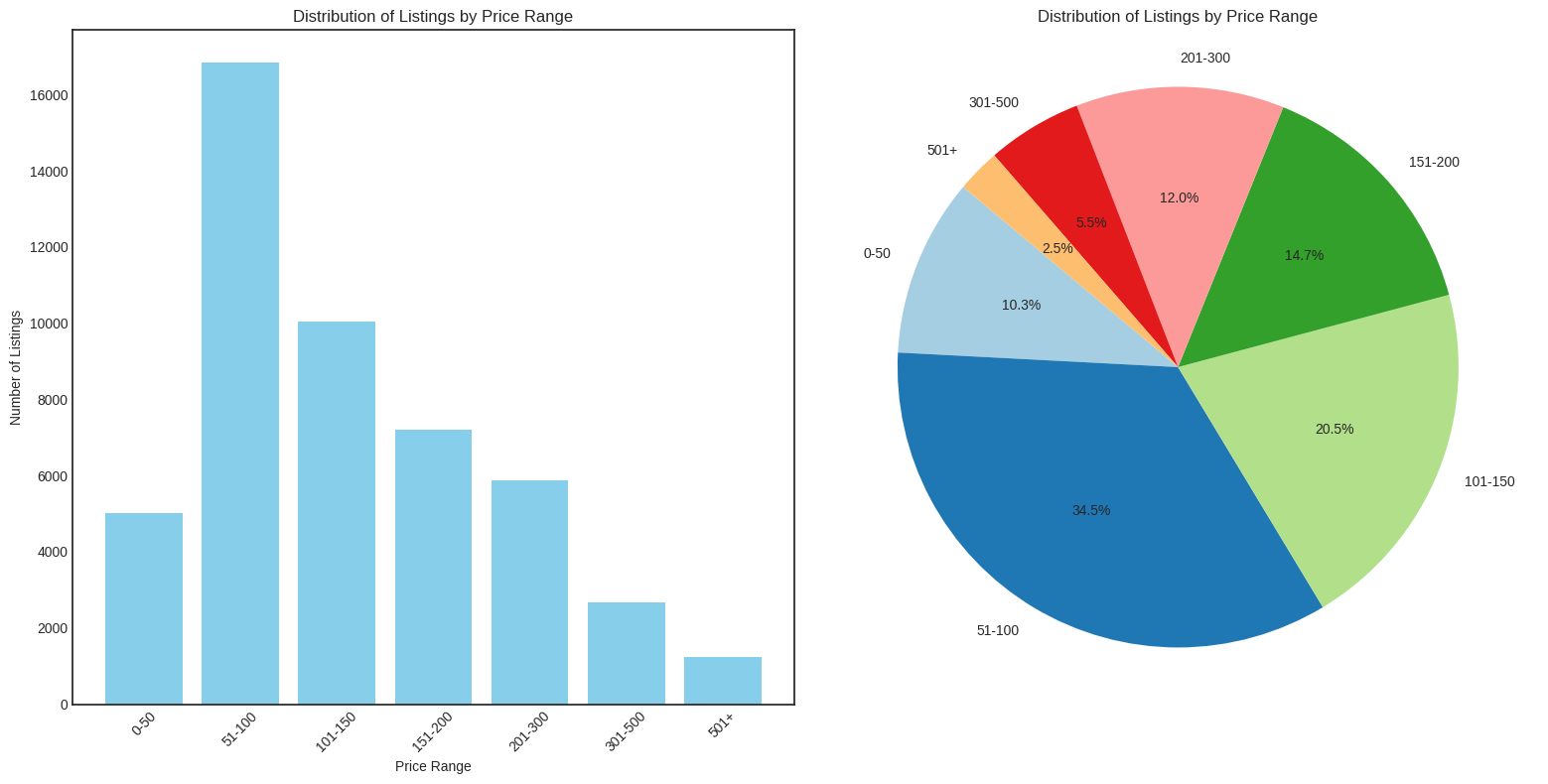
These areas have demonstrated high customer demand and should be targeted for marketing and property acquisition efforts.

Q2 What are the pricing ranges preferred by FLH customers?

Understanding preferred pricing ranges helps FLH set competitive prices that attract customers.







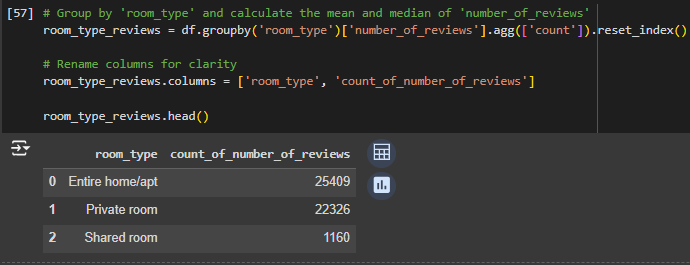
Preferred Pricing Ranges by FLH Customers

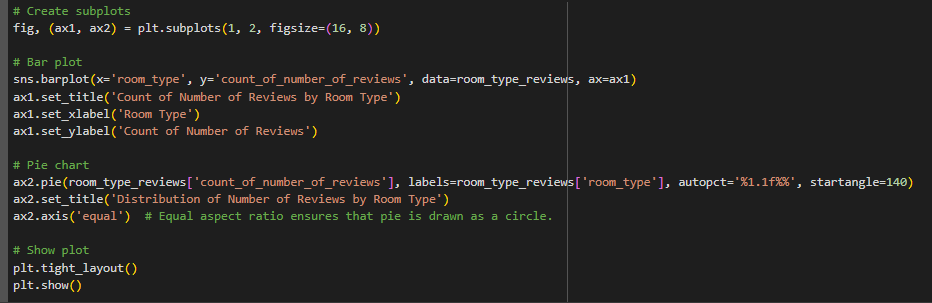
* The data shows that the majority of FLH customers prefer prices between $50 and $150 per night, with the most popular price point being $100 per night.
* Around 15% of customers are willing to pay between $150 and $200.
* Entire home/apt listings have the highest number of listings in this price range, followed by private rooms and then shared rooms.
* '$51-100': This range has the highest count of bookings (16850), indicating strong preference.
* '$101-150': Also popular, with 10033 bookings.

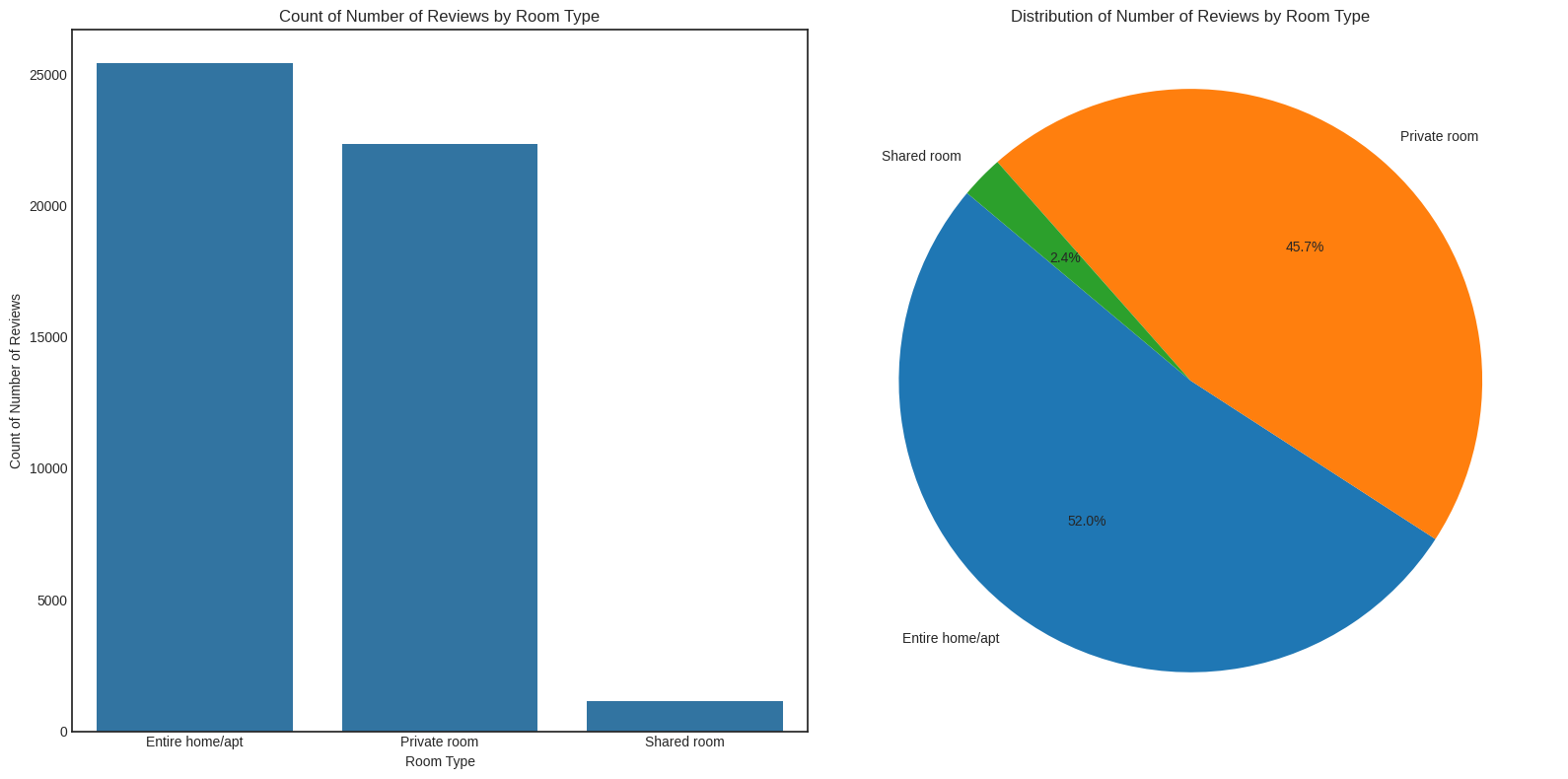
Focusing on properties within these price ranges could maximize revenue.

Q3 What are the types of properties that are most successful and least successful for FLH?

Analyze the 'room\_type' and 'number\_of\_reviews' columns to identify the types of properties that are most successful and least successful for FLH.







Identifying the most successful property types allows FLH to prioritize these in marketing and investment.

Most and Least Successful Property Types

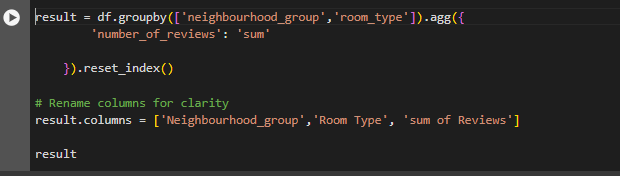
* Most Successful:

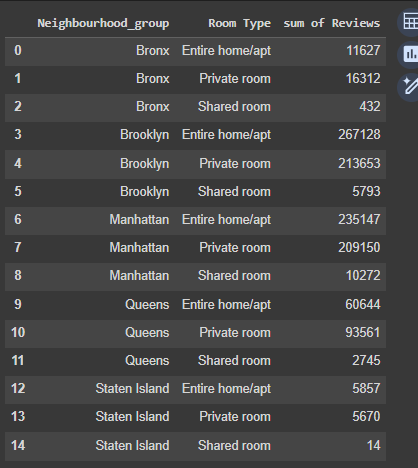
Entire home/apt have the highest count of number of reviews (25409), suggesting high customer satisfaction and repeat bookings.

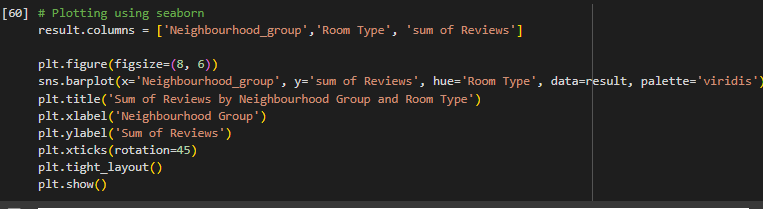
* Least Successful:

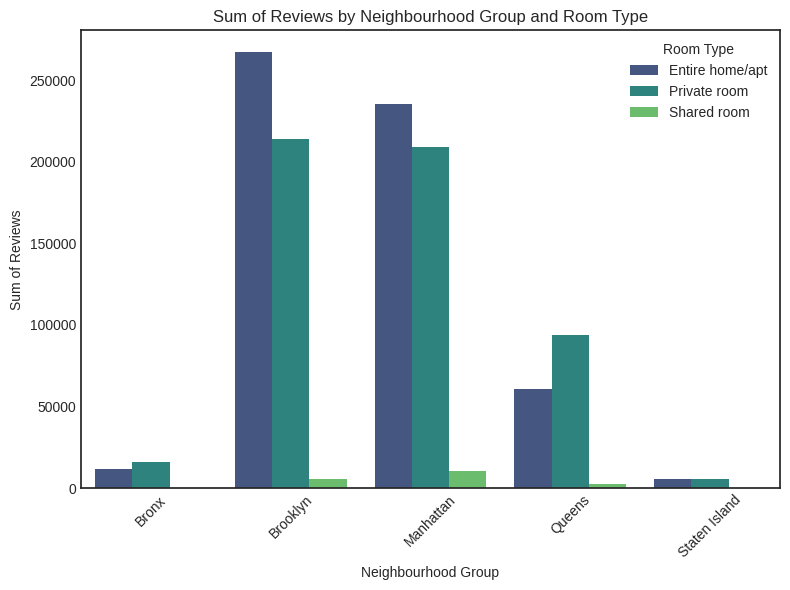
Shared Rooms show the lowest count of number of reviews (1160) indicating lower popularity or satisfaction.

Q4 What are customers looking for most in their stays with FLH?









**Neighborhood Group Popularity**:

* Brooklyn and Manhattan have the highest total number of reviews.
* Queens has a noticeable but lower number of reviews compared to Brooklyn and Manhattan.
* Bronx and Staten Island have the lowest number of reviews.

**Room Type Preferences**:

* Entire home/apt consistently has the highest number of reviews, making it the most popular room type.
* Private rooms are also popular, especially in Brooklyn and Manhattan.
* Shared rooms have the fewest reviews, indicating they are the least preferred.

**Comparison Between Neighborhoods**:

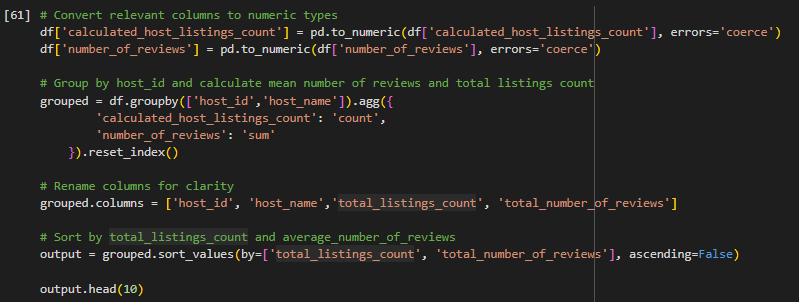
* In Brooklyn and Manhattan, reviews for Entire home/apt and Private rooms are close, with Entire home/apt leading.
* In Queens, Private rooms have more reviews than Entire home/apt, which is unique compared to other neighborhoods.

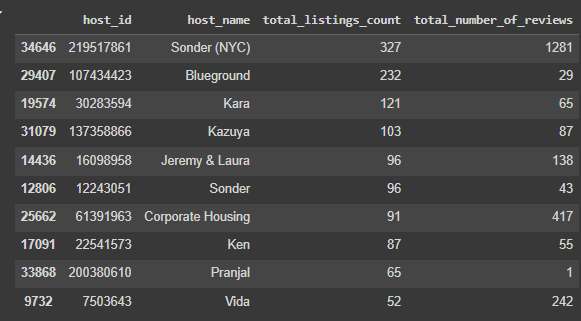
Most of the properties are offering either a Private room or the Entire home/apt, there are very few properties which are offering a Shared room. The reason for this can be that most of the tourists or visitors want privacy and prefer either a private room or to have the entire home/apt for themselves for their convenience.

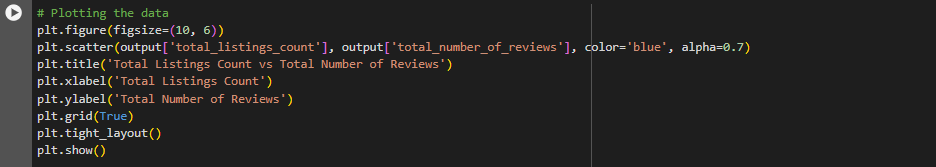
Queens shows a relatively higher interest in private rooms, which could be an opportunity for targeted marketing and improvements in that segment.

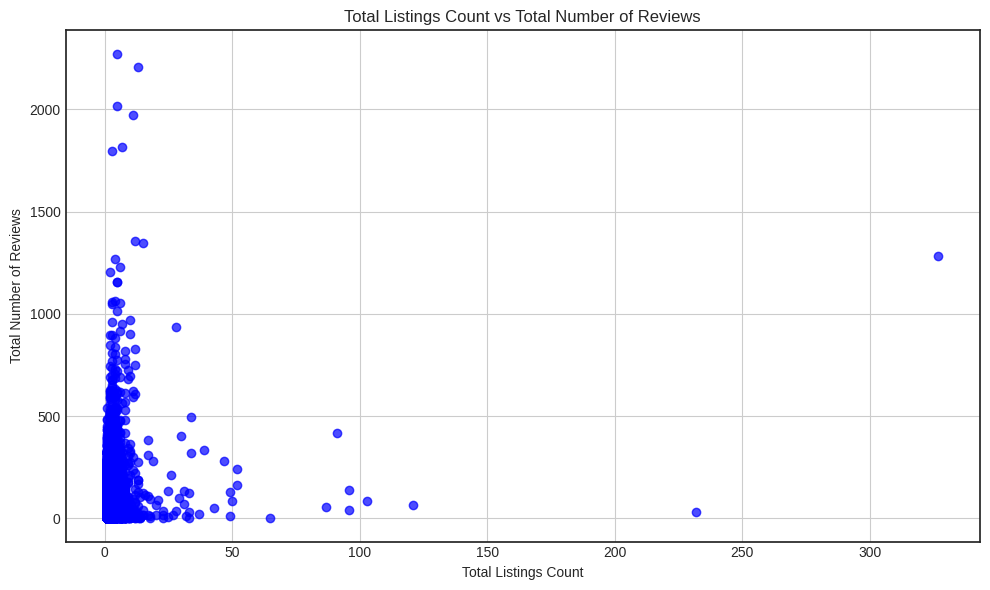
Q5 What attributes do the hosts have that FLH should target to add to their service to increase revenue?

Analyze the 'host\_name', 'host\_id', and 'calculated\_host\_listings\_count' columns to identify attributes of hosts that FLH should target to add to their service to increase revenue.





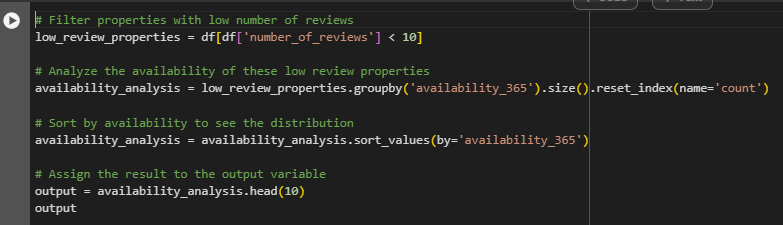


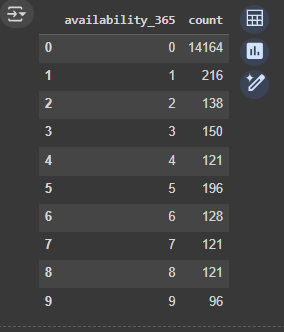


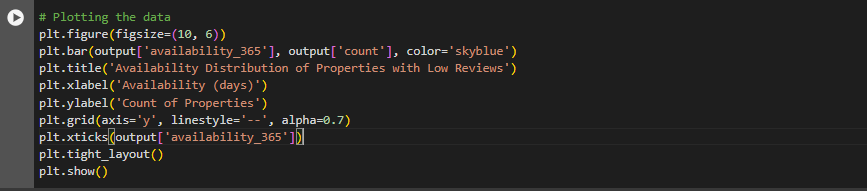
Attributes of Hosts to Target for Increasing Revenue

* Hosts with High Listing Counts: Hosts like Sonder (NYC)  (327 listings) and Blueground (232 listings) manage a large number of properties, indicating experience and capability in handling multiple listings.
* Focus on Experienced Hosts: Targeting hosts who manage multiple listings effectively can help ensure a consistent, quality experience for FLH customers, potentially increasing bookings and revenue.
* Identifying the top hosts is important as it can lead to collaborations, recognition, and potential incentives for these hosts. Recognizing and rewarding top hosts could encourage them to continue providing exceptional service, contributing to positive customer experiences and loyalty.

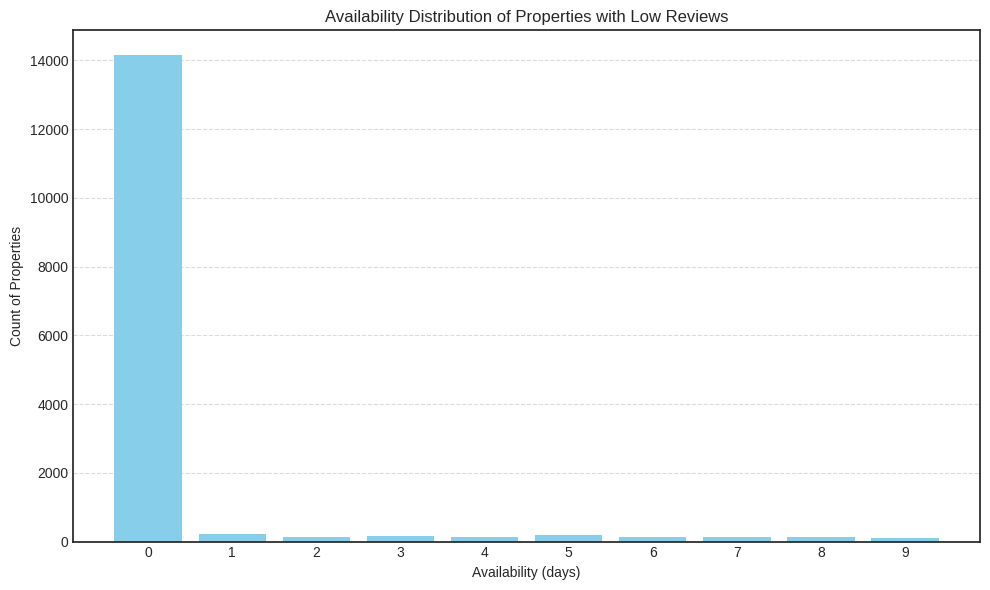
Q6 How to get unpopular properties more traction?







Observations:

1. **Highly Available Properties**: A significant number of properties (14,164) have an availability of 0 days, indicating they may not be available for rental throughout the year. This suggests these properties might have limited availability or are not actively rented out.

2. **Limited Availability**: As availability days increase from 1 to 9, the number of properties decreases, indicating that fewer properties have higher availability throughout the year. This highlights variability in rental availability among these properties.

3. **Seasonal or Periodic Availability**: The distribution shows peaks and valleys, suggesting that some properties may be available only during specific seasons or periods. This variability in availability can impact rental income potential across different times of the year.

4. **Potential Impact on Revenue**: Properties with higher availability (closer to 365 days) have the potential to generate more revenue if effectively managed. Properties with lower availability may require adjustments in pricing or marketing strategies to optimize revenue.

5. **Operational Insights**: This analysis can guide operational decisions such as adjusting cleaning schedules, implementing pricing strategies tailored to peak and off-peak periods, and forecasting revenue based on availability trends. Understanding these patterns helps in optimizing rental management and maximizing revenue from the properties.

To improve:

* Targeted Marketing: Focus marketing efforts on specific neighborhoods and room types that are identified as unpopular to boost their visibility.
* Pricing Strategy: Review and adjust pricing based on the average prices of similar properties in the same neighborhood to ensure competitiveness.
* Improving Property Features: Enhance the amenities and features of the properties based on feedback or common features of more popular properties.
* Optimizing Listing Descriptions: Use attractive and clear descriptions, high-quality photos, and update them regularly to attract more interest.

By systematically addressing these areas through data analysis and targeted strategies, FLH can effectively increase the visibility, attractiveness, and ultimately the revenue of its less popular properties.

CONNECTING MYSQL

