**Analyzing the *Airbnb* Market in New York City (Using R)**

**Introduction:**

In this project, we will gain insight into the world of Airbnb in New York City, conducting a detailed descriptive and exploratory analysis of the available data. Our objective is to acquire a better knowledge of how each variable in the dataset operates on its own and in combination with others. By doing so, we intend to produce assumptions that will be beneficial for future decision-making about Airbnb in the New York city.

To do this, we will employ a variety of strategies and resources, such as histograms, heatmaps, box plots, scatterplots, and other data visualization methods in R Studio. We will be able to find patterns and correlations between variables by utilizing these tools to display and analyze data in a useful and significant manner.

Thus, the goal of this study is to give a thorough and analytical overview of Airbnb in New York City. We intend to identify major trends and insights using our study that will be valuable for people working in the Airbnb field and customers interested in choosing properties using Airbnb or anybody interested in understanding the mechanics of the sharing economy in one of the world's top cities.

**Target Variable:**

The target variable chosen for study in the New York City Airbnb dataset is price. In the short-term rental market, price is a critical aspect that impacts both hosts and guests' decisions. Hosts want to maximize their revenue by optimizing their prices, while visitors want the most value for their money. Understanding the elements that drive a listing's pricing can give insights into the market's supply and demand dynamics, allowing hosts and guests to make more educated judgments.

The cost of an Airbnb listing in New York City is determined by a variety of criteria, including the location, room type, availability, and amount of facilities given. Listings in popular neighborhoods, as well as those near attractions and public transit, tend to fetch higher rates. The type of room is also important, because complete flats or homes are usually more expensive than private or shared rooms. Another critical consideration is availability, as costs might vary depending on the time of year and demand for short-term rentals. Finally, the quantity of facilities given might have an impact on the pricing, with listings that provide more amenities likely to be more expensive.

By evaluating the elements that determine the price of a listing, hosts may maximize their revenue by optimizing their listing pricing. To attract more visitors, hosts may, for example, modify their pricing based on the time of year or give discounts for longer stays. Guests may also utilize this information to discover the most value for their money by searching for listings that provide a desirable location and facilities at an affordable price.

**Problem Statements:**

Problem – 1: **How does the price of an Airbnb listing in New York City affect the number of reviews?**

Problem – 2: **What is New York City's most popular Airbnb neighborhoods?**

**Problem Description:**

**Problem – 1:** The price of an Airbnb property in New York City and the number of reviews have a complicated relationship. While higher-priced listings can attract visitors with higher expectations, leading to more negative reviews, they may also provide facilities and features that contribute to more good evaluations. Lower-priced listings, on the other hand, may attract visitors who are more budget-conscious and hence more inclined to give positive feedback if their expectations are fulfilled. Overall, the study indicates a negative relationship between price and the number of reviews, implying that lower-priced listings receive more reviews than higher-priced listings.

**Problem – 2:** We can evaluate the number of listings and reviews each neighborhood to find the most popular neighborhoods for Airbnb rentals in New York City. The most popular areas are likely to have the most Airbnb listings and the most reviews. However, bear in mind that other considerations such as location, facilities, and cost may all have an impact on popularity. The dataset's "neighbourhood" and "neighbourhood\_group" columns can be used to categorize entries by neighborhood and county. Then, for each neighborhood, we can compute the total number of listings and total number of reviews and arrange them in decreasing order to determine the most popular neighborhoods.

**Exploratory Data Analysis (EDA):**

**Data Analysis:**

Exploratory data analysis is a critical step in understanding the data's underlying patterns and trends. We started with data pretreatment in this project, where we removed unnecessary or missing data like outliers in dataset and standardized data formats to address any data quality concerns that may impair the accuracy of the research. The data visualization phase included creating visual representations of the data, such as scatter plots, histograms, and heatmaps, to detect patterns, trends, and correlations between various variables. The investigation focused on three primary issues: first, determining the elements that determine a listing's pricing. To address this, we employed correlation analysis to determine the magnitude and trend of the relationship between various variables such as the location, kind, and amenities of the listing, as well as its price. Second, we attempted to forecast a listing's availability.

.

**Modeling:**

**Descriptive statistics:**

As part of the Airbnb project, descriptive statistics can be utilized to give insights into key dataset aspects such as the average rental price, the most popular areas, and the frequency of various types of rental spaces. The mean, median, and mode, for example, may be used to explain the central tendency of the pricing of Airbnb rentals in New York City. The mean would represent the average rental price, while the median would represent the middle price point and the mode would represent the most common pricing point. These indicators can aid in understanding market pricing patterns. Similarly, descriptive statistics may be used to determine the most popular neighborhoods for Airbnb rentals in New York City.

After removing the outliers in the data set, we have the updated data now which is shown below in the picture.

Text

Description automatically generated with medium confidence

So, the summary of this updated data shows the mean, median, 1st quadrant, 3rd quadrant, minimum and maximum values of each column involved in this analysis, like price, number of reviews, reviews per month…etc. so with the help of this now we will be plotting graphs with the help of this values for the columns to identify the important factors which play a crucial role to determine the price.

Now with the help of the same updated data we have got the frequency and percentage of room types and the average price and price percentage based on room type.

Text

Description automatically generated with low confidence

Text

Description automatically generated with medium confidence

**Average Price per Room Type**

Text

Description automatically generated with medium confidence

**Top 10 expensive neighborhoods**

Text

Description automatically generated with medium confidence

**Price range with out constrains based on latitude and longitude columns**

**Conclusion:**

In the summary, analyzing the New York City Airbnb dataset, we can infer that various factors impact the pricing, availability, and popularity of listings. The price is significantly determined by the location and room type, while the availability of a listing is influenced by the number of reviews and the time of year. The amount of reviews a listing receives is impacted by various factors, including the price, location, and facilities given.

We observed that lower-priced listings have a larger number of ratings, indicating that New York City tourists are more value-conscious and prepared to sacrifice elegance for pricing. In terms of geography, Manhattan is the costliest neighborhood for Airbnb rentals, while Brooklyn and Queens are less expensive.

Our research also revealed significant characteristics linked with higher-priced postings, such as a private room or complete apartment, air conditioning, and a necessary minimum number of nights. Furthermore, we discovered that hosts who provide a large variety of facilities and respond quickly to visitor questions tend to obtain more positive evaluations. Based on our findings, we make advice to hosts on how to improve their listings and enhance their income. Hosts could think about giving individual rooms or full apartments, providing important facilities like air conditioning and Wi-Fi, and requiring a minimum number of nights. To achieve favorable evaluations, they should also reply quickly to guest inquiries and give a high level of service and facilities.