Project Summary
Airbnb Booking Analysis

The Airbnb booking dataset provides detailed information of the past bookings of room types with their respective neighborhoods and customer reviews. There are details regarding the price, availability, neighborhood, host, room type etc. Exploring the distributions as well as the relationships between the columns will give us a very thorough understanding of the factors determining the trends in booking.

We started by getting some basic information about the data types, null values as well as some descriptive statistics of the features. This allowed us to decide which columns and data to keep while removing the null values. The statistical analysis also allowed us analyze and keep the required data.

After cleaning the data, we simply studied the distribution of bookings with respect to the columns or features. This allowed us to get some important information on different neighborhoods and their property listings and their prices etc.

Some of the distributions were compared based on the existing knowledge of the relationship between the columns. This multivariate analysis helped us make some interesting observations, namely about the price variation, price variation between different neighborhoods, and most expensive neighborhoods and correlation between price and availability. etc.

Bar plot showed the mean price of different room types and inferred that Entire home/apt is the expensive one as it has more amenities, followed by a private room and the shared rooms lesser than the former rooms. Pie plot showed the room type which was the most booked/wished for and the percentage was higher for an entire room/apartment followed by a private room and shared rooms.

Heat map showed the correlation between availability and price. It showed that there is a moderate negative correlation between price and availability, which means low supply and high demand, which will in turn gives higher price.

A large percentage of the customers were found in Manhattan and Brooklyn area which may be because of number of reasons like climate, landscape, connectivity etc. and most booked room type was either an entire room/apartment or a private room.

Through this exploratory analysis, we are able to answer many important questions that can guide Airbnb to take the appropriate steps to improve their business as well as satisfy the customer demands.

Contributor Roles:
1. **Sharath Diwakar** (*mailmesharathd@gmail.com*)
   a. Data cleaning
   b. Data wrangling
   c. EDA
      i. Mean price of each room type
      ii. Most wished/booked room type
      iii. Mean price of each room type in all neighborhood groups
      iv. Footfall for each room type in all neighborhood groups
      v. Availability & price correlation
      vi. Location and density of properties in each neighborhood
      vii. Most successful host
      viii. Most expensive neighborhood in each neighborhood groups

GitHub Link:

https://github.com/Sharath2021/Airbnb-Bookings-Analysis