Capstone Project Submission

Instructions:
i) Please fill in all the required information.
ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
Member Name: • Shaik Ahmad Basha ahmadshaik982basha@gmail.com Contribution: • Exploring Data • Data Wrangling • Data Cleaning • Checking for Null Values • Checking for Duplicated Values • Analyze type of rooms highly booked • Analyze listing across various locations • Analyze average prices for different locations and room types • Analyze number of reviews for different locations and room types • Analyze number of night stays for different locations and room types • Analyze availability of rooms for different locations and room types • Performed EDA on Data
Please paste the GitHub Repo link.
GitHub Link: - https://github.com/ahmedshaik982/EDA_On_Airbnb_Bookings_Analysis
Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

Problem Statement:

For this project, we will be analyzing Airbnb's New York City (NYC) data of 2019. This dataset contains listings information such as listing name, host name, room types, minimum night stays, availability, area, reviews etc.

Our main objective behind this project is to explore and analyze the data to discover the key understandings. For this, we will explore and visualize the dataset from Airbnb in NYC using basic exploratory data analysis techniques.

Approaches:

The first step imported all the necessary libraries like NumPy, Pandas etc. and then collected the data. I started with understanding the data like what are the columns and their meanings and data types.

After that, the second step is data preprocessing. Data preprocessing is a process where raw data is converted into clean data. Our data has 48895 rows and 16 columns. I removed null values and unnecessary columns from the data for Exploratory data analysis

And then I moved on to Exploratory data analysis where I analyzed various features to get useful insights from the data. For Visualization, I used bar charts, heatmaps, scatter plots, pie charts, violin plots and multiple bar charts.

Conclusions:

Most people are preferring Manhattan and Entire home/apt and listings with low price.

The price of listings is mostly in the range of 0 to 1000 and a few listings are in range of 1000 to 10000. The number of listings with price less than 1000 are 48619 and more than 1000 are 239.

Manhattan area has the highest average price (196.87) and the least is Bronx (87.4967) and Entire home/apt has highest average price (211.806) than others
Brooklyn and Entire home/apt have the highest number of reviews has the highest number of reviews. Manhattan and Entire home/apt have more night stays

Staten Island has availability of rooms with a greater number of days than others and the shared room category has availability of rooms with a greater number of days. There are total of 1294 listings that are available for 365 days.