**Capstone Project Submission Airbnb Booking Analysis**

|  |
| --- |
| **Summary** |
| Airbnb is an open online platform where people list their own housing for rent. Since 2008, it has grown in popularity and specially for those community which frequently use to travel. It is becoming a strong competitor to the hotel industry. It has millions of listing, which generate lots of data. We can analyze these data for making business decision, for looking best room type etc.  As a first step I take the overview of data, where I specially made our focus on understanding what each column means. So that I can be clear from what perspective I have to analyze this data. After understanding different column, I marked few important columns. These columns are neighborhood group, room type, price, minimum nights, reviews per month. Then I did some basic visualization to see is there any correlation among columns.    Now I started cleaning our data. So, I first identified the null values and I replaced this null value according to their data type. After dealing with null values I moved on to those columns which I don’t need. Then replaced few data which don’t make sense with other values. Finally, our data is ready for EDA.   1. Area Analysis: The first focus I put on is ‘Area’. Here first I looked in to most listings in the neighbourhood. From this I get to know that Sonder (NYC) neighbourhood group has the most number of listings. Then I looked at neighborhood group which has the most number of listings, which came out to be Manhattan. From this analysis I made inference that if Manhattan and Brooklyn are both combined, almost 85% listings are done every day just in this regions. 2. Price Analysis: Here I first looked into the average price of listings. From this analysis I made inference that Manhattan is the costliest neighbourhood group although it was on top on the number of listings and Staten Island which was having the least number of listings is not the cheapest neighbourhood group among all. Airbnb’s with price range of 150 to 200 dollars are above 4k+ listings. Also if a person is low on budget, he may prefer to book an Airbnb in Bronx. 3. Review analysis: I get to know that Airbnb’s having 4 reviews per month are around 130 and Airbnb’s having 1 review per month are around 178. The inference which come out from this result if is people are rating the hosts with poor reviews more than good reviews. This shows that people will not prefer to come again to the rooms they were not comfortable. 4. Preferability analysis: To confirm the above inference I analysed the type of rooms people prefer to stay. I found out that ‘Entire Home’ is room type people prefers the most almost in every neighbourhood. The inference which come out from this result if host is having entire room then he will be making good money. I was astonished that irrespective of any neighborhood group, entire home is making way ahead revenue then other room types. In Brooklyn, Private rooms are more preferred instead of Entire Home.   This was all about my analysis that I did, and based on result and my inferences I am making following conclusions:   1. Entire home/apt is highly expensive. 2. Manhattan living cost is highest, Bronx living cost is lowest. 3. Cheapest listing is Bronx apart. 4. Manhattan have highest no. of listing. 5. In Manhattan entire home is mostly preferred but in Brooklyn ratio between entire home and private room is 50:50. 6. Sonder has maximum property in New York. |
| **Contributor Roles** |
| Ritik Vaidande ([vr171k@gmail.com](mailto:vr171k@gmail.com))   * Understanding Data   Understanding different column  Having overview of data   * Data Cleaning * Removing null values * Replacing zero value of price with mean value * Dropping unused rows. * Exploratory Data Analysis * Price analysis * Area analysis * Review analysis * Preferability analysis * Visualizations * Bar graphs and Pie charts * Pair plot * Inference * Summarization * Conclusions   **GitHub Repo and GDrive Link** |
| Github Link:-  <https://github.com/vaidande/Airbnb-Booking-Analysis>  Drive Link:-  https://drive.google.com/drive/folders/15i68SC23vqq06tuFNf40yXULLENQJ3DW?usp=sharing |
|  |