

A CASE STUDY TO IDENTIFY OPPORTUNITIES TO INCREASE OVERALL REVENUE OF AIRBNB

By Sagnik Chakravarty

Agenda or objective

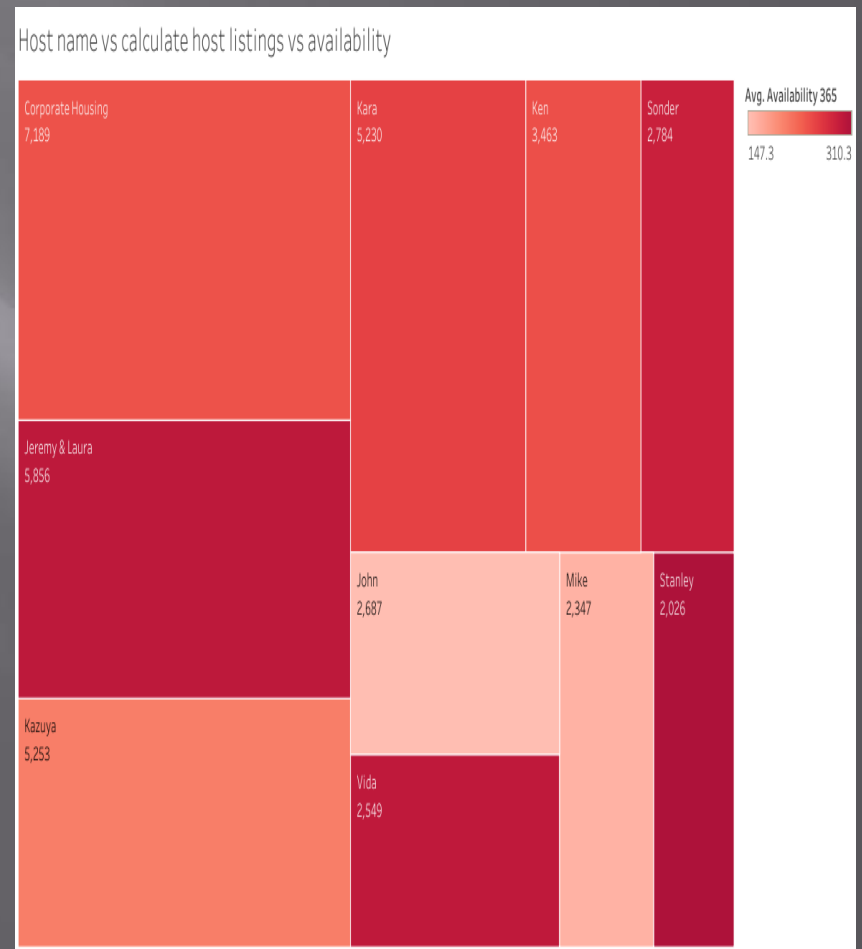
- ▣ Airbnb has seen a significant decline in revenue the past few months but now that the covid restrictions are slowly being lifted, people are starting to travel once again as the world slowly but surely moves back towards normalization.
- ▣ Therefore this is a prime time for us to get back into the upswing of business again and this study aims to identify opportunities for us to take advantage of in pursuit of this goal.
- ▣ This study aims to identify which types of hosts to acquire, the neighborhoods we need to target, price ranges preferred by our customers, most popular localities, etc so that our revenue can start it's upswing sooner rather than later.

Contents

- ▣ Best hosts to acquire.
- ▣ We should prioritize listings with higher availability.
- ▣ Manhattan and Brooklyn are the dominant neighbourhood groups which customers prefer.
- ▣ Hell's kitchen and Upper East side are the most popular neighbourhoods.
- ▣ Customer prefers price ranges between \$35 and \$250.
- ▣ Conclusion.
- ▣ Methodology

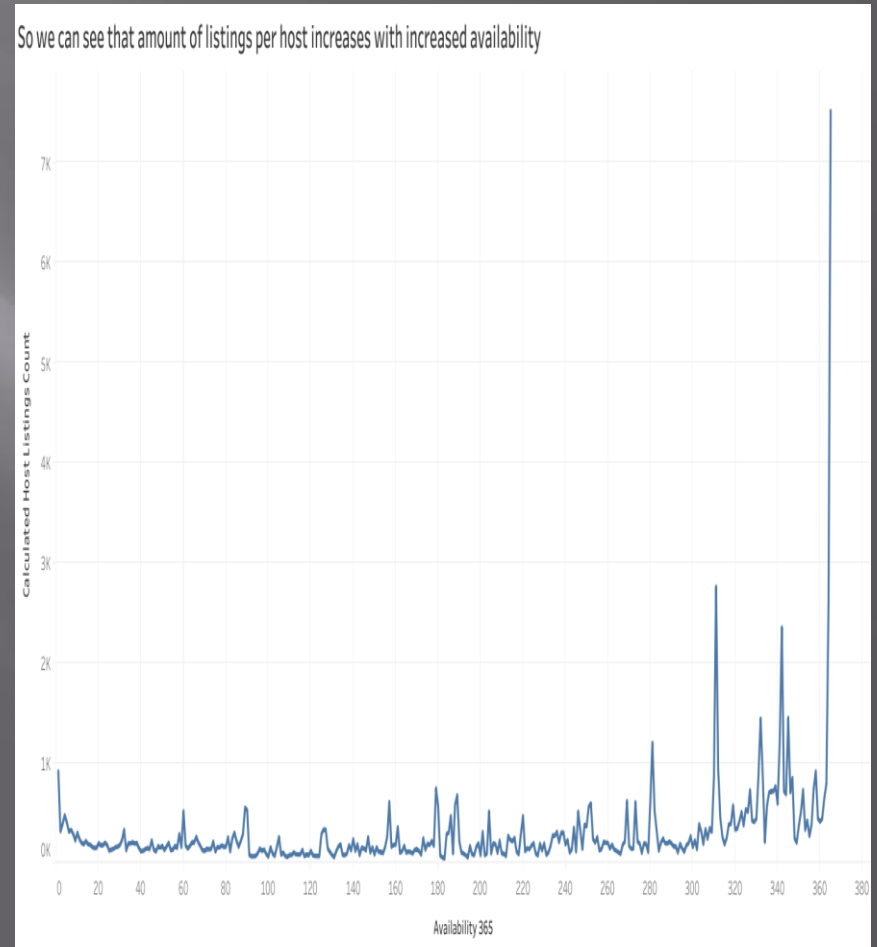
Best hosts to acquire

- ▣ Corporate Housing, Jeremy and Laura and Kara are the top 3 hosts we should target for acquisition.
- ▣ We created a heat map and checked the top 10 hosts in terms of highest amount of listings per host. We had a look at availability of their listings throughout the year as well and these 3 hosts were at the top. The others in that top 10 are being shown in the picture to the right as well.



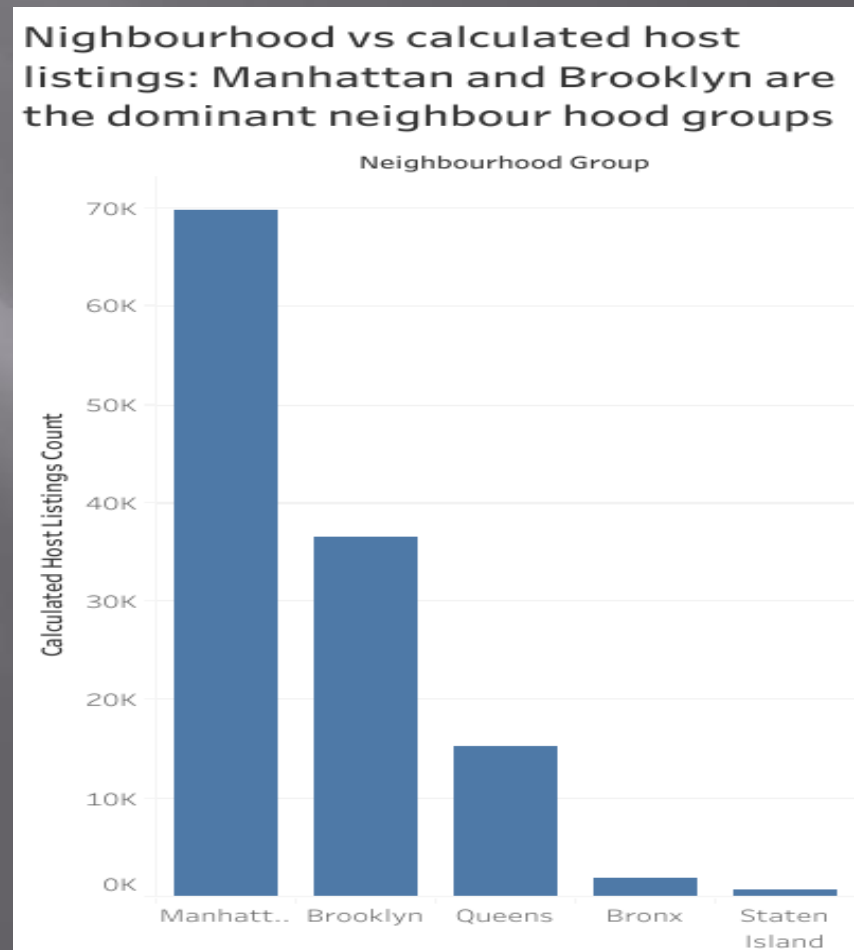
We should prioritize listings with higher availability

- ▣ We found that amount of listings per host increases with increased availability.
- ▣ As more listings mean more demand therefore higher engagement so we are treating this as our indicator.
- ▣ We plotted a line plot and saw that as availability increases, the amount of listings increase. So we should target listings that have higher availability throughout the year.

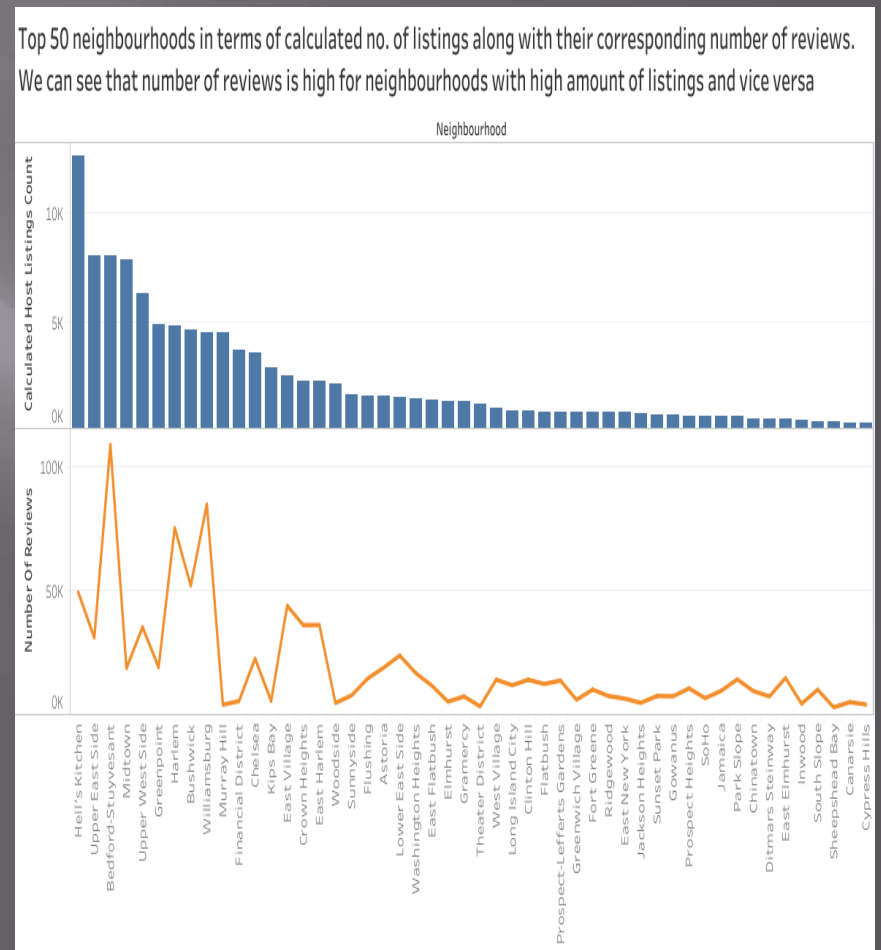


Manhattan and Brooklyn are the dominant neighbourhood groups which customers prefer

- ▣ Manhattan and Brooklyn have most amount of total listings per host. As number of listings is an indicator of demand for holiday homes so we can conclude that customers prefer to travel to Manhattan and Brooklyn the most.
- ▣ We found this by plotting a bar graph as shown in the right.

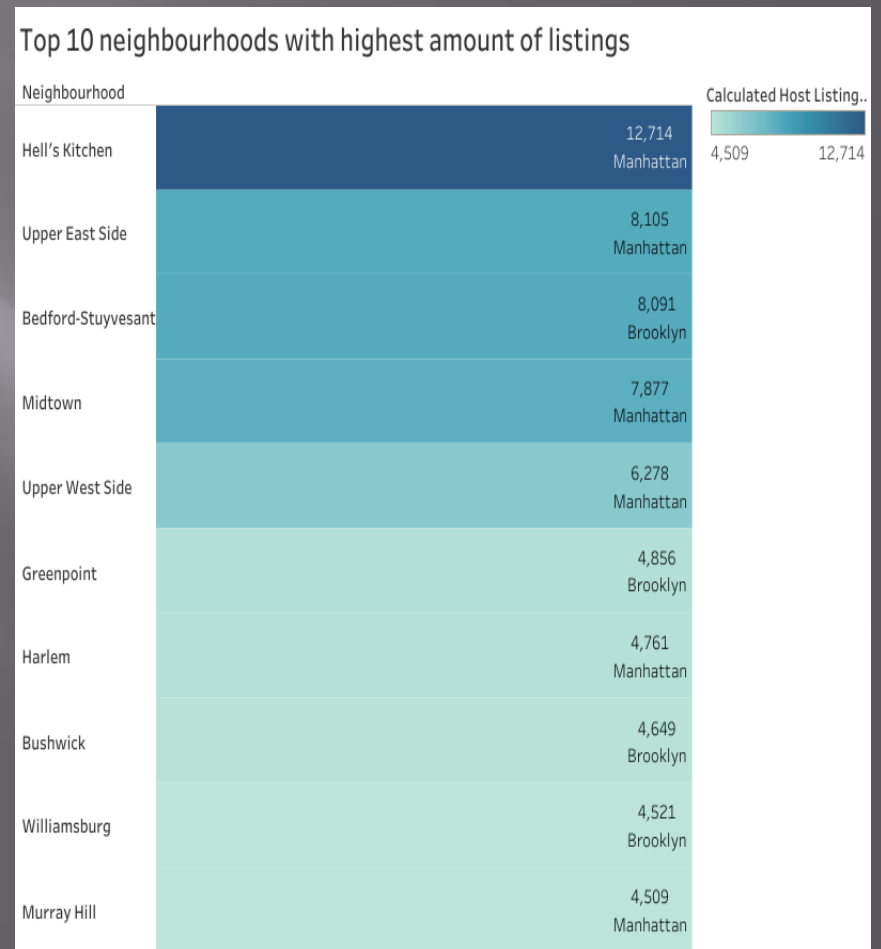


- We also found that neighbourhoods that have higher amount of listings also get more number of reviews. As we know that reviews are a good indicator of customer engagement so we can conclude that higher amount of listings lead to higher customer engagement i.e. proving our earlier point about amount of listings being an indicator of customer demand.



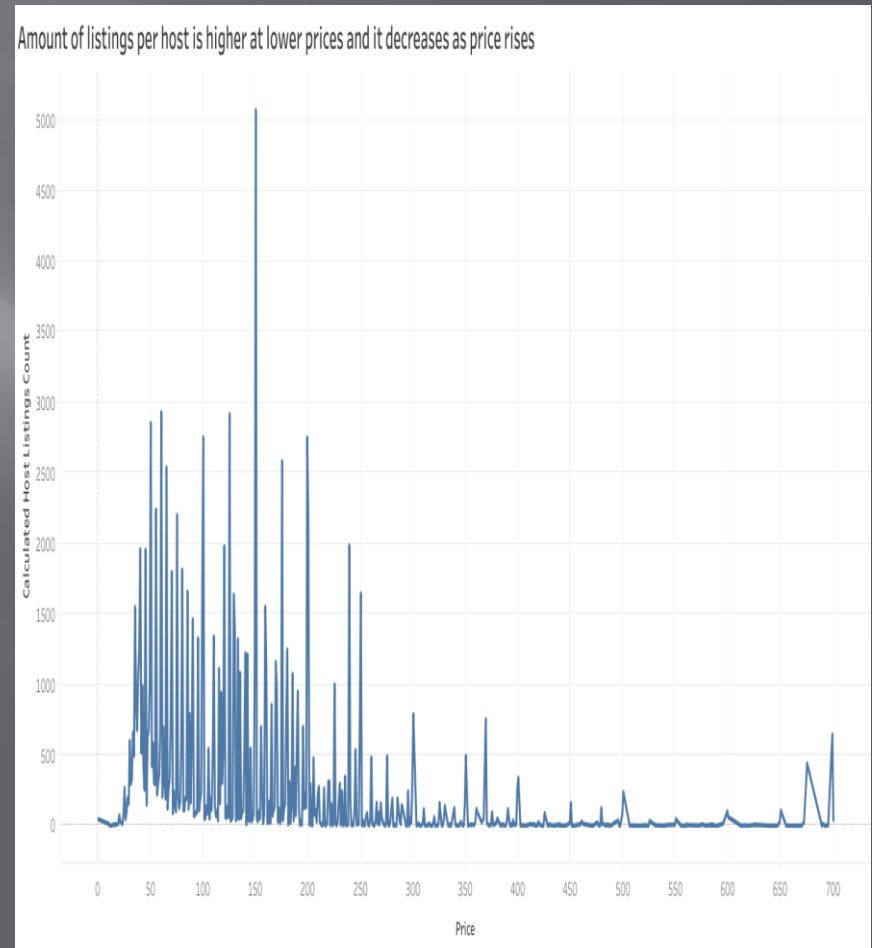
Hell's kitchen and Upper East side are the most popular neighbourhoods

- Hell's kitchen, Upper East side and Bedford Stuyvesant have the most amount of listings per host. Therefore we can conclude these are the top 3 neighbourhoods or localities with highest demand from our customer base.
- We made a top 10 neighbourhoods and all ten of them are predominantly from Manhattan and Brooklyn thereby further proving our point that Manhattan and Brooklyn are the most popular neighbourhood types.



Customer prefers price ranges between \$35 and \$250.

- ▣ We can see that with increase in prices the amount of listings per host decreases. We already concluded before that amount of listings is an indicator of demand thus conforming to the inverse relationship between price and demand.
- ▣ From the graph to the right we can see that the amount of listings is high for price range: \$35 and \$250.
- ▣ Hence we should be targeting this price range for attracting customers.



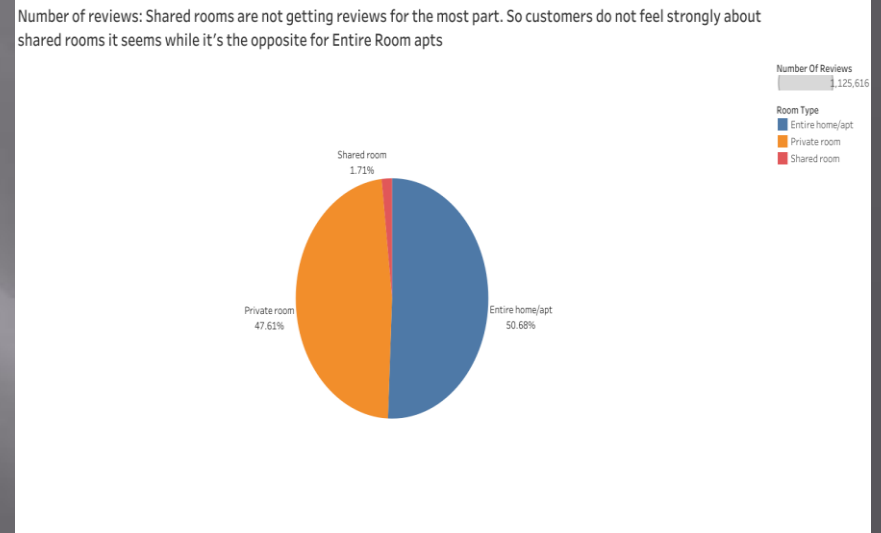
Entire home apartments are the most popular room type

- Entire home apartments have the most amount of listings with Private rooms following at second. Shared rooms have the least amount of listings per host indicating it has the least demand.
- We can also see average price is highest for entire home apartments while shared rooms have the least average price.
- So taking these observations into consideration we can conclude that it's better to focus on entire home apartments and less on shared rooms.

Entire home apt seems to receive the most amount of listings and have the highest price on average as well while shared room seems to have the least amount of listings and are the cheapest



- We can also see from this pie chart that shared rooms are not getting reviews for the most part. So customers do not feel strongly about shared rooms it seems while it's the opposite for Entire Room apartments.



Conclusion

- ▣ Corporate Housing, Jeremy and Laura and Kara are the top 3 hosts we should target for acquisition.
- ▣ We found that amount of listings per host increases with increased availability. So we should target listings that have higher availability throughout the year.
- ▣ Manhattan and Brooklyn are the dominant neighbourhood groups which customers prefer.
- ▣ We also found that neighbourhoods that have higher amount of listings also get more number of reviews. As we know that reviews are a good indicator of customer engagement so we can conclude that higher amount of listings lead to higher customer engagement i.e. amount of listings is a valid indicator of customer demand.
- ▣ Hell's kitchen and Upper East side are the most popular neighbourhoods.
- ▣ Customer prefers price ranges between \$35 and \$250.
- ▣ Entire home apartments have the most amount of listings with Private rooms following at second. Shared rooms have the least amount of listings per host indicating it has the least demand.
- ▣ We can also see from this pie chart that shared rooms are not getting reviews for the most part. So customers do not feel strongly about shared rooms it seems while it's the opposite for Entire Room apartments.

Methodology

- ▣ We first downloaded the AB_NYC_2019 dataset and imported it into a jupyter notebook for cleaning the data.
- ▣ After having a look at the dimensions of the data i.e. number of columns, rows, data types, etc we then moved on to the data cleaning stage.
- ▣ We checked for duplicates in the data. None were found so we moved on to checking for missing values.
- ▣ We checked for missing values in all the columns and checked percentage of the missing values for each column as well.
- ▣ We saw that columns name and host_name have 0.03% and 0.04% amount of null values which is very miniscule. So we had no issue in dropping these rows from the dataset.
- ▣ We also see that last_review and reviews_per_month has 20.54% null values. We know that last_review is supposed to be a date column and if we replace the missing values with the mode or a 'missing' category. It won't help us with the further analysis as it won't be treated as a date column. So we dropped these rows.
- ▣ Now that the all the null values are taken care of we moved to the outlier checking process.
- ▣ We plotted box plots for the numeric variables to check for outliers.
- ▣ We identified the columns that may have outliers from the boxplots and then conducted univariate analysis on them to further investigate if there are legitimate outliers or not.
- ▣ For e.g. in the column minimum_nights we found that there is a massive difference between the max value(1250) and 99% percentile value(31) which means there is something wrong. So we capped minimum_nights to 40.
- ▣ After dealing with outliers we finally arrived at the clean dataset which had 38003 rows and 16 columns.
- ▣ We then finally exported the clean dataset as an excel workbook so that we can import the same to tableau for exploratory analysis and generating insights.
- ▣ We then imported the clean dataset into tableau for the analysis, creating graphs and generating insights.