Airbnb, NYC Case study

- **Presented By :**

1. Chetan Tippa
2. Sphurti Gaikwad
3. Vaishnavi Pande

Agenda

- Business understanding/ Background
- Objectives
- Key Findings
- Recommendations
- Appendix
    1. Data sources
    2. Data cleaning
    3. Data methodology

➢Business understanding / background :

The revenue of Airbnb has significantly decreased during the last few months.

Now that the restrictions have started to lift and more people are beginning to travel, Airbnb wants to make sure that the business is fully prepared for this shift.

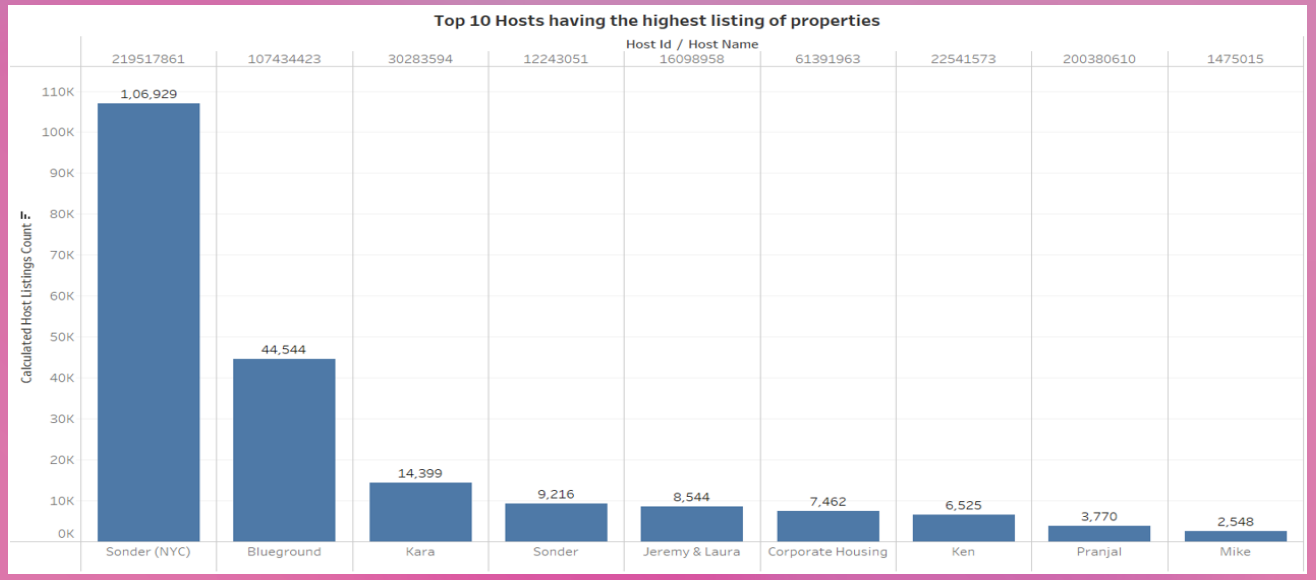
So, analysis has been done on a dataset consisting of various Airbnb listings in New York.

➢Objectives :

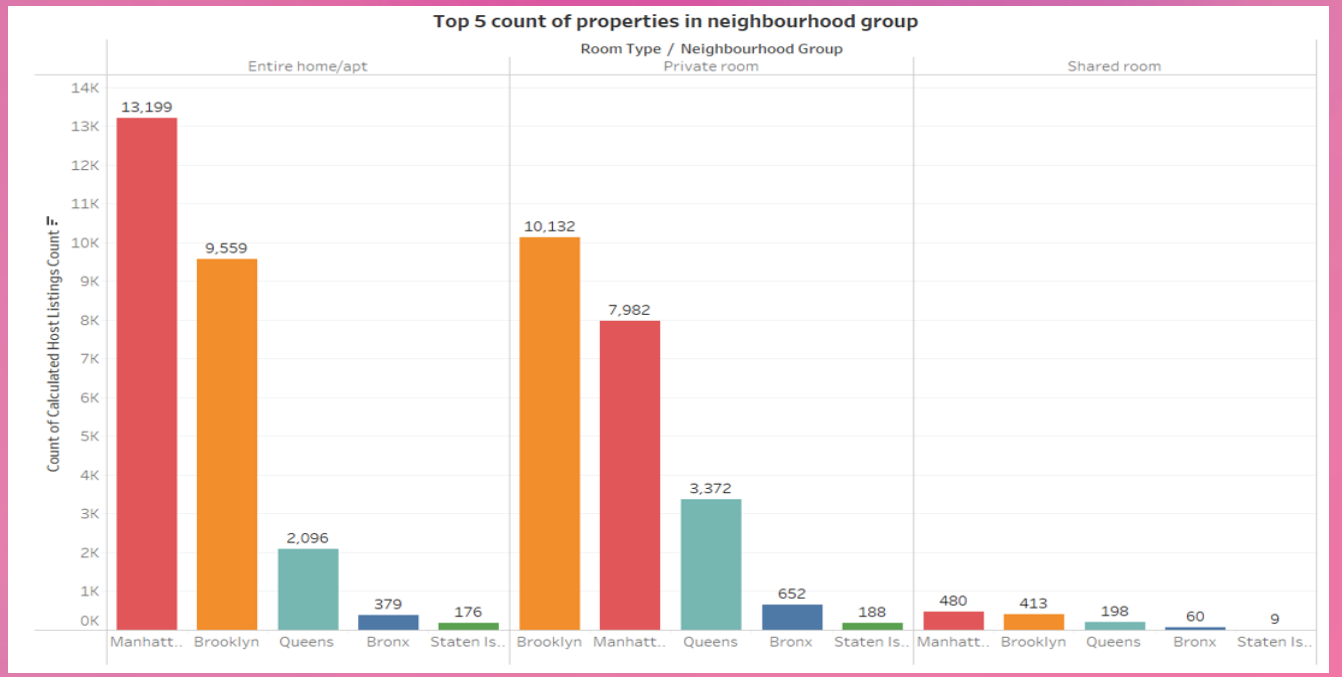
- **Main objective is to improvise the Airbnb business to generate more revenue.**
- **To acquire the properties according to the customer preferences.**
- **Raising the number of bookings for the unpopular properties with added incentives.**

➢Key Findings :

- Most popular host or top 10 hosts having highest listing of properties



- The most popular neighbourhood groups in the room listings are Manhattan and Brooklyn, and the majority of people prefer the entire home or private room.



➢Key Findings :

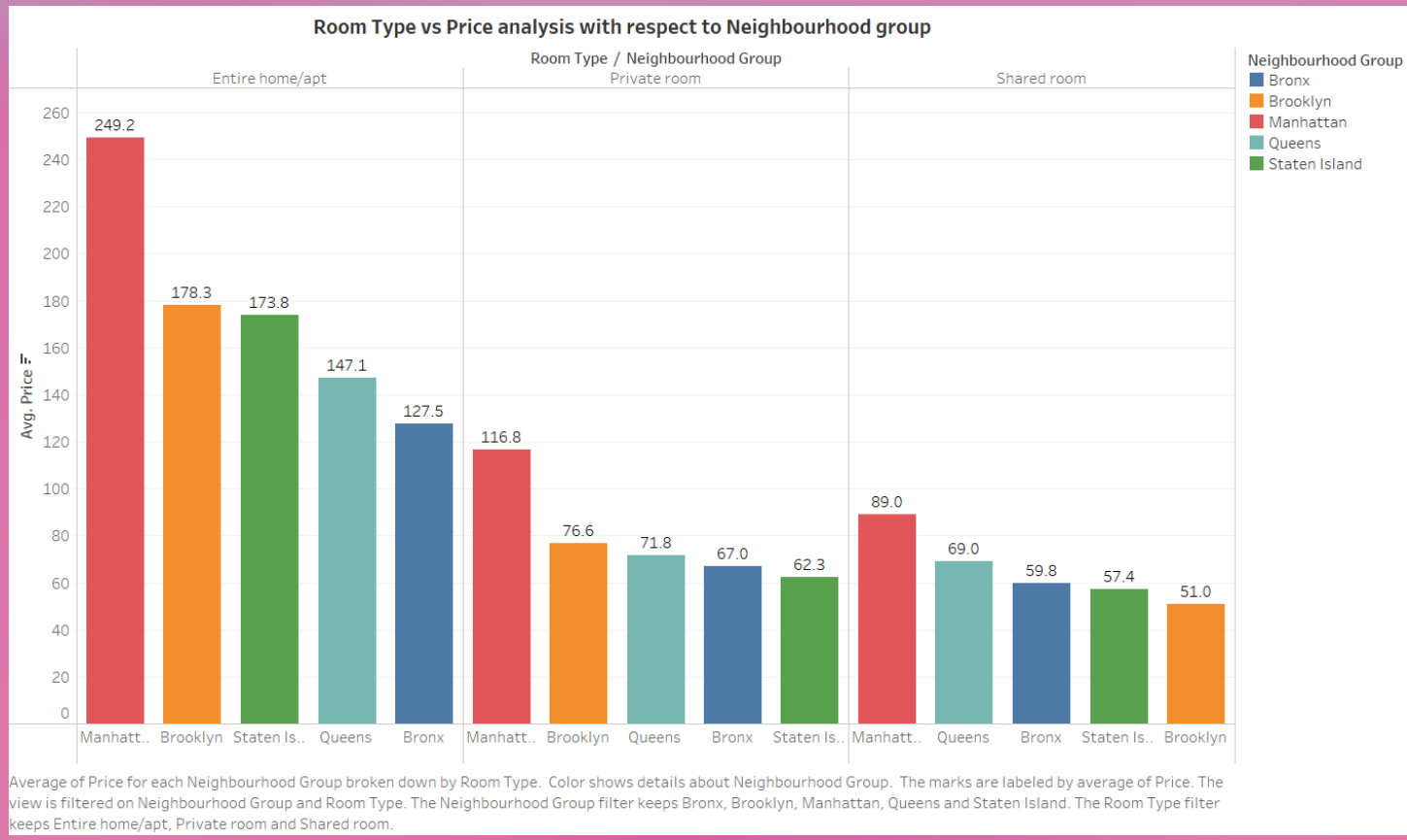
- These are the top 10 locations based on price, even if Williamsburg's accommodation rates are high due to the neighborhood's appealing atmosphere.

- Customer preference based on minimum nights of stay :
  - Customer mostly opt for below & up to a month of stay.
  - Customers rarely opt for a minimum stay of more than a month.

➢ Key Findings :

- **Room Type vs Price analysis with respect to neighbourhood group :**

- **Manhattan, Brooklyn are the top two neighbourhood groups with the highest average price in Entire home /apt & Private room category.**
- **In the Shared room type category, Manhattan & Queens have the highest average price**

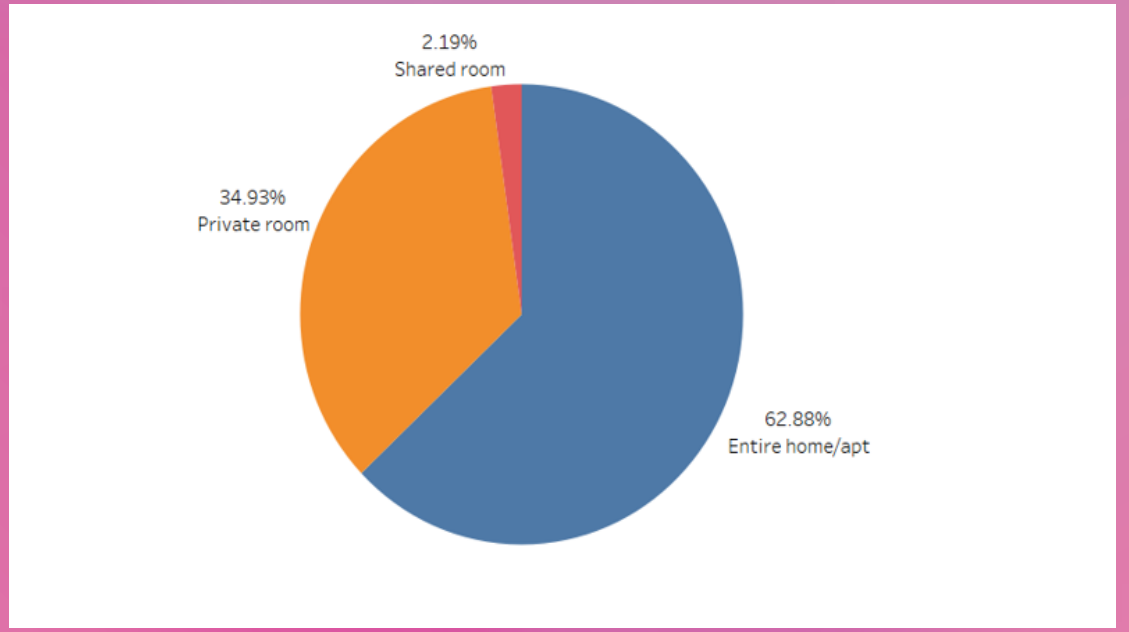


➢Key Findings :

- Customer preferences based on location & type of the room

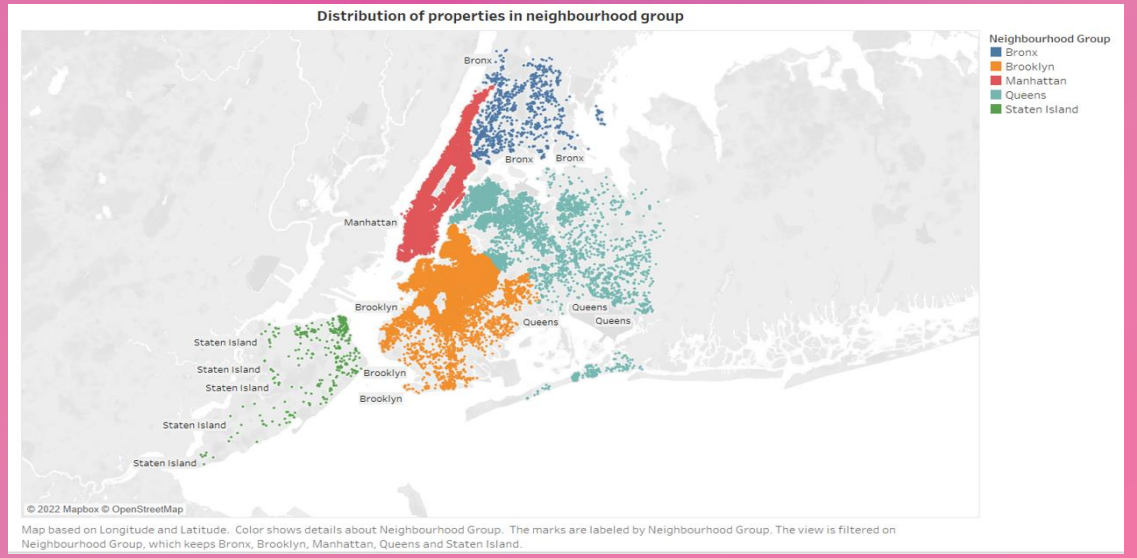
  Overall room type percentages :

  - Entire home/apt = 62.88%
  - Private room = 34.93%
  - Shared room = 2.19%



- Property distribution in the neighbourhood group :

- Staten Island is sparsely populated with properties.
- Manhattan, Brooklyn, Queens, and the Bronx are all more densely populated than Staten Island.



➢ Recommendations :

1. Acquire private rooms and entire apartments, especially in Staten Islands since it is popular.

➢ Recommendations to get unpopular properties more traction :

1. We can provide complimentary airport pick-up and drop-off services.
2. Complimentary breakfast can be provided.
3. Offer membership discounts on Airbnb and encourage more people to apply for those unpopular properties.
4. We can offer discounts by collaborating with local recreational facilities such as indoor games, swimming pools, snooker, libraries, and Carom, among others.

➢Appendix :

Appendix 1 - Data sources

- **Data set provided for the assignment**

  AB_NYC_2019.csv : The provided dataset contains details regarding all the properties in New York which are Airbnb. Dataset also contains the customer reviews.

- **Size of the provided dataset**

  AB_NYC_2019.csv has 48895 rows & 16 columns.

➢Appendix :

Appendix 2 - Data cleaning

AB_NYC_2019.csv Dataset :

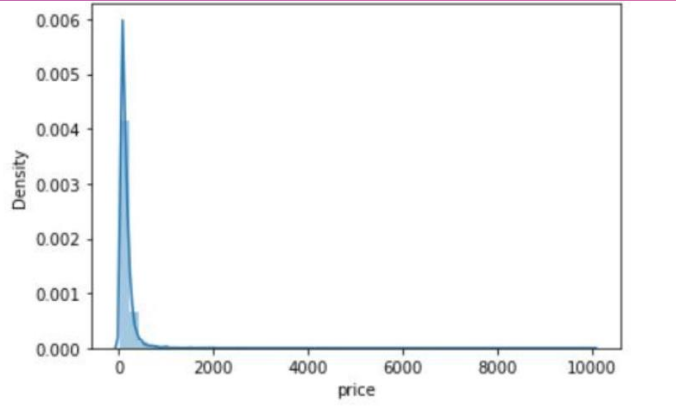
a. Checked the Null values in the dataset. Found some columns with the null values i.e. names , host_name, last_review , and review_per_month.

b. Checked the outliers in the dataset.

Appendix 3 - Data methodology

- **Exploratory Data Analysis:**

- Checked for the Null values/Missing values in the dataset. Missing values are present in the “name, host_name, last_reviews, reviews_per_month columns. We dropped the column last_review & impute missing values in reviews_per_month as 0.
- Checked the outliers in the dataset.
- There are 11 properties which have zero price. Values might not have been entered or might be an error. Binning of price ranges to be done to check the value to be imputed for the properties which have zero value. We will also have the null values present in the host_name and name columns as they are not required for now.
- The price range is heavily skewed with values mostly between 0 to 2000. Impute median values on the 11 properties which have zero price.



- Cleaned dataset then converted to csv for further tableau analysis.

  “data.to_csv(‘air_bnb.csv’)”

- **Data Analysis :**
  - Try to analyse the data using different columns based on their price, availability 365, minimum nights, and customer reviews.

- **Conclusion after analyzing the dataset :**
  - We discovered that people like visiting New York's centre, from which they can enjoy the city's beauty.
  - The number of shared room listings is limited, but their average price is low and availability is high.
  - Number of reviews and reviews per month are more at less price than the higher price as there is less chance of people going for a high price room.
  - Manhattan and Brooklyn Neighborhood groups are quite expensive.
  - People are drawn to the host Blueground and stay for longer periods of time.
  - Minimum number of nights to stay reduces with increase in price.
  - Pay attention to popular areas like Manhattan and Brooklyn where people are interested.

Thank you