

Exploratory Analysis on NYC locations and Airbnb availability
(Fordham STEP program)

Arpita Choudhury

Amy Sarker, Ishrat Iqbal

Abstract:

OBJECTIVE:

Airbnb has become the most popular destination and choice for many people traveling and experiencing the world. However, the progression and information regarding the success of Airbnb amongst hosts and guests are still to be investigated and analyzed. This study aims to analyze a dataset based on NYC locations and their Airbnb availability and determine the factors for their traffic and growth to help hosts and guests determine valuable information and disparities.

METHODS:

We created charts on Tableau to better understand different variables in data and investigate specific hypotheses we formulated. For instance, we compared the price of Airbnbs within different boroughs in NYC and their room type. Additionally, we organized and compared specific neighborhoods according to availability, price, number of reviews, hosts, recent reviews, average listings count, and minimum nights.

Research Overview

Introduction

Since its launch in 2008, Airbnb (Air Bed and Breakfast) has skyrocketed in popularity as an alternative to hotels. Guests and hosts have used Airbnb to expand on traveling possibilities and present a more personalized, unique, and worthy way of experiencing the world. Founders Joe Gebbia and Brian Chesky started Airbnb to make extra cash, listing their apartment in San Francisco as the first Airbnb. Airbnb has shown massive success by providing many benefits to guests and hosts. In this project, we will focus on finding out more about hosts, geographical availability, and other necessary metrics to analyze the factors that affect Airbnb's traffic in NYC.

Problem Statement

Is there any difference in Airbnb traffic in NYC areas, and what could be the reason for it? How do Airbnbs in different NYC locations differ from each other? Which locations have the most traffic? Does gentrification play any role in the explanation of these differences in availability and location?

Objective

This research analyzes the various information on Airbnb hosts, geographical availability, and other necessary metrics to conclude the relationship between NYC locations and Airbnb traffic. This project will help better understand what Airbnb locations have shown the most progress

and success in supporting guests and hosts in their experience and gaining profits. Additionally, we hope to help Airbnb that is not progressing, and there are reasons to overall improve the experience of Airbnb.

Methodology

The steps we can take to solve the problem would be to translate the dataset into specific models in Tableau that would address certain patterns and correlations between NYC locations and Airbnb's traffic. We can then research more about specific locations and their geographics to better understand the context and reasoning behind why so many guests and hosts are choosing certain Airbnbs than others.

Hypothesis

1. Airbnbs in general are pricey in NYC
2. Location significantly affects the Airbnbs price
3. Customers prefer apartments in prestigious locations
4. Gentrification can affect certain neighborhoods in terms of availability and price
5. Maximum number of minimum nights spend is largely affected by location and room type
6. The more the reviews, the higher the price of that location
7. Majority of the Airbnb's listings are in Manhattan
8. Most of the hosts are available throughout the year

Independent and Dependent Variables:

Independent variables: Neighborhood, latitude, longitude, host_name, host_id, id, name, neighborhood group, last review, number of reviews, reviews per month, neighborhood, calculated_host_listings_count

Dependent variables: availability_365, minimum nights, price, room_type

Data Description

Time interval: 2011-2019

Data Type -Text, Integers, Numeric

Data Scales - Nominal, Ratio, Interval

Data size = 9 years *16 columns* 38,843 rows
= 5,593,392

Data source: <https://www.kaggle.com/datasets/dgomonov/new-york-city-airbnb-open-data>

Comment: A total of 10052 null values were noted, and after removing them, the number of rows totalled to be 38,843.

Result (Tableau Analysis)

Figure 1: Airbnbs in general are pricey in NYC

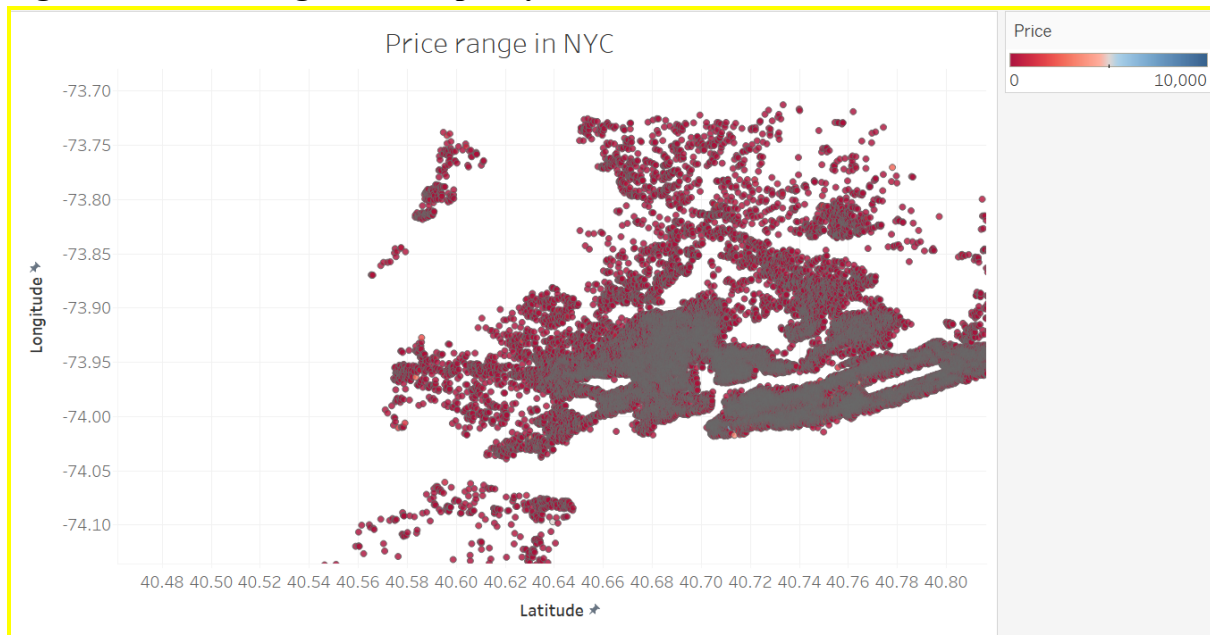


Figure 1 showcases the price range throughout New York City. From this graph, most of the points are in a reddish color which, from the price legend, illustrates the idea that most of the listings are below \$800.

Figure 2: Location significantly affects the Airbnbs price

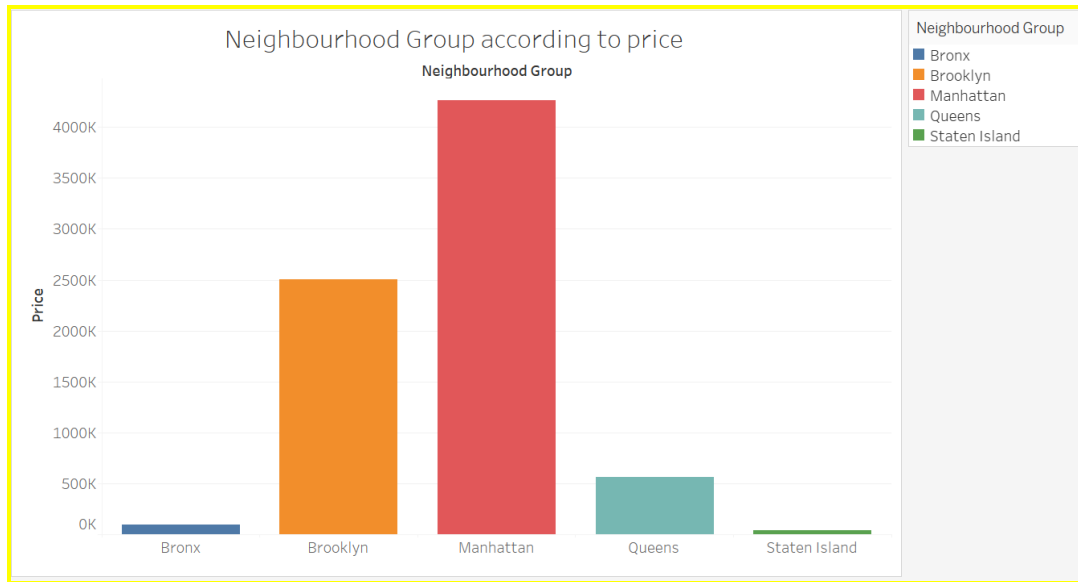


Figure 2 conveys how certain neighborhood groups like Manhattan have higher Airbnb prices than those like Staten Island and the Bronx. For example, Manhattan has its Airbnb listed for more than 4000K, whereas Queens has Airbnb listed for only 500K. We can conclude that the more tourists attractions, better sights, and activities for people to do in an area, the higher the Airbnb price will be because hosts know that Airbnbs are popular amongst people within these boroughs and will continue to increase as gentrification and construction continues to expand in these areas and provide better experiences and opportunities for people to enjoy.

Figure 3: Customers prefer apartments in prestigious locations

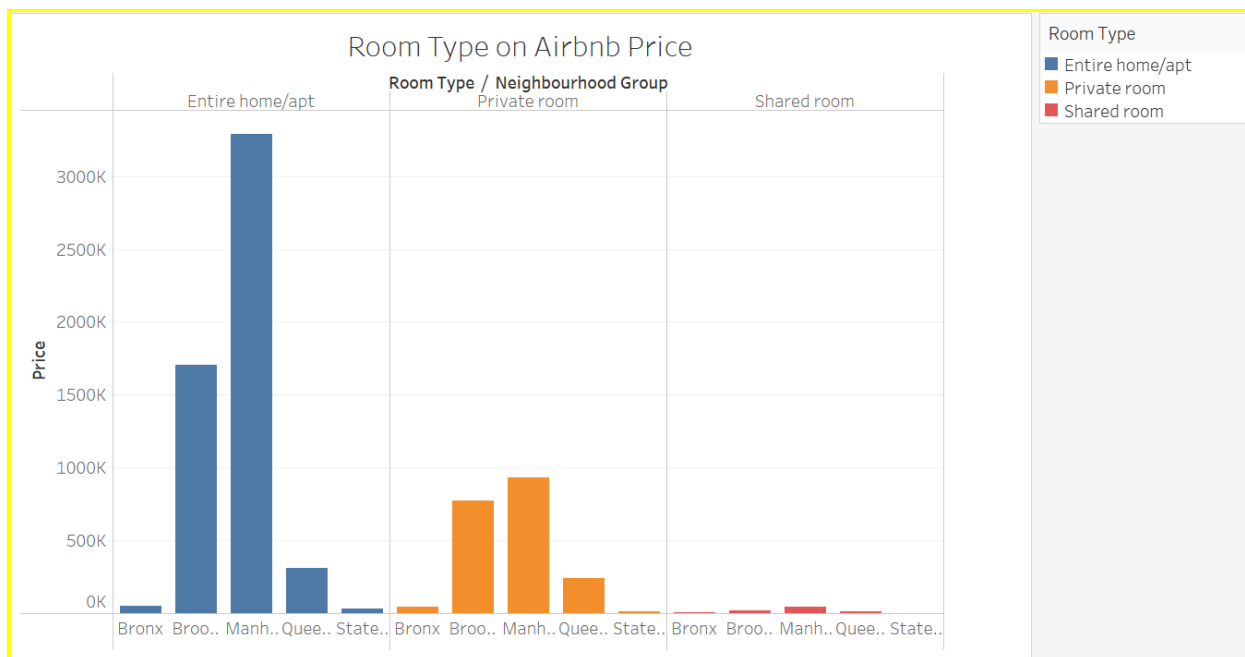


Figure 3 showcases the data on the average price of Airbnb within specific boroughs. The average price of Airbnb is in direct comparison with the Airbnb room type. According to this bar graph, we can see that Entire homes/Apartments are the most pricey, particularly in Manhattan, costing more than 3000K. Shared rooms are the least expensive in all boroughs. We can conclude that hosts are more likely to choose Entire home/apartment Airbnb in areas such as Brooklyn, Manhattan, and Queens because the cost is higher, and that may be due to the better views and sights tourists are looking for. Guests with a cheaper budget should aim for private rooms in areas such as the Bronx or Staten Island because although the sights are not as great, the budget and place to stay are not as bad as a shared room may have been.

Figure 4: Gentrification can affect certain neighborhoods in terms of availability and price

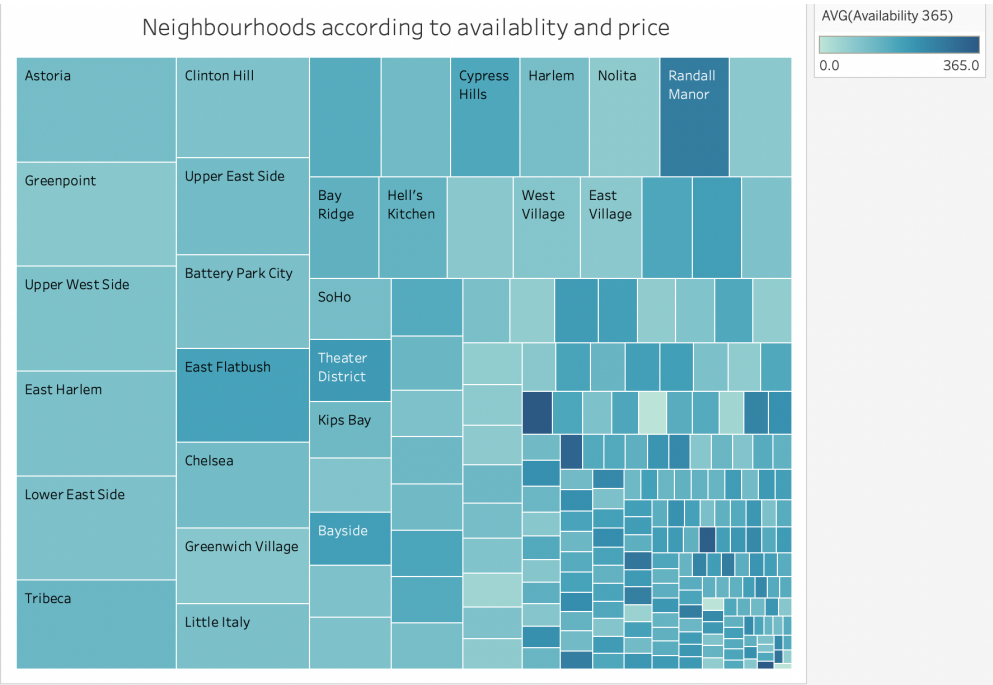


Figure 5 showcases different locations in NYC, and the color of their box represents the average availability, and their box size represents how pricey their Airbnbs are. For instance, Astoria has an average Airbnb availability of 109 and a max price of \$10,000. On the other hand, a place on the tree map such as Midtown has an average availability of 157 and a maximum price of 5,100. Based on this tree map, we can conclude that poorer areas such as Astoria and Greenpoint are becoming tourist attractions, raising the cost of Airbnb and increasing availability. This leads us to a more resounding conclusion and the issue at stake, gentrification.

Figure 5: Maximum number of minimum nights spend is largely affected by location and room type

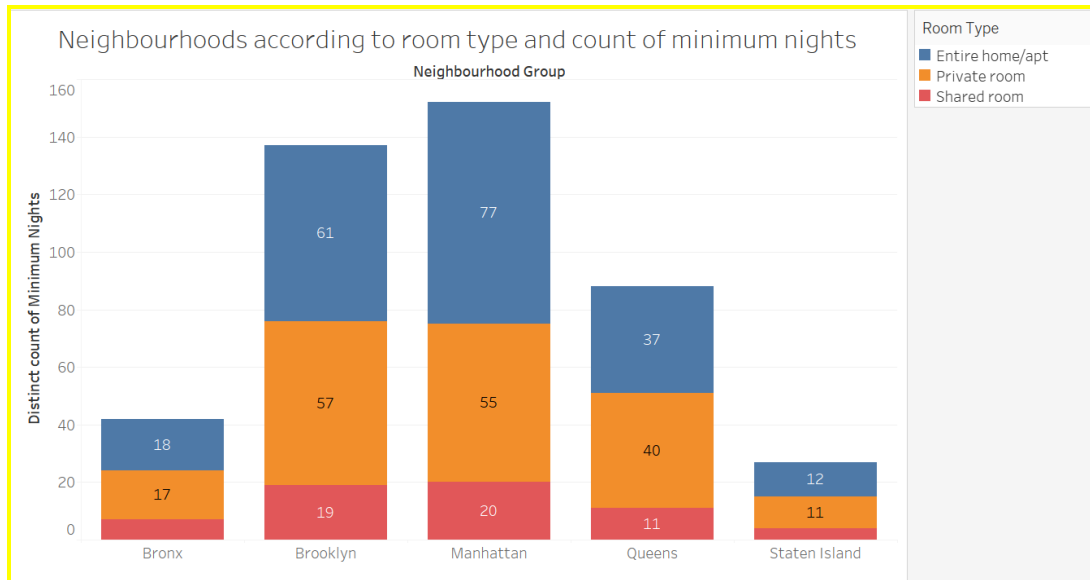


Figure 6 shows the distribution of room types along with their correlated minimum nights expenses. The data in the bar graph reveals that Manhattan is the most popular for a combination of all room types and with their attached minimum nights compared to areas like the Bronx, Brooklyn, and Staten Island. This suggests that people or tourists prefer to have entire homes and apartments in Manhattan as they usually stay for a minimum of 77 nights. That is a lot of nights, but Mahantten has a lot for tourists to see. We can infer that entire homes and apartments, which are the most luxurious and expensive for tourists, are most common in areas like Manhattan. But, shared and private rooms are more affordable and better options for tourists visiting areas like the Bronx, which does not offer much to offer in terms of activities and enjoyable opportunities that tourists travel for and to.

Figure 6: The more the reviews, the higher the price of that location

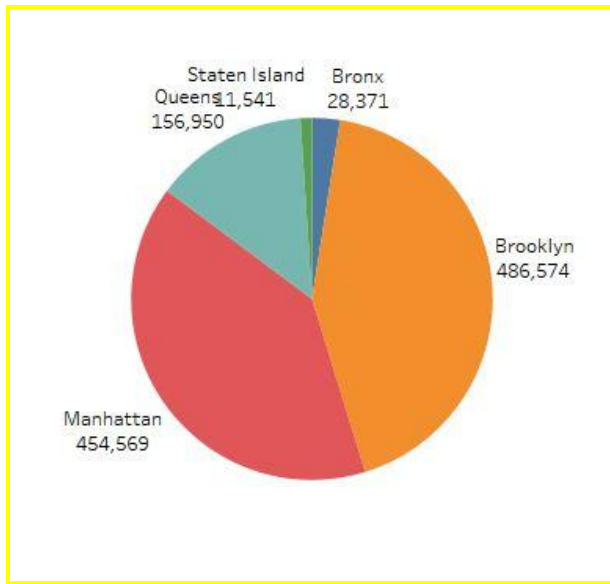


Figure 7 demonstrates the number of reviews from each borough. According to this graph, Brooklyn and Manhattan have more reviews. This can be because customers will likely rent Airbnbs in Manhattan and Brooklyn rather than in boroughs like the Bronx and Staten Island.

Figure 7: Majority of the Airbnb's host listings are in Manhattan

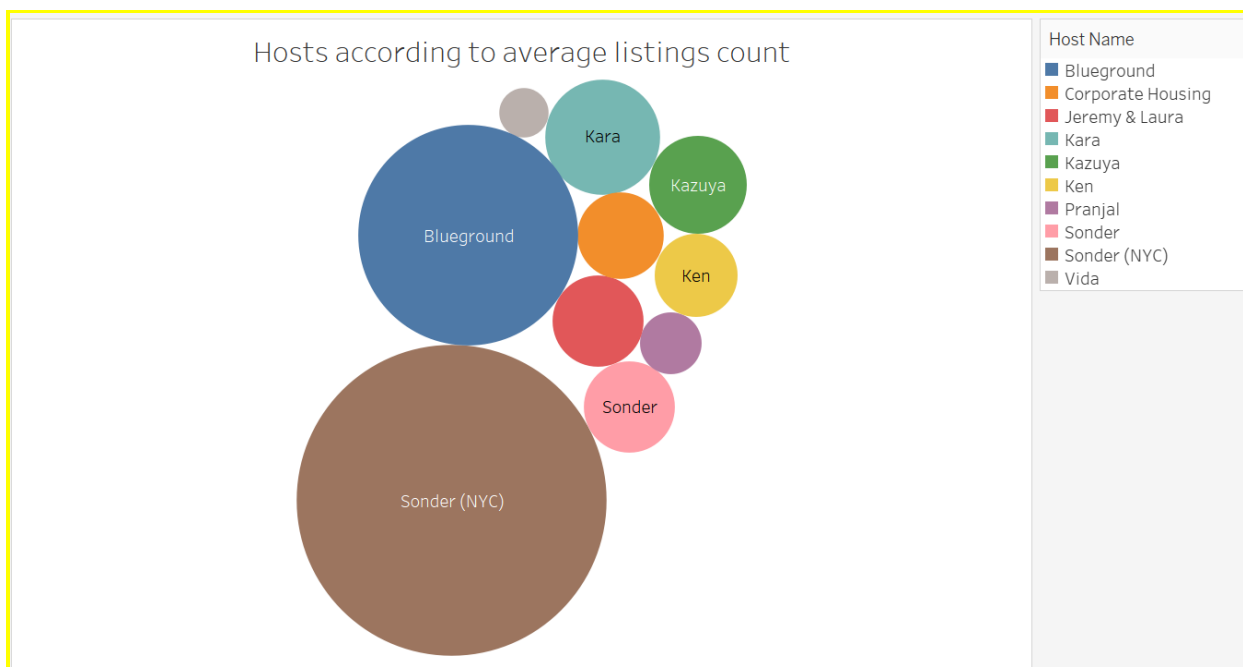


Figure 9 showcases how Host Name, Sonder, has the highest average Airbnb listening count compared to all other hostnames such as Ken, Kazuya, and Kara, with much smaller-sized circles and host listening counts. This conveys that NYC, most likely Manhattan, contains the most host listings as more people want Airbnb to live in this area. Hosts and guests can profit most from choosing Airbnb in NYC, and Sonder must have good reviews and success.

Figure 8: Most of the hosts are available throughout the year

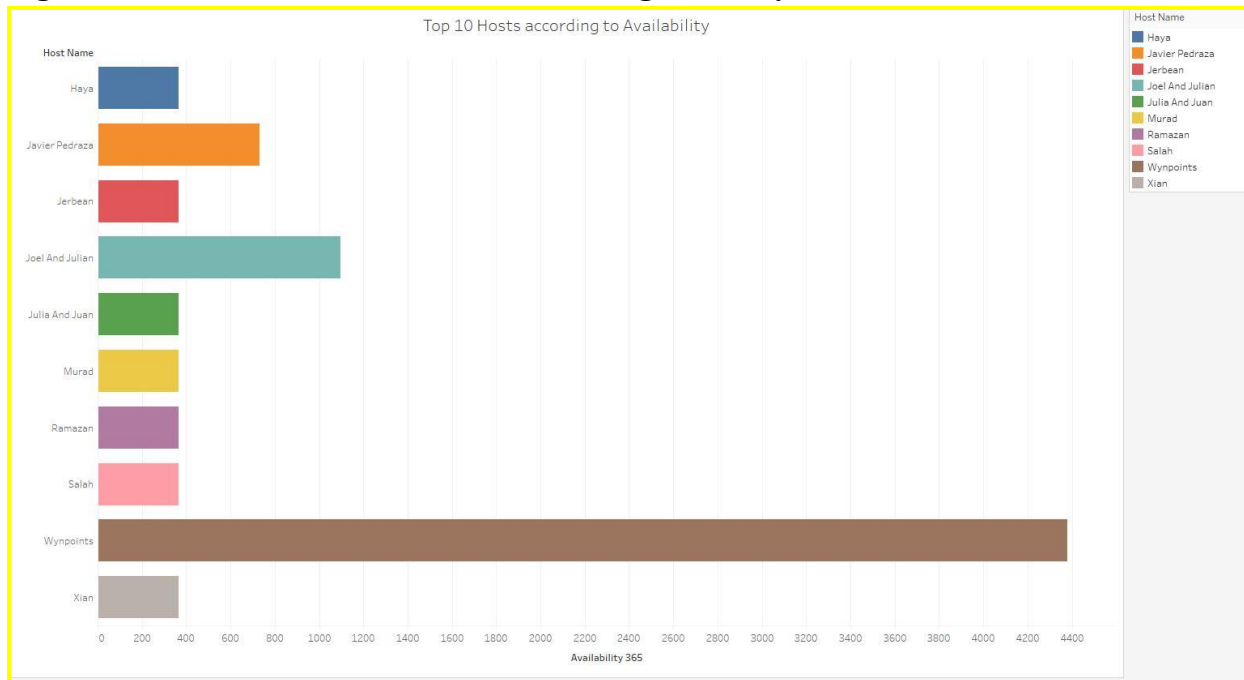


Figure 10 demonstrates the availability rate of each host from this data set. From this graph, these are the top ten hosts according to availability. Furthermore, Wynpoints have the highest number of availability which is 4,380 implying that they might have several listings available all throughout the year. Apart from the three hosts with higher numbers, the average number among the rest is 365.

RESULTS:

Our Figure 2 bar graph showed that Entire homes and apartments were the most pricy in areas like Manhattan rather than a place like the Bronx. We also determined that areas like Astoria have shown lower availability, with an availability of 109 and a price of \$10,000, despite their costly Airbnbs. On the other hand, Chelsea, Manhattan, known for its prestige and beauty, has a higher and cheaper Airbnb availability. According to the dataset, Chelsea has an average availability of 114 and a price of only \$6,800. We found this odd, but after further research discovered that this is because of gentrification and the reconstruction and building of wealth in poorer but beautiful areas such as Astoria. Also, Manhattan and Brooklyn were prone to

receiving more reviews than any other borough, with up to 486,574 reviews compared to the 28,371 in the Bronx, which suggests that Manhattan Airbnb were likable and recommended to others. This data is shown through various graphs, such as a scatter plot, line graph, pie chart, bubbles chart, and bar graphs.

CONCLUSIONS:

Airbnb in Manhattan and Brooklyn are prone to receiving the most profits from guests and hosts. With gentrification and rebuilding these particular areas in NYC, Airbnb's availability and prices will likely increase with traffic and growth, as new sights and activities will become available. These will provide many gains to guests and hosts. However, the poorer communities in these areas should be aware of this progress of gentrification as it has both negative and positive effects.

Further Analysis

Our further analysis is meant to provide a deeper understanding of the long-term effects of the growth of Airbnb in places such as Manhattan and Brooklyn, which have shown to have the most Airbnb traffic and success. It has less availability, meaning more people are booking Airbnb, despite its cost. Our Tree Map particularly sparked a conclusion that gentrification explained the rise in Airbnb prices and success. Places like Astoria are now profiting the most from Airbnb than Chelsea. So, Airbnb's success is suitable for most hosts; however, what is gentrification? How can we educate the public on this issue? What are the positives and negatives of gentrification in the real world? Are there possible solutions or justices that we do as a society? Here are some articles and research that discuss the questions brought from the research further.

1. Gentrification: Why is it a Problem?

- Results: Due to gentrification, skyrocketing housing costs in the Bay Area have driven many people of color, the elderly, and people with disabilities from their homes. From 2010 to 2014, area households with annual incomes of \$100,000 grew by 17%, while households making less decreased by 3%. As a result, "affordable housing" has become virtually nonexistent in the San Francisco-Oakland area.

2. What came first, the gentrification or Airbnb? Identifying spatial patterns of neighborhood change using Airbnb data

- Results: In this study, we implement an integrated perspective towards Airbnb as one of the drivers of neighborhood change and gentrification in the DMA. This means that we see Airbnb as a technologically supported platform that can create a new form of rent gap or highlight the existing rent gap in the market.

Concurrently it offers tools to close this technology-driven rent gap by commodifying neighborhoods and selling them to tourists. It also facilitates the (over)consumption of demand for short-term rental accommodation in specific urban locations, which are (inter)nationally notable and culturally desirable.

3. Gentrification doesn't have to force minority residents out of their homes. Activists say there are 3 ways to protect communities.
 - Results: taxing vacant properties to free up housing stock, preventing gentrification through zoning and other housing regulations, and empowering residents to resist gentrification

Scope and Limitations

Scope:

1. General knowledge about Airbnb hosts and different NYC areas could be obtained from this analysis.
2. The price variation in different boroughs
3. Gaining knowledge about the hosts availability

Limitations:

1. Guests could contact the host outside of the platform to avoid paying commissions.
2. Verification is needed to ensure that most reviewers are not from family/friends or paid reviewers.
3. Host picture and name could intentionally or unintentionally influence booking decisions.

Conclusion and Future Research

The graphs demonstrate the connection between location and Airbnb traffic. To illustrate, in Manhattan, NYC, the Airbnbs are more expensive compared to other boroughs. Furthermore, apartments have a higher price than other rooms, such as shared and private rooms. It can be concluded that Airbnbs in Manhattan and Brooklyn are more extensive options for those renting Airbnbs in NYC. With the growth of Airbnb's comes the growth of a more deeper and serious issue: gentrification. We must work to bring solutions to the issue of gentrification and help those in need.

Further Studies

Factors involved in Airbnb traffic or that partake in Airbnb progress, such as hotels, can be studied further to analyze its impact. Human and location demographics can also be further

analyzed deeper to investigate correlations and possible conflicts or disparities for future solutions.

Policy or Managerial Implications

This research implies more Airbnb traffic in Manhattan than in other boroughs. However, other factors need to be considered. For example, which places have the most Airbnb traffic, and how do the demographics and the city's funding participate in the Airbnbs?

References

Deane, S. (2022, March 28). *2022 airbnb statistics: Usage, demographics, and revenue growth*. Stratos Jet Charters, Inc. Retrieved July 19, 2022, from <https://www.stratosjets.com/blog/airbnb-statistics>

Rabiei-Dastjerdi, H., McArdle, G., & Hynes, W. (2022, May 18). *Which came first, the gentrification or the airbnb? identifying spatial patterns of neighborhood change using Airbnb Data*. Habitat International. Retrieved July 19, 2022, from <https://www.sciencedirect.com/science/article/pii/S0197397522000790>

Longley, R. (2021, April 23). *Gentrification: Why is it a problem?* ThoughtCo. Retrieved July 19, 2022, from <https://www.thoughtco.com/gentrification-why-is-it-a-problem-5112456>

Dgomonov. (2019, August 12). *New York City airbnb open data*. Kaggle. Retrieved July 19, 2022, from <https://www.kaggle.com/datasets/dgomonov/new-york-city-airbnb-open-data>

Williams, M. (2020, September 15). *Gentrification doesn't have to force minority residents out of their homes. activists say there are 3 ways to protect communities*. Business Insider. Retrieved July 19, 2022, from <https://www.businessinsider.com/personal-finance/how-to-protect-longtime-residents-fro>

[m-gentrification-2020-9#:~:text=According%20to%20community%20leaders%20and,pool%20their%20 capital%20to%20buy](#)