**Airbnb Listings Exploratory Data Analysis Project**

**Goal: The goal of the project is to analyse the New York Airbnb data set.**

**Problem Statement:**

* **Let's Explore and analyze the Data Set and find some insights(A few Questions Listed Below)**

### 1. What can we learn about different hosts and areas?

### 2. What do we learn from room type and their prices according to area?

### 3. What can we learn from Data? (ex: locations, prices, reviews, etc)

### 4. Which hosts are the busiest and why is the reason?

### 5. Which Hosts are charging higher prices?

### 6. Is there any traffic difference among different areas and what could be the reason for it?

### 7. What is the correlation between different variables?

### 8. What is the room count in overall NYC according to the listing of room types?

* **Objective**

The objective of the project is to build a dashboard that will help to understand Airbnb data and its pattern.

* **Introduction**

Airbnb is a service that allows property owners to rent out their spaces/condos to travelers looking for a place to stay.

Airbnb was started in 2008 by Brian Chesky and Joe Gebbia based in San Francisco California. The platform is accessible via a website and mobile app. This dataset describes the listing activity and metrics in Amsterdam, Netherlands for 2019. Content This data file includes all the needed information to find out more about hosts, geographical availability, and necessary metrics to make predictions and draw conclusions.

The objective of the project is to perform data visualization techniques to understand the insight of the data. This project aims to apply Exploratory Data Analysis (EDA) and to get a visual understanding of the data.

* **About the Project**

Airbnb is an American company that facilitates an online marketplace for lodging, primarily homestays for vacation rentals, and tourism activities. It connects travelers with local hosts who want to rent out their homes to people who are looking for accommodations in that locality. On the other hand, this platform enables hosts to list their available space and earn extra income in the form of rent and it also enables travelers to book unique homestays from local hosts, saving them money and giving them a chance to interact with locals.

* **Dataset Information**

The dataset consists of 782320 rows and 16 columns followed by:

1. Id: This is the id no. of our datasheet.

2. Name: It stands for the name of the property listed by the host.

3. Host\_id: It is the identity number of the hosts who have registered on

the Airbnb website.

4. host\_name: It stands for the name of the property listed by the host.

5. neighbourhood\_group:

6. Neighborhood: These are the names of the neighborhood or locations

present in the city.

7. Latitude: These represent the coordinates of the latitude of the

property listed.

8. Longitude These represent the coordinates of the longitude of the

property listed.

9. room\_type: This represent the various types of room listed by host.

10. Price: This is the rent of the property listed in euro.

11. minimum\_nights: This represent the minimum nights to stay.

12. number\_of\_reviews: This represent the number of customers reviewd

the property.

13. last\_review: This represent the date when the property was last

reviewed.

14. reviews\_per\_month: This represent that how many people reviews

15. calculated\_host\_listings\_count: It shows that how many hosts are ‘ ‘there.

16. availablity\_365: This represent that how many rooms are available at

all time.

* **TOOLS USED**

**Python**

**Matplotlib**

**Seaborn**

**Pandas**

* **ARCHITECTURE DESCRIPTION.**

The architecture for an Airbnb data analysis project typically involves several components and steps to process and analyze the data. Here's a description of the architecture for an Airbnb data analysis project:

* **Data Collection:** The first step is to gather the necessary data from Airbnb. This can include data on listings, reviews, bookings, host information, and more. The data can be obtained by downloading from Kaggle's publicly available websites.
* **Data Cleaning and Preprocessing:** Before performing analysis, the collected data needs to be cleaned and pre-processed. This involves tasks such as handling missing values, removing duplicates, removing unnecessary columns standardizing data formats, and transforming the data into a suitable structure for analysis.
* **Data Analysis and Visualization:** Once the data is prepared, various data analysis techniques can be applied to gain insights. This can involve statistical analysis and exploratory data analysis. Visualization tools, such as Python libraries like Matplotlib or Seaborn, are used to create visual representations of the data and communicate findings effectively.

**Data Transformation:** In the Transformation Process, we will Transform our original datasets excel fil into jupyter notebook for data Exploration and performing Exploratory Data Analysis using python programming language.