

Airbnb Bookings Analysis

Sourav Chowdhury (Data science Aspirants)

AlmaBetter, Bangalore

Abstract:

Airbnb is an online marketplace for renting out homes/villas/ private rooms. The website charges a commission (3 to 20 percent) for every booking. Since 2008, guests and hosts have used Airbnb to expand on traveling possibilities and present a more unique, personalized way of experiencing the world.

The data used in the Analysis is the outcome for business decisions, understanding of customers' and providers' (hosts) behavior and performance on the platform, guiding marketing initiatives, implementation of innovative additional services and much more.

1. Problem Statement

Following are a few questions that we aim to answer through our analysis:

- **What can we learn about different hosts and areas?**
- **What can we learn from predictions? (ex: locations, prices, reviews, etc)**
- **Which hosts are the busiest and why?**
- **Is there any noticeable difference of traffic among different areas and what could be the reason for it?**

Variable breakdown:

1. Name : Name of the rental property.
2. Host name : Name of the owner of the property.
3. Neighbourhood : Area to which property belongs.
4. Neighbourhood group : Grouped areas consisting various neighbours.
5. Room type : Type of room it could be Private, Shared and Entire home/apt.
6. Price : The rental price of the property.
7. Minimum nights : Minimum no. of nights the property can be booked.
8. Number of reviews : Reviews received from customers.
9. Calculated host listings count : No. of listings by host.
10. Latitude/Longitude : LAT/LONG of the particular listing.

Steps involved:

- EDA
- Removing null values
- Targeting each question of problem statement one by one.
- Drawing graphs using the matplotlib library.
- Conclusion

2. INTRODUCTION

Having the data of various places of Newyork city. We can do every sort of Analysis on this data so as to find out a perfect Business solution to increase the listings or increase the pricing for various hosts. From here we can also churn out the places where there are much number of listings and which place is most busiest and to tackle down the problem there we can provide a solution where to increase the listings. As we can see that there are neighbourhood groups which eventually will help to get the predictions based on the tourist places as well as the type of property in more demand.

As the hospitality sector grows up there is also competition arises So as to make a competitive decision data is the best source. The accuracy of the prediction is important as inaccurate in predicting will lead to campaign fails. So applying Exploratory Data Analysis, will churn out the data.

In this project we are:-

1. Analyzing the trend which place is most prioritized and why.
2. Reasons of not giving preference to a certain type of property.
3. Prices as per the Neighbourhood groups.

3. Different Hosts and Areas

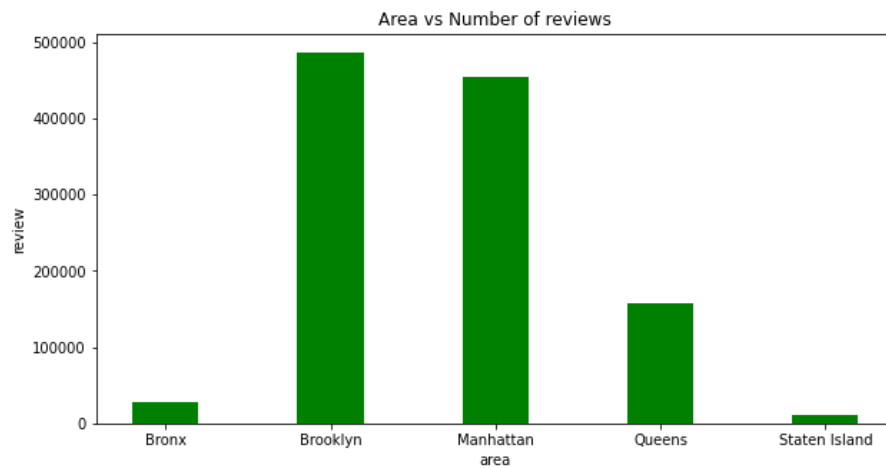
After grouping by neighbourhood groups, we have found that most of the listings are from Manhattan created Sonder (NYC), Blueground, Kara.

	host_name	neighbourhood_group	calculated_host_listings_count
13214	Sonder (NYC)	Manhattan	106929
1833	Blueground	Manhattan	53360
7273	Kara	Manhattan	14669
13213	Sonder	Manhattan	9216
6538	Jeremy & Laura	Manhattan	9216
2900	Corporate Housing	Manhattan	8281
7478	Kazuya	Queens	8137
7544	Ken	Manhattan	7500
11396	Pranjal	Manhattan	4225
9853	Mike	Manhattan	2824

4. Reviews of different Neighbourhood groups.

After doing the analysis on reviews we have found that roklyn is the most preferenced neighbourhood group and Staten is the least one to stay.

Neighborhood with most number of reviews ,means that these neighbourhood has most number of customers.

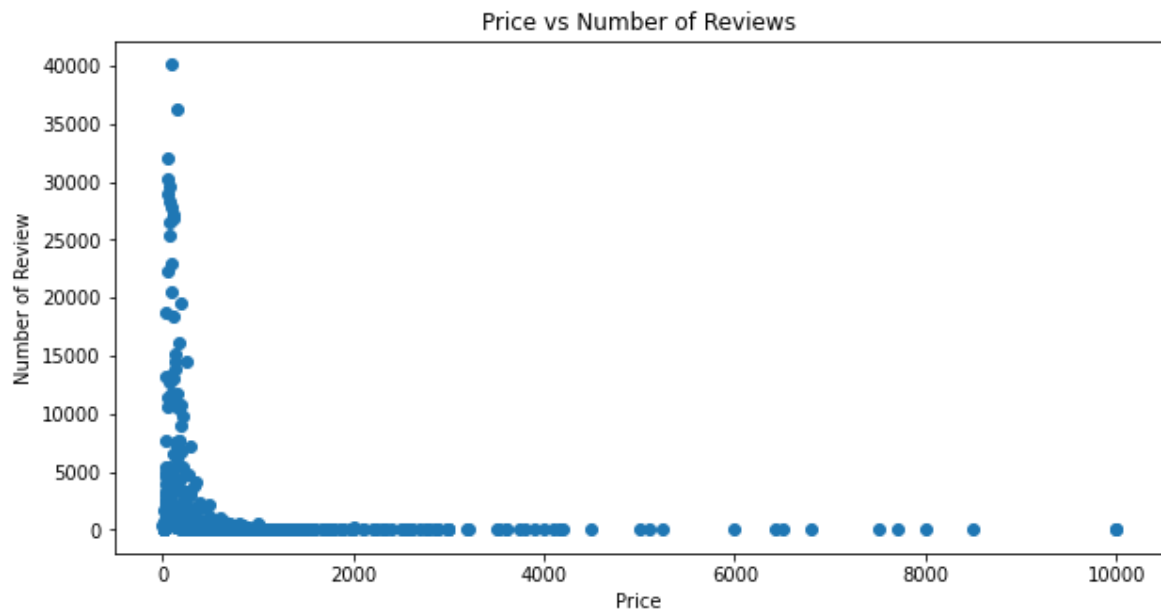


5. Price Vs Number of Reviews

This analysis help us to know which is the most preferable property type.Also the reason of choosing that type of property.

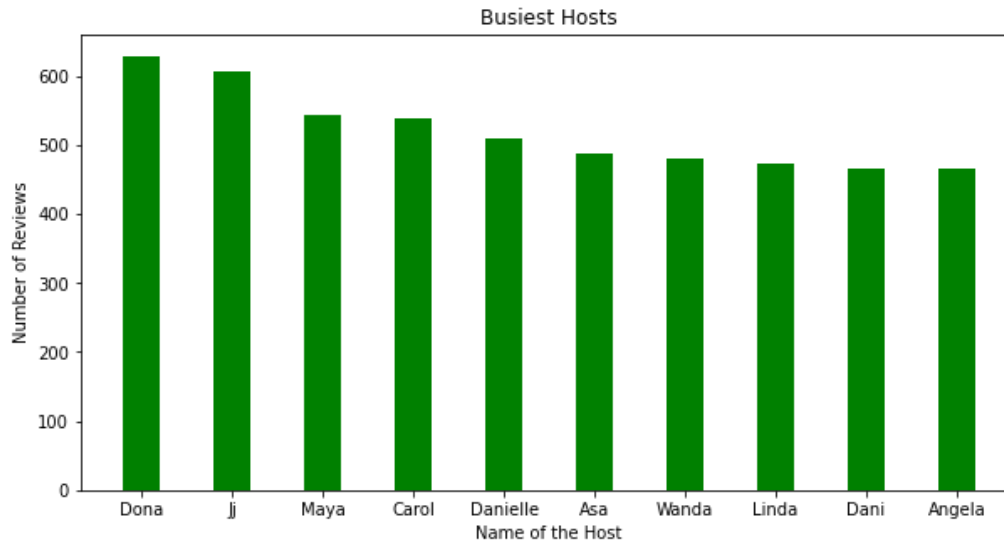
Most people prefers to stay where prices are less.

	price	number_of_reviews
0	0	377
1	10	241
2	11	113
3	12	8
4	13	9



6. Busiest Hosts and Reason

We have dig out into the host name and number of reviews . SO the host name having most number of reviews is the busiest host.



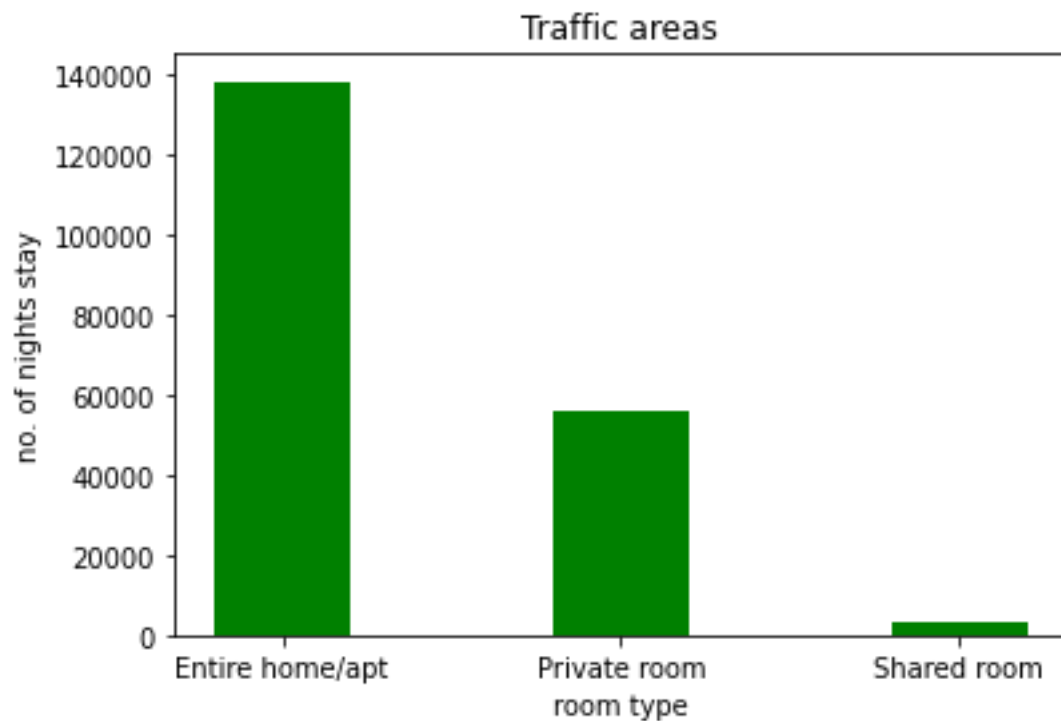
	host_name	room_type	number_of_reviews
3434	Dona	Private room	629
6332	Jj	Private room	607
8977	Maya	Private room	543
2164	Carol	Private room	540
2975	Danielle	Private room	510
1214	Asa	Entire home/apt	488

Busiest Hosts are : DONA > JI > MAYA > CAROL Because these hosts listed room type as Entire home and Private room which is preferred by most number of people.

7. Traffic among different areas and the reason behind it.

As analysis by grouping by neighbourhood groups and room type , found that people are preferring Entire home/apt or private room which are present in Broklyn,Manhattan and Queens, and which are in less price.

	neighbourhood_group	room_type	minimum_nights
6	Manhattan	Entire home/apt	138081
3	Brooklyn	Entire home/apt	62393
4	Brooklyn	Private room	56098
7	Manhattan	Private room	43463
10	Queens	Private room	17263



7. Conclusion :

The analysis of CSV file helped us to determine important factors affecting customer's preference of booking the rental property and analysed the trend where Manhattan is most preferred location of stay. People prefer to stay in the places where the prices are less.

References:

Geeksforgeeks

Wikipedia

Kaggle