**AIRBNB BOOKING ANALYSIS**

**Ankush, Keerthi**

**AlmaBetter, Bangalore**

**Introduction**

Airbnb, Inc. is an American company that operates an online marketplace for lodging, primarily homestays for vacation rentals, and tourism activities. Based in San Francisco, California, the platform is accessible via website and mobile app.

Our experiment can help understand what could be the reason for the booking of airbnb by feature selection, data analysis and prediction with python.

**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..

The main purpose for analyzing this data is to discover key understandings like about different neighborhoods, different room types, highest room price and many more.

* **About data**

So, in our dataset we have

* Id- We have different unique I'd
* Name - Name of listing
* Host\_id - unique host Id
* host\_name - name of hosts
* Neighbourhood\_groups - Collection of neighborhood
* Neighborhood - Different neighborhoods
* latitude - latitude coordinates
* longitude - longitude coordinates
* room\_type - Different types of room
* price - price in dollars
* minimum\_night - amount of minimum night spend
* Number\_of\_reviews - number of review
* last\_review - last review
* reviews\_per\_month - number of reviews per months
* calculated\_host\_listing\_count - amount of listing per host
* availablity\_365 - number of days when booking is available

**Steps involved:**

* **Exploratory Data Analysis**

After loading the dataset we use python, sql and excel to analyze our data to get our results. This process helped us figuring out various aspects and relationships among the target and the independent variables.

* **Import some important library**

After loading our dataset we import some important libraries too like pandas, numpy, seaborn etc. These libraries help us alot to analyze and to get the results.

* **Null values Treatment**

Our dataset contains many null values which might tend to disturb our analysis hence some of them we dropped and some of them we replace values with 0 and average at the beginning of our project inorder to get a better result.

.

* **Visualization**

We use different types of graphs like bar graph, pie chart , box plot graph and many more so that we can visualize data in a better way. It also helps to understand the statements

* **Conclusion**

That's it! We reached the end of our exercise.

Starting with loading the data so far we have done EDA , null values treatment, important library and visualization.

* Most Airbnb services are located in Manhattan and Brooklyn.
* Manhattan has the highest mean price almost 200.
* The Bronx has a minimum mean price around 90.
* Fort Wadsworth (Staten Island) has the highest room price among all neighborhoods.
* Entire home/apt has the highest price of all room types.
* Manhattan and Brooklyn have a maximum number of hosts
* Sonder(NYC) is the most listed host with more than 300 listings.
* Row NYC is the most reviewed host with more than 50 reviews per month.
* Shared room available 42% of the time which is highest in all room types.
* Queens and Staten Island have the maximum traffic among all.

**References-**

* Stack overflow
* GeeksforGeeks