airbnb

AirBnb Booking Analysis (EDA)

Presented By – Piyush Shivhare (Individual)

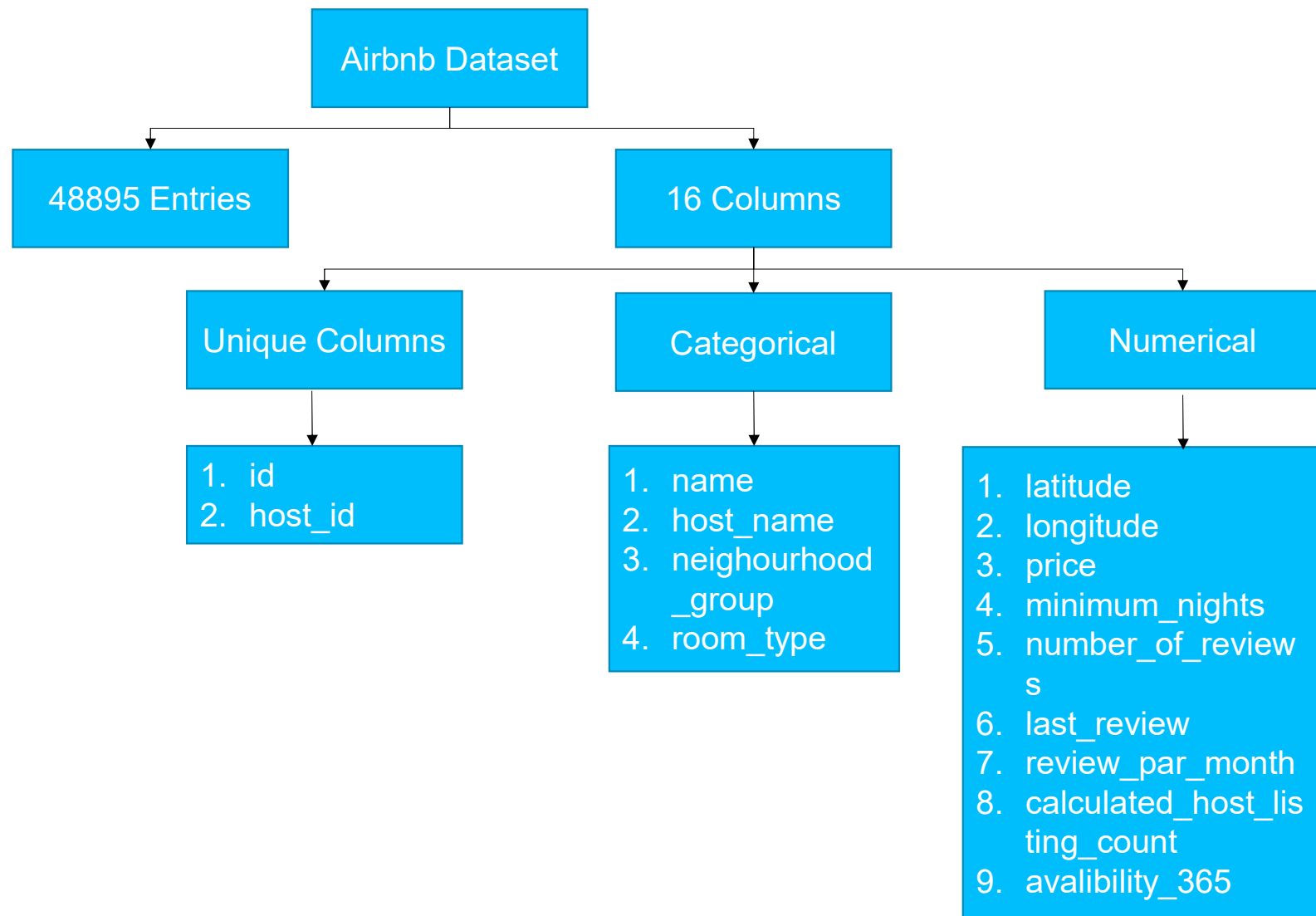Cohort – Warsow

Problem and Objective

The potential business problem that could be addressed is how to utilize the data provided by Airbnb listings to improve the platform's overall performance and profitability.

The business problem is that Airbnb needs to improve user experience and satisfaction by providing clear and accurate information about the listings available on the platform. Users have varying preferences and requirements, such as budget, location, and occupancy, and Airbnb needs to ensure that users are able to easily find listings that match their needs. By addressing this business problem, Airbnb can increase the number of bookings and repeat customers, which can ultimately lead to increased revenue and profitability.

What we will do in EDA

Collecting the Data → Exploring the Dataset → Cleaning and Transforming the Data → Visualization → Gaining Insights → Conclusion

```
                          ┌─────────────────────┐
                          │   Airbnb Dataset    │
                          └─────────────────────┘
                    ┌───────────────┴───────────────┐
          ┌──────────────────┐            ┌──────────────────┐
          │  48895 Entries   │            │   16 Columns     │
          └──────────────────┘            └──────────────────┘
```

┌──────────────────┐   ┌──────────────────┐   ┌──────────────────┐
│ Unique Columns │   │   Categorical   │   │   Numerical   │
└──────────────────┘   └──────────────────┘   └──────────────────┘

┌──────────────────┐   ┌──────────────────┐   ┌──────────────────┐
│ 1. id   │   │ 1. name   │   │ 1. latitude   │
│ 2. host_id   │   │ 2. host_name   │   │ 2. longitude   │
└──────────────────┘   │ 3. neighourhood│   │ 3. price   │
                         │   _group   │   │ 4. minimum_nights│
                         │ 4. room_type   │   │ 5. number_of_review│
                         └──────────────────┘   │   s   │
                                                  │ 6. last_review │
                                                  │ 7. review_par_month│
                                                  │ 8. calculated_host_lis│
                                                  │   ting_count │
                                                  │ 9. avalibility_365│
                                                  └──────────────────┘

Column Attributes

id
The identity number of the listed property.

name
Name of the listed property.

host_id
The identity number of the host who listed the property.

host_name
Name of the host who listed the property

neighbourhood_group
Aria of the Particular Location.

neighbourhood
location

latitude
Latitude range

longitude
Longitude range

room_type
Type of the room listed

price
Rant of the room

minimum_nights

Minimum Number of nights the customer ranted the property

number_of_reviews

Number of reviews by customer

last_review

Last review date

reviews_per_month

The number of reviews per month that were properly received.

calculated_host_listings_count

Number of listings done by a host

availability_365

The number of days available in a year.

- There are 16 columns and no duplicate values in the dataset.
- Some columns have missing or null values, including name, host_name, last_review, and reviews_per_month.

✓ Some columns have missing or null values:
✓ Name: 16 null values (Replaced with 'Unknown')
✓ Host_name: 21 null values (Replaced with 'Unknown')
✓ Last_review: 10052 null values (Replaced with '0')
✓ Reviews_per_month: 10052 null values Replaced (Replaced with '0')

Missing Values

✓ The availability column has a minimum value of 0 for 17533 entries.
✓ This means that around 17533 entries have an extreme case where the apartments have 0 availability days.
✓ We will not change this column to avoid manipulating apartments that are usually busy.

Availability Column

✓ There are a few values with 0 in the 'price' column.
✓ This could be due to dynamic pricing or the host's decision not to share the price with Airbnb.
✓ A price of $10,000 or higher may not necessarily be an outlier and cannot be discarded without analyzing the context.

Price Column

Top 10 Hosts Having Highest Numbers of Apartments

Number of Apartment according to host_id
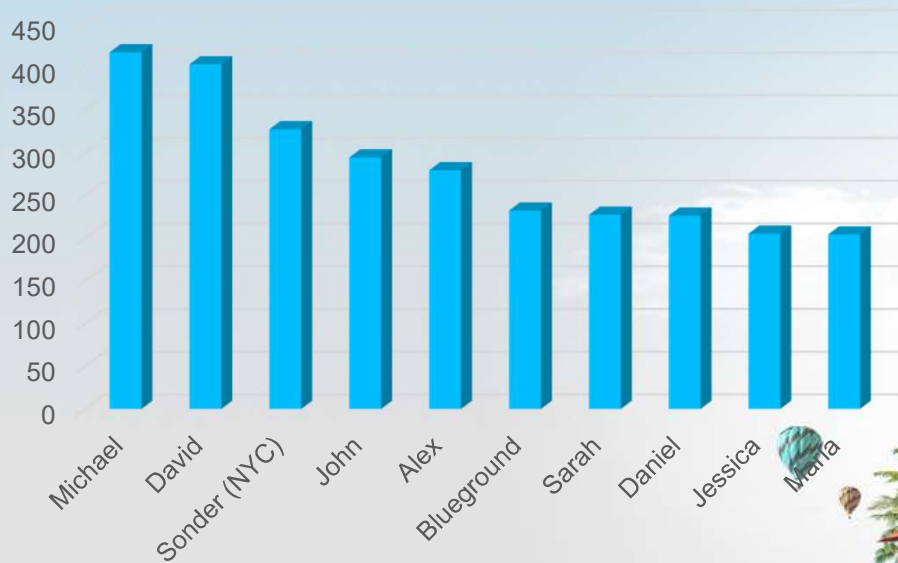
Number of Apartment according to host_id
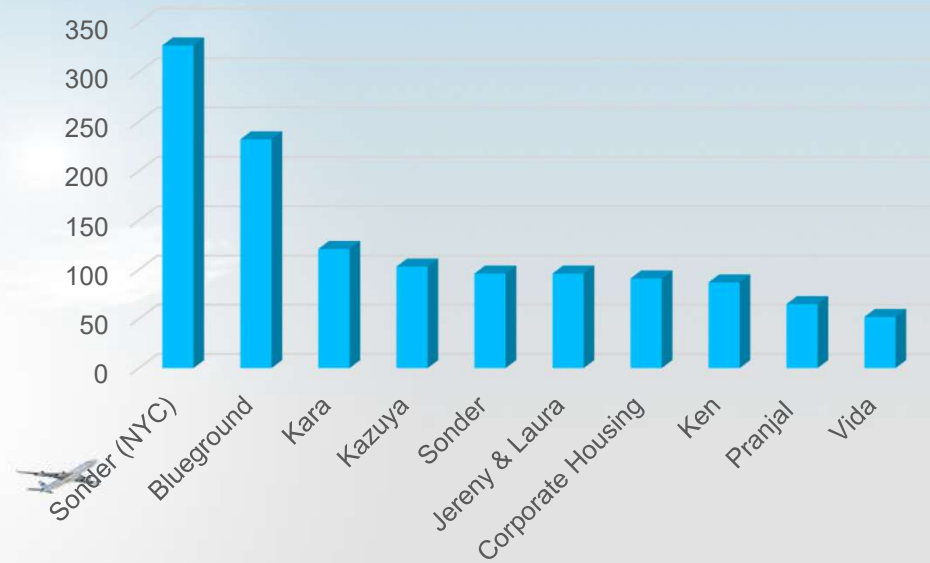
Number of listings in a Particular Neighbourhood Group

- Pie chart showcasing the distribution of Airbnb listings by neighborhood group in NYC in 2019.
- Manhattan and Brooklyn are the two most popular neighborhoods, accounting for 44.4% and 41.3% of listings respectively.
- Queens, Bronx, and Staten Island account for the remaining 15.3% of listings.
- Analysis of neighborhood popularity can help Airbnb and its users make informed decisions about marketing efforts, pricing, and location choices.
- Understanding neighborhood popularity can enhance the user experience and increase business growth and profitability.

1091 373
5666
21661
20104

■ Manhattan ■ Brooklyan ■ Qweens ■ Bronx ■ Staten Island
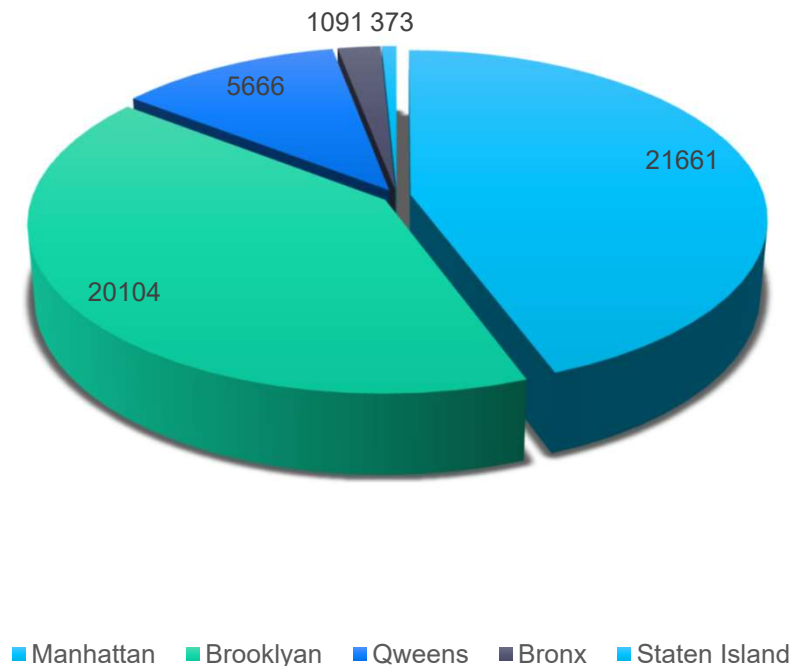
Room Type in each Naighbourhood Group

- Chart shows the number of listings for each room type (entire home/apt, private room, and shared room) in NYC.

- Brooklyn has the highest number of private room listings with 10,132, followed by Manhattan with 7,982 private room listings.

- Manhattan has the highest number of entire home/apt listings with 13,199, followed by Brooklyn with 9,559 entire home/apt listings.

- Shared room listings are least common, with Staten Island having the least number of listings for shared rooms with only 9, followed by the Bronx with 60 shared room listings.

- These insights can help Airbnb hosts understand guests' preferences and adjust listings accordingly.

- Airbnb can use this information to tailor marketing strategies according to demand for different room types in different neighborhoods.

- Overall, these insights can have a positive impact on business growth
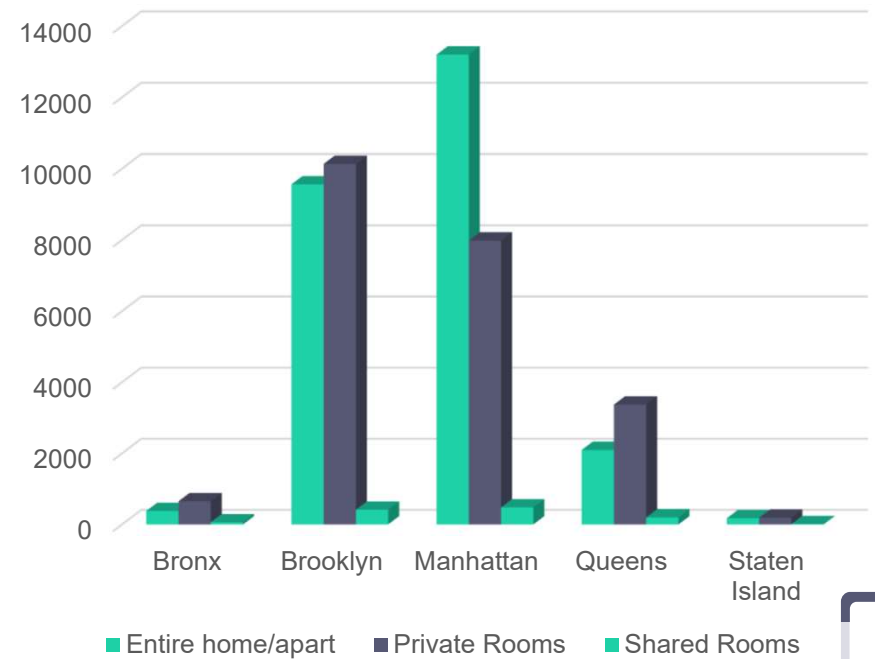
Listing of Rooms
by Room Type

- "Entire home/apt" and "Private room" are the most popular property types on Airbnb in New York City, with 25409 and 22326 listings respectively.
- "Shared room" has the lowest number of listings, with only 1160 listings in total.
- The insights gained from the chart can help hosts and businesses tailor their offerings to attract more guests and increase their revenue.
- Understanding the demand for different types of accommodations can help hosts and businesses make data-driven decisions about their listings and pricing strategies.
- By focusing on the most popular property types, hosts and businesses can increase their occupancy rates and revenue potential.
- However, hosts and businesses may need to consider offering a mix of property types to cater to a wider range of guest preferences and needs.
- Overall, the insights gained from the chart can have a positive impact on business growth and success on the Airbnb platform.

Number of Apartment

22326

25409

1160

■ Entire home/apart   ■ Shared Room   ■ Private Room   ■
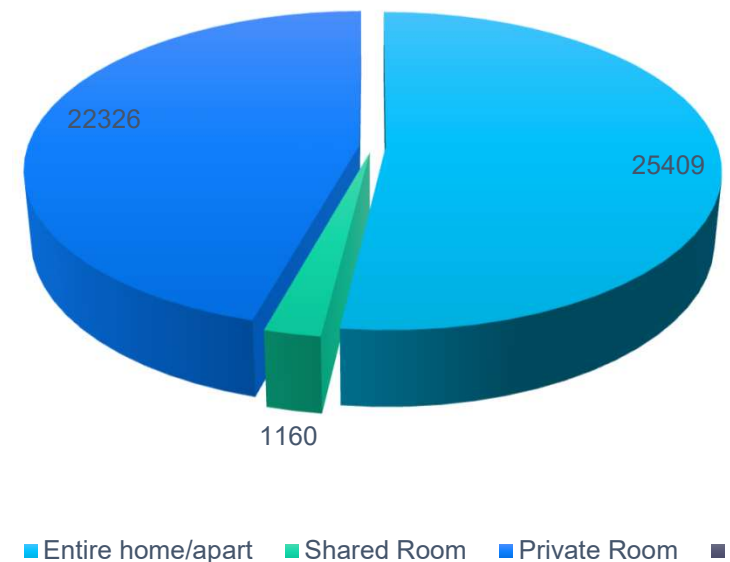
Top-10 Neighbouhoos which are having highest number of listing for Airbnb

- Williamsburg has the highest number of apartments listed on Airbnb in New York City, with a total of 3920 listings.
- The high number of listings in Williamsburg indicates a strong demand for short-term rentals in this neighborhood.
- Property owners can use this insight to invest in properties in Williamsburg to cater to the demand for short-term rentals.
- Airbnb can use this insight to develop targeted marketing strategies for Williamsburg, highlighting the high demand for short-term rentals in this neighborhood.
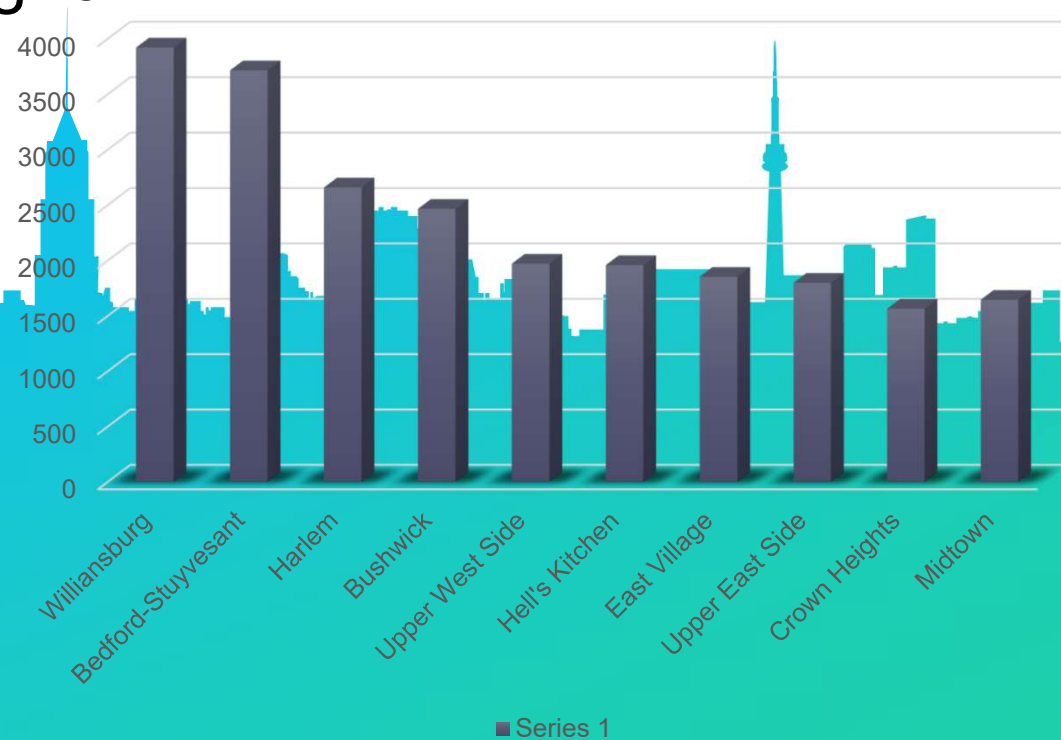
Distribution of Price

- The histogram shows that the price distribution can help hosts and property managers develop effective pricing strategies.
- The price distribution of vacation rentals is right-skewed, indicating more listings with lower prices and fewer listings with higher prices.
- Majority of the listings are priced under $500, with a significant drop-off in the number of listings above this price point.
- Some listings have very high prices, with only a small number of listings priced above $2000 per night.
- Effective pricing strategies balance affordability and competitiveness, leading to positive business growth.
- However, pricing should not be solely based on outliers; other factors such as location, amenities, and guest preferences should also be considered.
- Setting the price too high may result in difficulties attracting guests, negative reviews, and negative business impact.
- In conclusion, analyzing the price distribution is useful in developing pricing strategies, but a holistic approach is necessary for sustainable business growth.
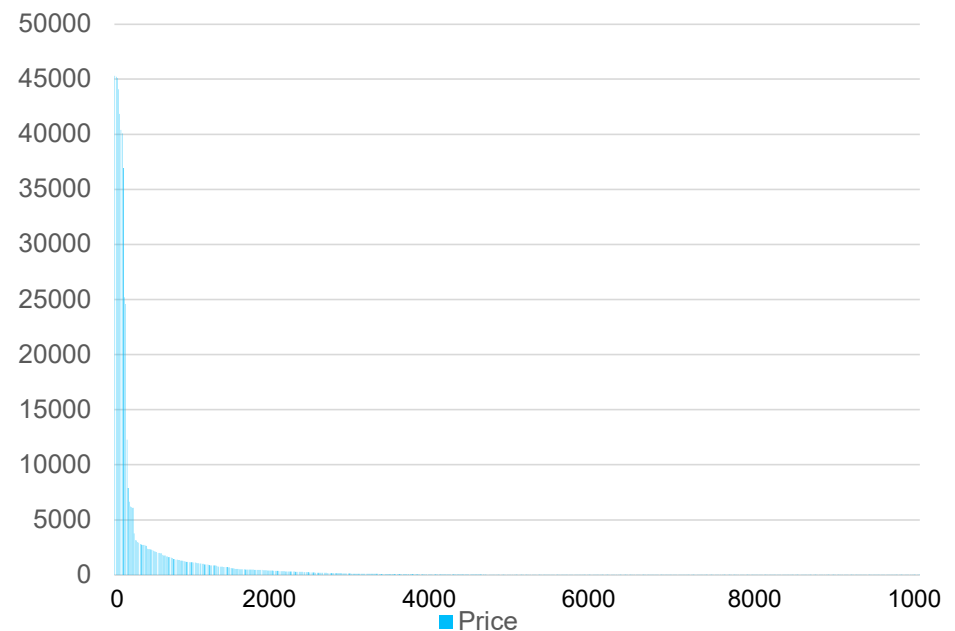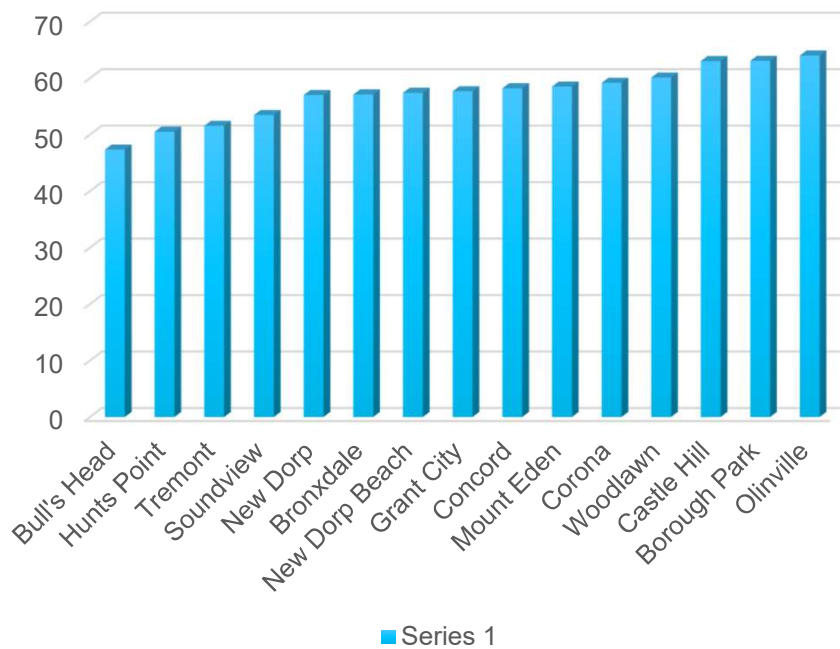
Most and Least Expensive Airbnb Neighborhoods

Expansive

- The most expensive neighborhood in the Airbnb listings is Fort Wadsworth with an average price of around 800 dollars.
- Other expensive neighborhoods include Woodrow, Tribeca, Sea Gate, Riverdale, and Prince's Bay.
- Most of the expensive neighborhoods are located in Fort Wadsworth.
- This insight could be useful for travelers looking for high-end accommodations and Airbnb hosts interested in renting out properties in these expensive neighborhoods.

■ Series 1

Chart: Series 1 (bar chart of average listing prices by neighbourhood)

Y-axis: 0, 10, 20, 30, 40, 50, 60, 70

X-axis: Bull's Head, Hunts Point, Tremont, Soundview, New Dorp, Bronxdale, New Dorp Beach, Grant City, Concord, Mount Eden, Corona, Woodlawn, Castle Hill, Borough Park, Olinville

Legend: ■ Series 1
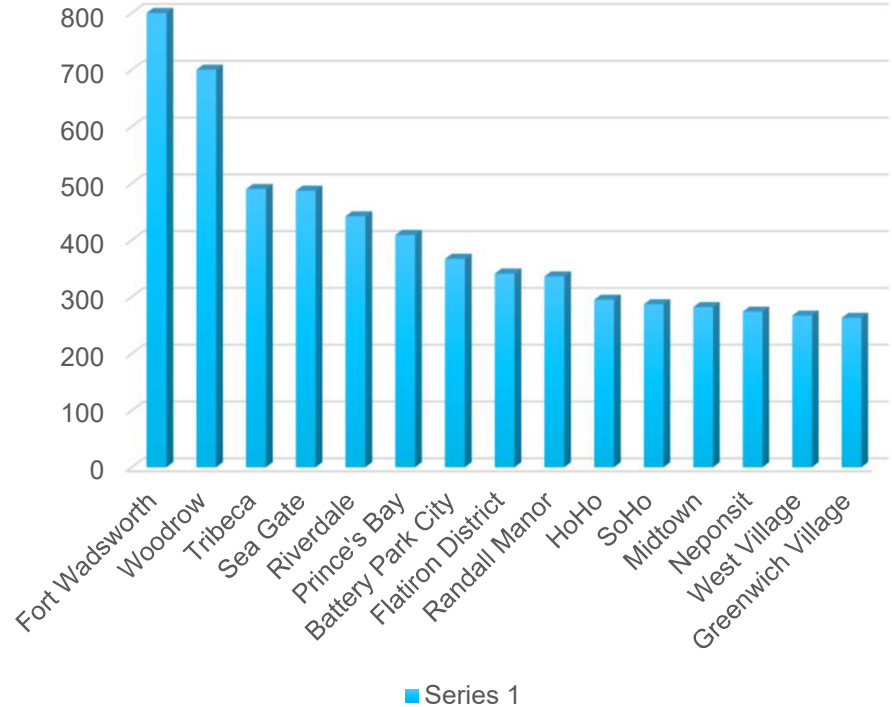
- The neighbourhoods with the lowest average prices for listings in New York City are Bull's Head, Hunts Point, Tremont, Soundview, and New Dorp.

- These neighbourhoods are relatively less expensive than other neighbourhoods in New York City and might be a good option for budget travelers.

- This chart can provide insights for budget travelers or those looking for affordable accommodation options.

- The chart can also be useful for hosts who want to set competitive prices and attract budget-conscious travelers.

- Overall, understanding the pricing trends in different neighbourhoods can help hosts and travelers make informed decisions and create positive business impacts.

Cheap

Price VS Number of Reviews

- The scatter plot on this slide represents the relationship between the number of reviews and price for Airbnb listings in New York City.
- The x-axis shows the number of reviews for each listing, while the y-axis represents the price of the listing.
- There is a weak negative correlation between price and the number of reviews, which means that as the price of the listing increases, the number of reviews tends to decrease slightly.
- However, there are also many listings with high prices and high numbers of reviews, indicating that price is not the only factor that affects the number of reviews.

price
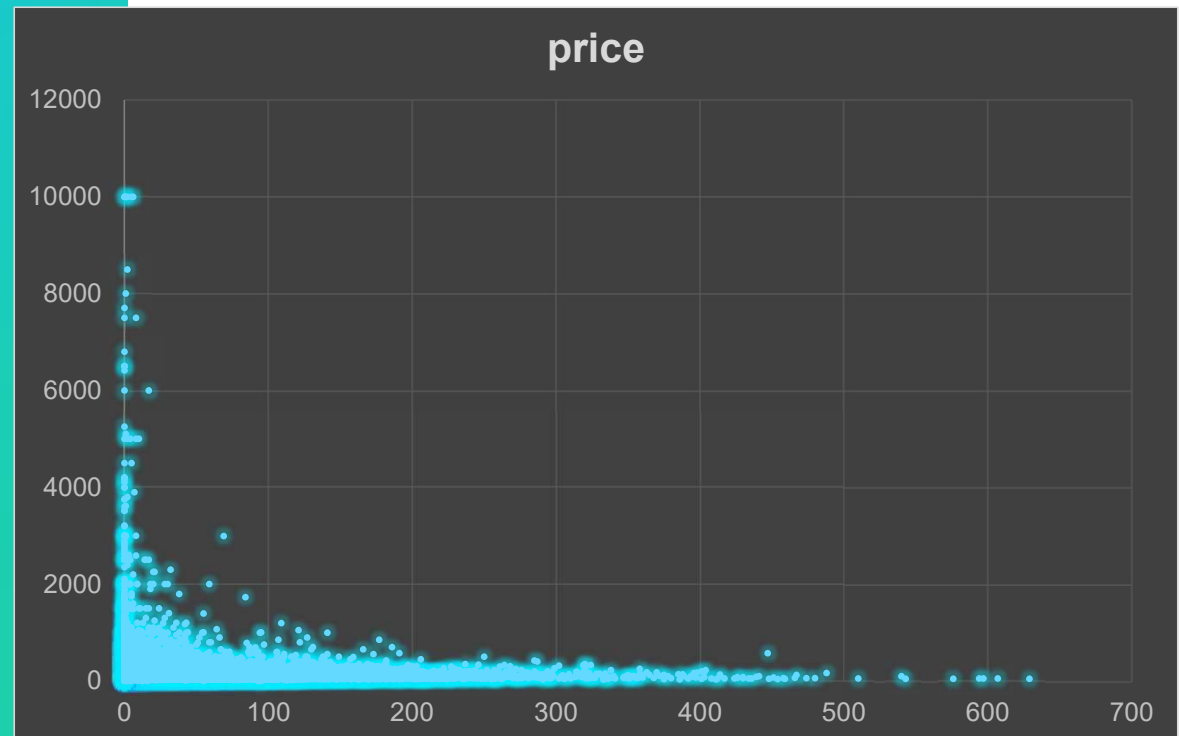
# Neighbourhood Group and Number of Reviews

Series 1
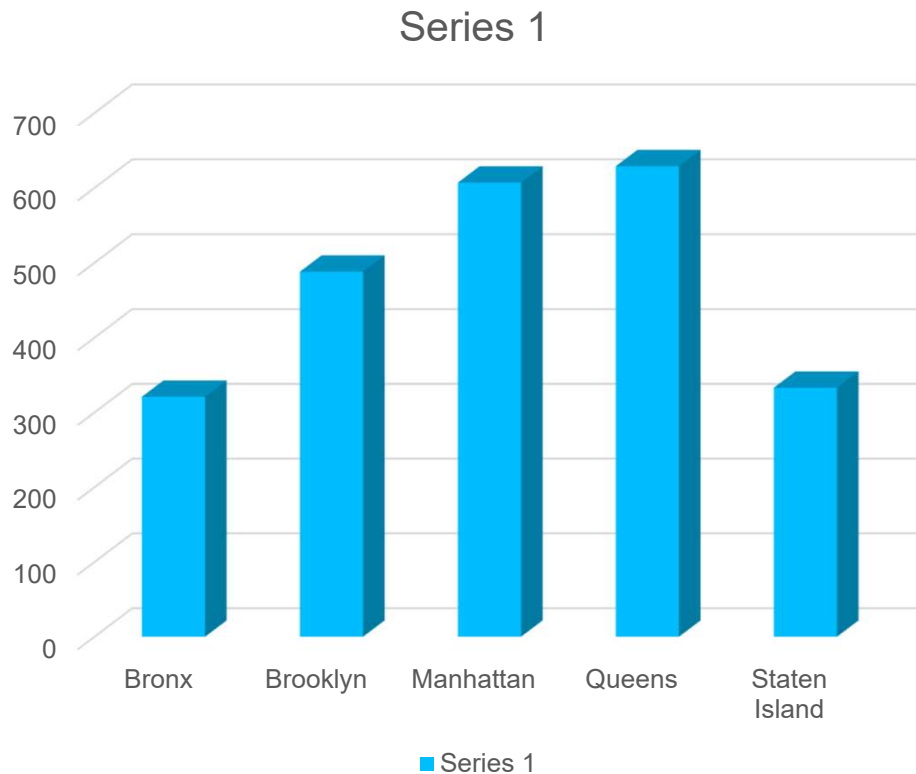
- The Queens neighborhood group has the highest number of reviews among all the neighborhood groups, followed by Manhattan, Brooklyn, Staten Island, and the Bronx.
- This indicates that Brooklyn and Manhattan have a higher number of Airbnb listings and are more popular among guests, leading to more reviews.
- These insights can help Airbnb to identify areas with high demand and interest from customers and focus on marketing and promoting listings in these areas.
- This can potentially attract more hosts to list their properties, leading to an increase in the number of bookings and revenue for both Airbnb and hosts.
- By analyzing reviews per neighborhood group, Airbnb can develop effective marketing strategies and drive positive business growth.
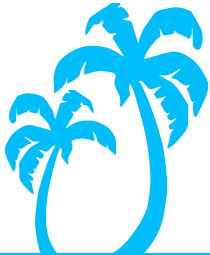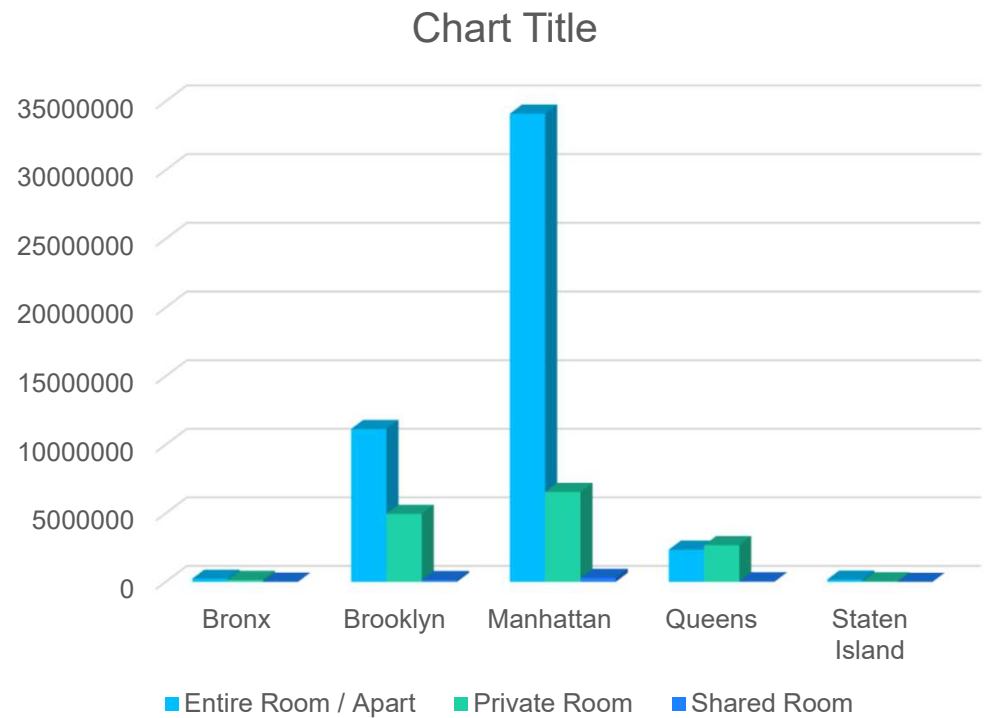
Average Price of Each Room Type in Neighbourhood Group
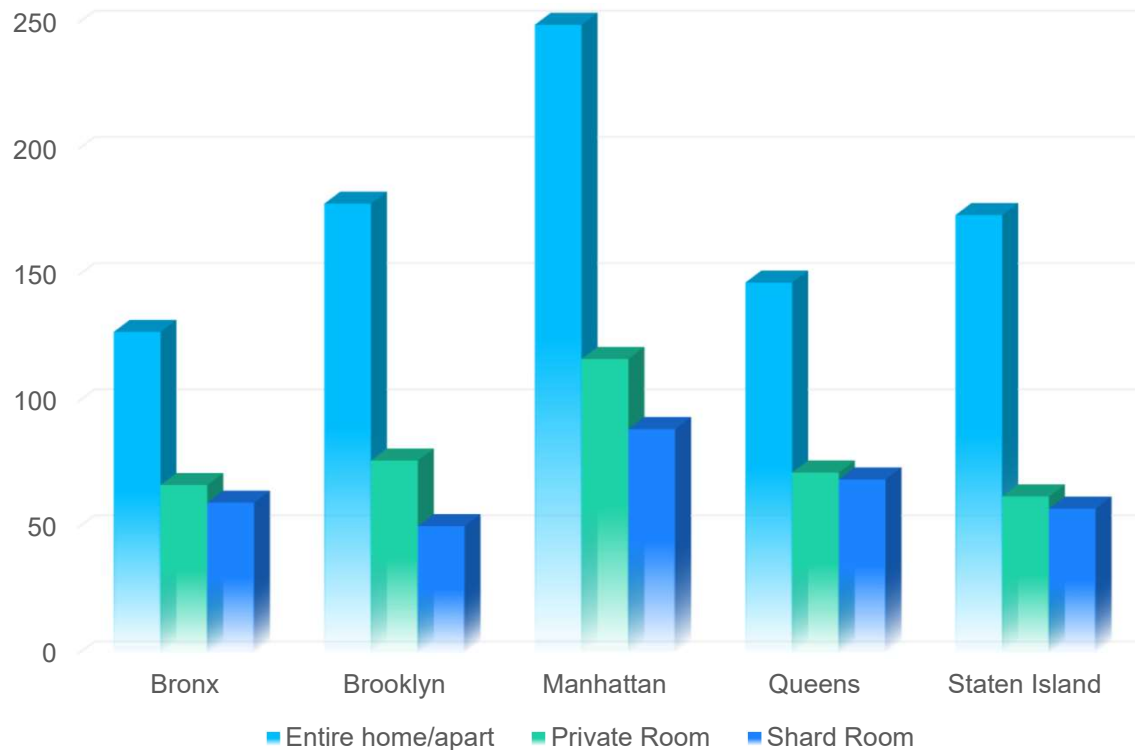
- The average price of an entire home/apt is the highest across all neighborhood groups, with Manhattan having the highest average price of $249.24, followed by Brooklyn with an average price of $178.33.

- Private rooms have a lower average price compared to entire homes/apt, with Manhattan having the highest average price of $116.78, followed by Brooklyn with an average price of $76.50.

- Shared rooms have the lowest average price across all neighborhood groups, with Manhattan having the highest average price of $88.98, followed by Queens with an average price of $69.02.

- Among all neighborhood groups, the Bronx has the lowest average price for entire homes/apt, private rooms, and shared rooms.

Revenue Generated

- Manhattan generates the highest revenue among all neighborhoods, followed by Brooklyn and Queens, while Staten Island and the Bronx generate the least amount of revenue.
- Entire homes/apartments generate the highest revenue across all neighborhoods, followed by private rooms and shared rooms.
- This suggests that hosts may be able to generate more revenue by renting out entire homes/apartments, rather than individual rooms.
- These insights can be useful for Airbnb hosts and businesses to identify neighborhoods and property types that are most lucrative for generating revenue.
- By focusing on these neighborhoods and property types, hosts can develop effective pricing strategies and maximize their earnings, leading to a positive impact on business growth.

Chart Title

Legend: ■ Entire Room / Apart ■ Private Room ■ Shared Room

Categories: Bronx, Brooklyn, Manhattan, Queens, Staten Island

Y-axis: 0, 5000000, 10000000, 15000000, 20000000, 25000000, 30000000, 35000000

The Correlation Heatmap is a visual representation of the correlation between different numerical features in the dataset.

- Positive correlation = Red
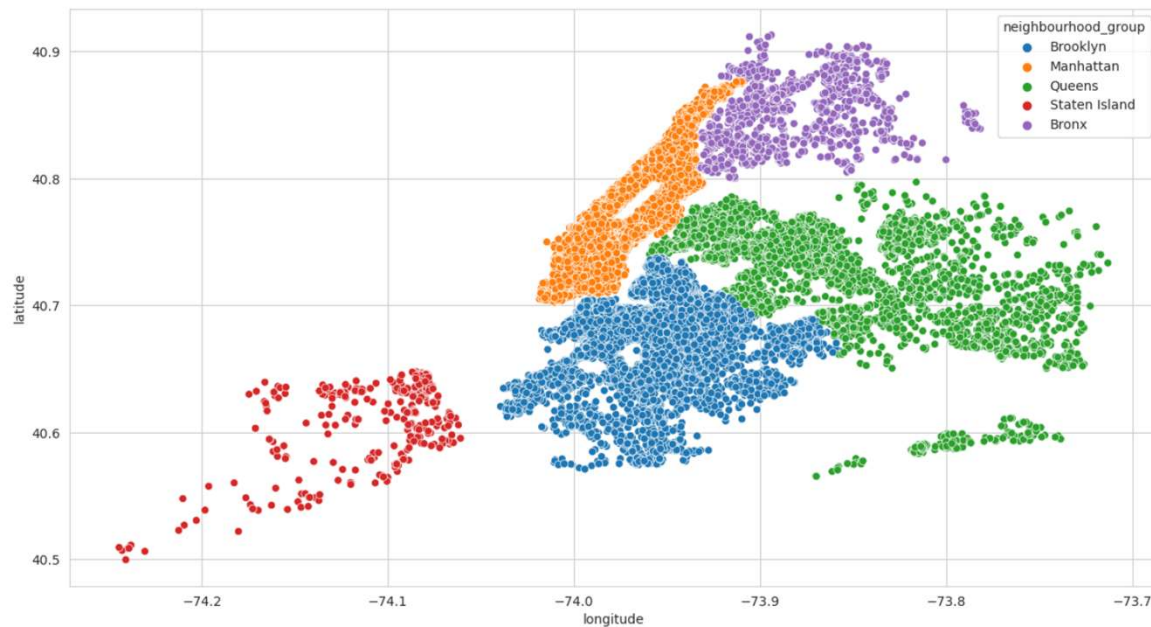- Negative correlation = Blue

The annotation in each cell shows the correlation coefficient value.

This information can be useful for understanding how different features are related to each other and identifying potential patterns or relationships in the data.

Correlation Heatmap

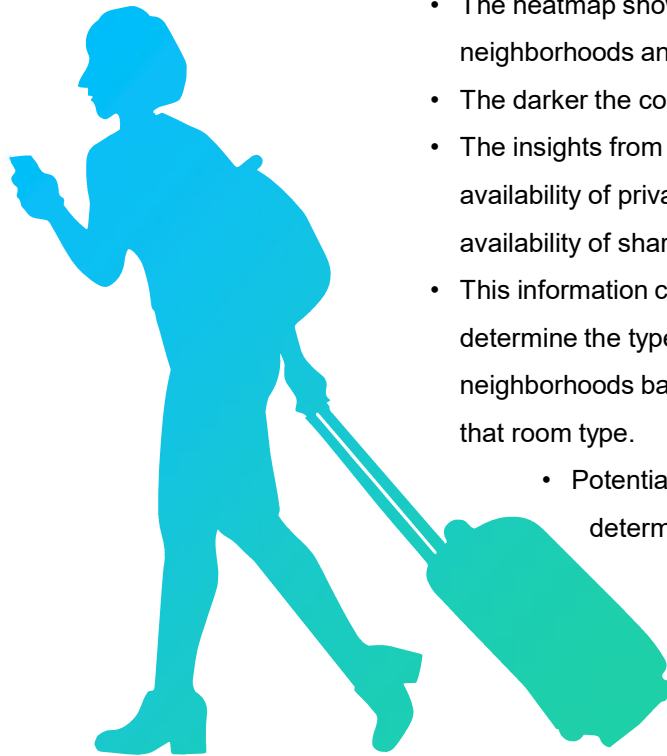

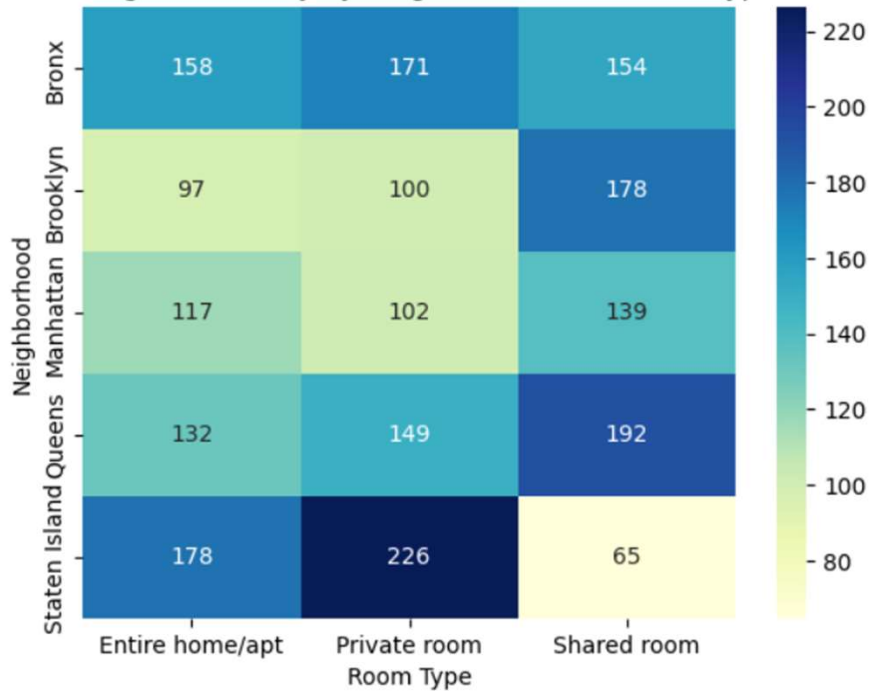Correlation Heatmap of Airbnb NYC Dataset
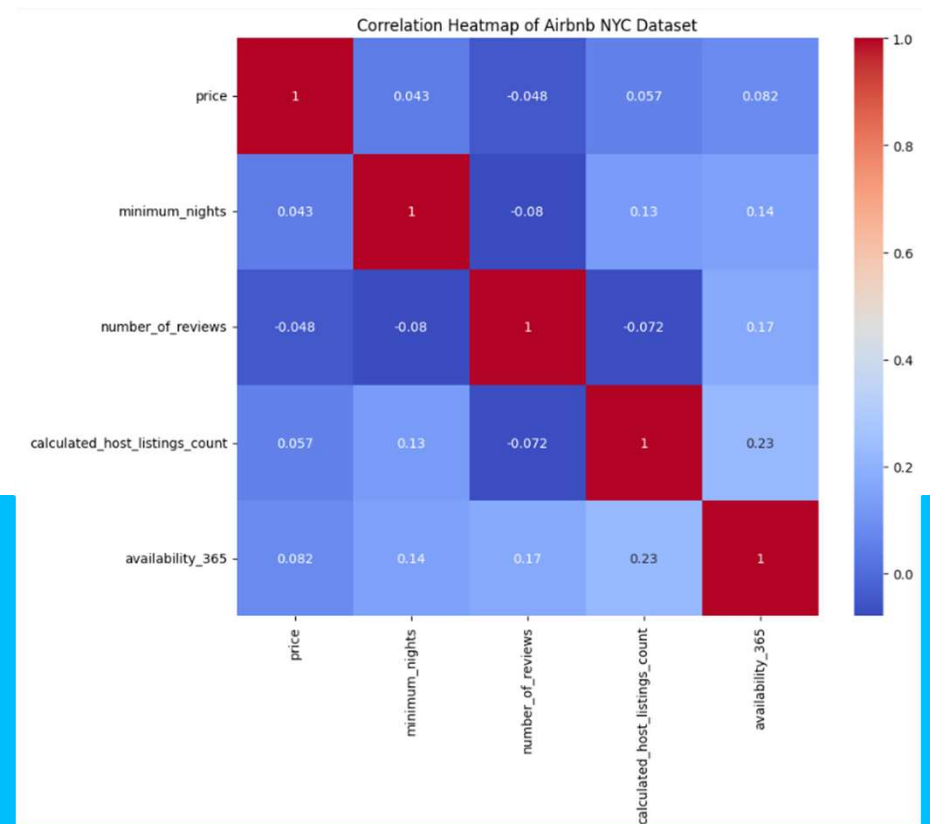
Location of Apartment

- The scatter plot shows the distribution of Airbnb listings across the NYC map, with each dot representing a listing and the color of the dot corresponding to the neighborhood group.
- Majority of the Airbnb listings are concentrated in Manhattan and Brooklyn, with fewer listings in the Bronx, Queens, and Staten Island.
- Listings in Manhattan tend to be more concentrated in the lower part of the island, while listings in Brooklyn are more spread out.
- This information can be useful for both hosts and guests, as it can help hosts decide where to list their properties and guests choose where to stay based on their preferences and budget.
- The chart provides a visual representation of the distribution of Airbnb listings across NYC, which can help hosts and guests make informed decisions.
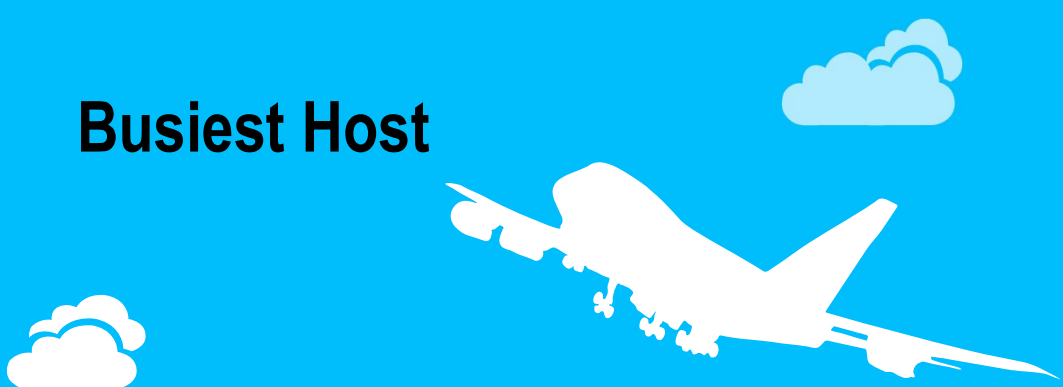
Availability

Average Availability by Neighborhood and Room Type

Neighborhood	Entire home/apt	Private room	Shared room
Bronx	158	171	154
Brooklyn	97	100	178
Manhattan	117	102	139
Queens	132	149	192
Staten Island	178	226	65

Room Type

- The heatmap shows the availability of listings in different neighborhoods and room types.
- The darker the color, the more listings available.
- The insights from the chart indicate that there is high availability of private rooms in Staten Island and less availability of shared rooms in Staten Island.
- This information can be useful for Airbnb hosts to determine the type of room to offer in different neighborhoods based on the demand and availability of that room type.
  - Potential guests can also use this information to determine which neighborhoods have a higher likelihood of finding available rooms.
  - These insights can help hosts to optimize their listing offerings and potentially increase bookings and revenue.

Busiest Host

Host Name	Room Type	Number of Reviews
Dona	Private Room	629
Jj	Private Room	607
Maya	Private Room	543
Carol	Private Room	540
Danielle	Private Room	510
Asa	Entire home/apt	488
Wanda	Private Room	480
Linda	Private Room	474
Dani	Entire home/apt	467
Angela	Private Room	466

- The busiest Airbnb hosts based on their listed room types are Dona, Ji, Maya, Carol, and Danielle.
- These hosts have listed their rooms as Entire home and Private room, which are popular choices among the majority of Airbnb guests.
- These insights can help Airbnb optimize their business strategy by focusing on recruiting more hosts who offer entire homes or private rooms, which are in high demand among guests.
- However, there could be potential negative impact if Airbnb becomes too reliant on these hosts and does not diversify its offerings.
- Airbnb will need to balance the popularity of these room types with other offerings on their platform to ensure sustained growth and customer satisfaction.
- By understanding customer preferences and optimizing their offerings, Airbnb can drive positive business impact and growth.

Solution

- Increase availability of listings in neighborhoods with high demand.
- Optimize pricing based on availability using dynamic pricing strategies.
- Focus on recruiting more hosts who offer entire homes or private rooms.
- Expand into underserved neighborhoods to increase supply and meet customer needs.
- By implementing these strategies, Airbnb can increase revenue in NYC by increasing supply, optimizing pricing, and improving the user experience for both guests and hosts.
- This can help Airbnb stay competitive in the market and achieve its business objectives.

Conclusion

- The charts provide various insights useful for both Airbnb hosts and guests.
- Hosts can determine where to list their properties and what types of rooms to offer based on demand and availability, while guests can choose where to stay based on their preferences and budget.
- The charts also provide some indication of potential correlations between variables, such as the slight negative correlation between price and availability in 365 days.
- These insights can help hosts and guests make more informed decisions about pricing and availability.
- Overall, these insights can help Airbnb optimize its business strategy, by focusing on recruiting more hosts who offer popular room types, while also diversifying its offerings to mitigate potential negative impacts.
- By leveraging the insights gained from these charts, Airbnb can improve its platform, increase revenue, and enhance the overall user experience for both hosts and guests.
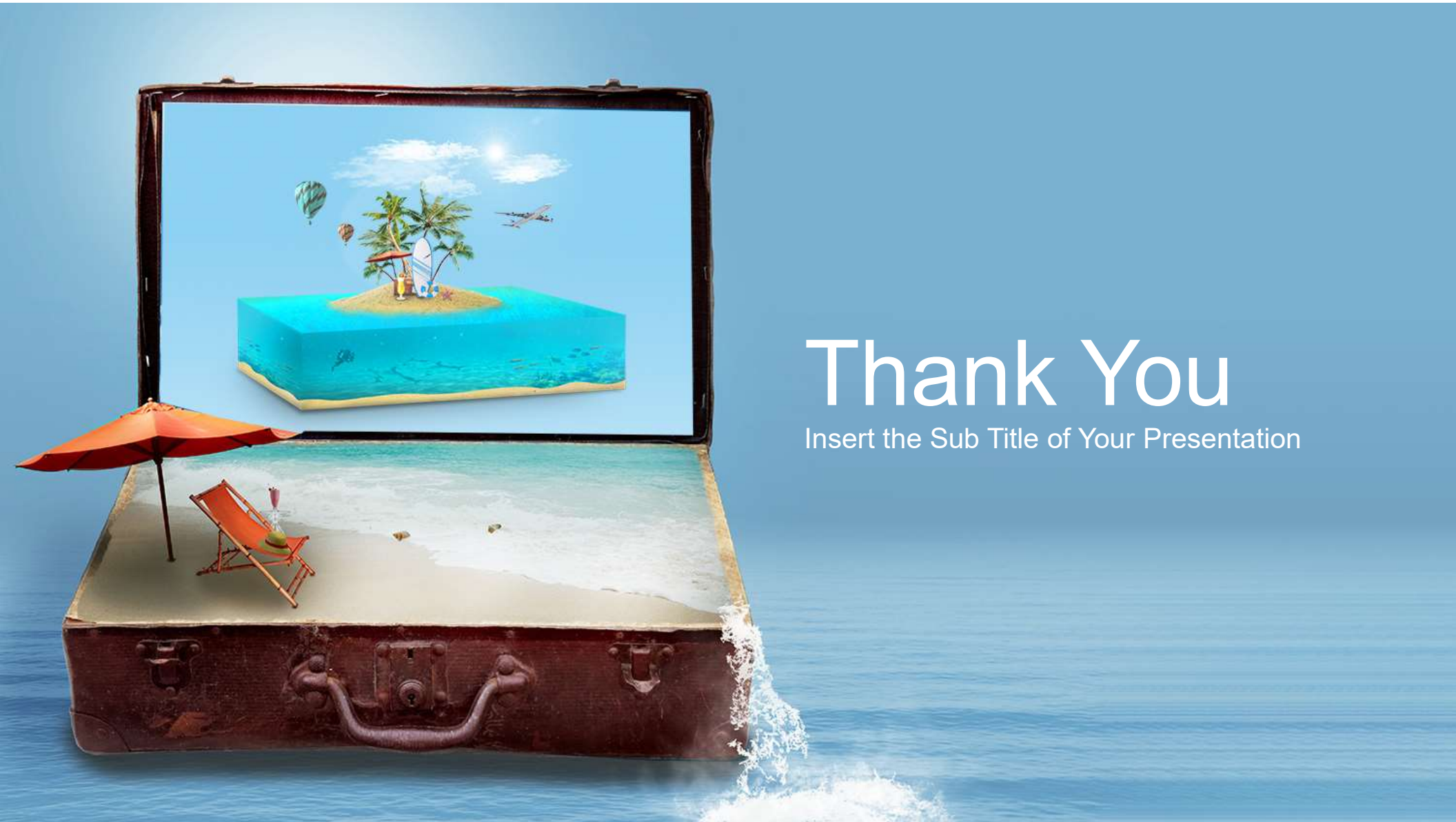
Thank You

Insert the Sub Title of Your Presentation