

# Capstone Project Submission

## Instructions:

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

### **Team Member's Name, Email and Contribution:**

Contribution Role :

**1. Akash Salmuthe ([akashsalmuthe30@gmail.com](mailto:akashsalmuthe30@gmail.com)) :-**

**1. Data Cleaning and Data Preprocessing**

**2. General and Business Analysis**

- All neighbourhood groups
- Availability of number of days for room when listing is available for booking
- Room type
- Neighbourhood group based on latitude and longitude
- Price exploration
- Relation between neighbourhood group and availability of number of days for room
- Property owned by each neighbourhood group
- Property distribution on room type
- Most successful host
- Most expensive neighbourhood in each group

**Please paste the GitHub Repo link.**

Github Link :- <https://github.com/AkashSalmuthe/Airbnb-Bookings-Analysis---Akash-Salmuthe>

**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:**

Since 2008, guests and hosts have used Airbnb to expand on traveling possibilities and present a more unique, personalized way of experiencing the world. Today, Airbnb became one of a kind service that is used and recognized by the whole world. Data analysis on millions of listings provided through Airbnb is a crucial factor for the company. These millions of listings generate a lot of data - data that can be analyzed and used for security, 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.

This dataset has around 49,000 observations in it with 16 columns and it is a mix between categorical and numeric values.

Explore and analyze the data to discover key understandings (not limited to these) such as :

- 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?

**Approaches:**

First thing first, I 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.

**Conclusion:**

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.

**Drive link :-**

**<https://drive.google.com/drive/folders/1oFeg27Ucx4dTIL85Qofb84MEjoNLFHhO?usp=sharing>**