**Airbnb Report.**

**What is Airbnb?**

Airbnb is an online marketplace connecting people who want to rent out their homes with people looking for accommodations in specific locales. Airbnb offers people an easy, relatively stress-free way to earn income from their property.

*(https://www.investopedia.com/articles/personal-finance/032814/pros-and-cons-using-airbnb.asp)*

**Airbnb Statistics:**

There are currently over 4+ million Airbnb hosts worldwide and 6+ million listings on the platform. There are over 100,000 cities worldwide that have Airbnb listings in them. Airbnb has more than 150 million worldwide users that have booked over 1 billion stays*.*

*(https://www.matthewwoodward.co.uk/work/airbnb-statistics/#:~:text=Airbnb%20Key%20Statistics%202022,-Airbnb%20is%20a&text=There%20are%20currently%20over%204,booked%20over%201%20billion%20stays)*

**Report Proposal:**

I aimed to work on the dataset carefully, choose research areas, make graphs based on those areas and interpret them.

**Data Cleaning:**

The data was messy and unclean. Therefore I used Tableau 2022.3 to clean them. Clicking on the “Use Data Interpreter” icon on the left pane of the screen was the first thing to do. Then null values appeared on most columns. I could get rid of them by filtering ‘only non-null values one by one for each column. Data types of Neighborhood and Neighborhood Group were a string. I detected their geographic role and changed the data type. Some columns were unnecessary, such as Table name, Sheet. I could get rid of them by hiding the columns.

**QUESTIONS**:

1. **How does the average price change with respect to the construction year(2003 and 2022) in each neighborhood group?**

**Chart, line chart

Description automatically generated**

From the graph, it is understandable that each neighborhood group has a different distribution from 2003 to 2022. The sharpest increase in the average price of houses is observed in Staten Island, from 628.4 to 692.5 dollars. The positive price change in the Bronx is 39$, then Manhattan(19$), and finally Brooklyn(2.7$). Obviously, between 39$ and 19$, 2.7$ is such a small change number, we can conclude that Brooklyn has a weak positive change from 2003 to 2022.

Surprisingly, the Queens neighborhood group has changed negatively over the years(-5.7$). The reason can be many things, such as a low rating(due to a low rating host deciding to lower the price) or people not wanting to take risks when they rent a house in Queens (renting a 2022 constructed house is riskier because those new ones do not have many reviews so that people can decide easily)

1. **What do you think: Did more prices mean more reviews in 2020? Answer this question for each neighborhood group.**

Map

Description automatically generated

We should consider this to interpret the data correctly: the darker the color, the higher the price. (Values are filtered for 2020)

It is obviously seen from the shade of the colors that the most expensive houses based on the average price were in the Bronx (709$), followed by Staten Island (688$) in 2020. However, looking at the number of reviews, we see that those neighborhood groups were the lowest. Queens(cheapest) was the leader with 41 reviews, Manhattan 26 was second, and Brooklyn 32 was third. This means price and review had a negative relationship, and there are several reasons for that.

1. The wealthiest people who rented Airbnb houses in 2020 were careless about giving a review or helping others. The first reason is mostly it did not matter for them to look at the review and compare houses based on their features and prices; in the end, they chose the most expensive ones. There is an easy and straight logic: the most expensive is the best. The second reason is that those people may not have enough time to give a review since they are busy all the time. That is why the number of reviews was low in expensive areas.

2. During the pandemic (2020), most people were not in a financially good position. That is why they tended to rent cheaper houses. That is why in more affordable areas, reviews are high.

1. **In which type of rooms did people prefer spending more time in 2020-2022? How did their preference change during this period? Compare them and explain the reason briefly.**

**Chart, line chart

Description automatically generated**

In 2020, the average minimum nights were higher in the entire home/apt(nearest to 8), followed by shared rooms (6-7) and private rooms (4). Comparing the changes from 2020 to 2022, the entire home shows an almost increasing pattern; shared room distributes strictly decreasing trend, but private room increases in 2020-21, and decrease in 2021-2022(not very sharply)

From this interpretation, we can conclude that people who rent entire homes/apt-type houses tend to travel for an extended period, and their tendency increases over time. The reason: those who rent entire homes/apt mostly travel with their family/friends; spending time with them is fun, and that is why those people do not get bored staying in the same place for 8-9 days.

As you know, 2020-2022 was a pandemic period. That is why people were more cautious then and did not want to live with foreigners (maybe they had a coronavirus, which was a huge risk at that time). It can be a logical reason for decreasing pattern of shared rooms from 2020 to 2022.

About the third room type-private room, I can say that living in those types of houses is more expensive, that is why people could not afford to stay in those places for long nights. We cannot conclude its time pattern.

1. **Which is more popular in Manhattan: traveling lonely or with other people?**

**A picture containing table

Description automatically generated**

It can be observed that buyers mostly choose entire home apartments(11-12) followed by shared (3-4) and private rooms(3) when they are renting Airbnb houses in Manhattan. Knowing that people choose

1)entire home/apt when they are traveling with family/friends

2)shared room when they want to meet new people or make new friendships

3)private room when they want to have a little privacy and be alone

We can make a conclusion: traveling lonely is not popular in Manhattan.

1. **Compare the unverified/verified hosts in Bronx (2008) based on their reviews, price, and numbers.**

Chart, pie chart, bubble chart

Description automatically generatedTable

Description automatically generated

**Chart, bar chart

Description automatically generated**

In order to answer those questions, we should analyze three graphs. The first graph shows the number of verified/unverified hosts in the Bronx in 2008. Unverified:19, verified:31. The second table tells us the average price of houses possessed by each type of host. Unconfirmed:500$, confirmed:402$. Finally, the third bar chart gives us information about the number of reviews based on the type of hosts in the Bronx(2008). Unconfirmed:28.3, verified:29.2.

One of the surprising facts is that in Bronx(2008), only 38% of the hosts were unconfirmed. Confirmed hosts were more popular at that time. Although unverified hosts have higher prices than verified hosts, they got fewer reviews. The reason can be that buyers are scared to rent a house from unconfirmed hosts because it is not safe at all in some cases. That is why only a few people reserved Airbnb houses from those hosts in the Bronx in 2022. the number of verified hosts was higher than unverified hosts; unconfirmed ones got more reviews. Maybe they rented those houses but wanted to avoid giving a review on a public platform.

1. **What are the bottom ten neighborhoods that had the lowest service fee in 2016? How did the service fee affect the review rate number?**

Chart, treemap chart

Description automatically generated

Description: Darker the color, the higher the average service fee.

Bigger the size, the higher the average review rate number

The first numbers represent average review rate numbers, the second the name of the neighborhood, third the average service fee.

The neighborhoods are the top 10 ones with the lowest service fee in 2016.

It is observed that Emerson Hill(94.50$) had the highest service fee in 2016, it is followed by Douglaston(91.90$) and Manhattan Beach(86$). However, those are the neighborhoods that had quite low review rate numbers.( Emerson Hill:1, Douglaston:2.2, Manhattan Beach:1.7).

When we look at the graph, we can immediately see that Bay Terrace and Staten Island both had a 5/5 review rate number, and their service fee was relatively low compared to other ones in the graph(12$)

The reason:

1. There is a high possibility that in 2016 when people had to pay relatively high service fees in advance, their expectations about houses increased and higher. That is why reality did not meet their expectations, so review rate numbers were not high.
2. On the contrary, paying low service fees made people happier; they felt like paying with a discount or smth. Hence, review rate numbers were high enough.

**Descriptive Statistics.**

**Price:**

**Table

Description automatically generated**

**Interpretation:**

The prices of half of the Airbnb houses were below 667$, and the costs of half of them exceeded 667$ in Manhattan in 2007.

The minimum price of the houses was 69$ while the maximum price was 1186$ in Manhattan in 2007.

25% of the prices were 517$ or less, and 75% were 887$ or less in Manhattan in 2007.

**Service Fee:**

**Table

Description automatically generated with medium confidence**

**These values are valid for private room types.**

The mean of the service fees is 125$

The service fees for half of the private rooms are below 124$, and half of them are above 124$

Variation of the service fees is quite high:4394$.

St.Dev is also high, meaning service fee values are more spread out rather than being clustered around the mean(125$).

The Sum of the service fees for private rooms from 2003 to 2022 is approximately 2,3M.

**Neighbourhood Groups:**

**Chart, bar chart

Description automatically generated**

This bar graph shows the frequency distribution of neighborhood groups.

It is seen that the lowest one is Staten Island (361) and the highest one is Manhattan(approximately 17k).