Airbnb Analysis

Project Title	Airbnb Analysis
Skills take away From This Project	Python scripting, Data Preprocessing, Visualization, EDA, Streamlit, MongoDb, PowerBI or Tableau
Domain	Travel Industry, Property Management and Tourism

Problem Statement:

This project aims to analyze Airbnb data using MongoDB Atlas, perform data cleaning and preparation, develop interactive geospatial visualizations, and create dynamic plots to gain insights into pricing variations, availability patterns, and location-based trends. The objectives are to:

1. Establish a MongoDB connection, retrieve the Airbnb dataset, and ensure efficient data retrieval for analysis.
2. Clean and prepare the dataset, addressing missing values, duplicates, and data type conversions for accurate analysis.
3. Develop a streamlit web application with interactive maps showcasing the distribution of Airbnb listings, allowing users to explore prices, ratings, and other relevant factors.
4. Conduct price analysis and visualization, exploring variations based on location, property type, and seasons using dynamic plots and charts.
5. Analyze availability patterns across seasons, visualizing occupancy rates and demand fluctuations using suitable visualizations.
6. Investigate location-based insights by extracting and visualizing data for specific regions or neighborhoods.
7. Create interactive visualizations that enable users to filter and drill down into the data.
8. Build a comprehensive dashboard using Tableau or Power BI, combining various visualizations to present key insights from the analysis.

Data: Data

1. **Create a MongoDB Atlas Account:** Sign up for a MongoDB Atlas account by visiting the MongoDB Atlas website and follow the registration process to set up your account and create a new project.
2. **Set Up a Cluster:** Within your MongoDB Atlas project, set up a cluster. Choose the cloud provider and region for hosting your data, configure the cluster specifications, and create the cluster. This will serve as the database environment for storing the sample data.
3. **Load the Airbnb Sample Data**: Once your cluster is set up, access the MongoDB Atlas dashboard. In the left-hand navigation menu, click on "Database Access" to create a database user with appropriate permissions for accessing and loading data. Then, select "Network Access" to set up IP whitelisting or configure other security measures.
4. **Import Sample Data:** From the MongoDB Atlas dashboard, navigate to the "Clusters" page and click on your cluster. In the cluster view, select the "Collections" tab and click on the "Sample Data" button. Choose the "Load Sample Dataset" option, and MongoDB Atlas will import the Airbnb sample data into your cluster. The sample data typically includes collections for listings, reviews, and users.

Example AIRBNB data structure:

```
{"_id": "unique_listing_id",

 "name": "listing_title",

 "description": "listing_description",

"host_id": "unique_host_id",

 "host_name": "host_name",

 "neighbourhood": "neighbourhood_name",

 "location": {

         "type": "Point",

          "coordinates": [longitude, latitude]

          },
```

```
    "price": "listing_price",

     "availability": {

                    "start_date": "YYYY-MM-DD",

                    "end_date": "YYYY-MM-DD"

                     },

    "amenities": ["amenity_1", "amenity_2", ...],

    "rating": "average_rating",

     "reviews": [

                  {

                   "reviewer_id": "unique_reviewer_id",

                   "reviewer_name": "reviewer_name",

                   "comment": "review_comment",

                    "rating": "review_rating"

                   }, ...

                   ], ...

          }
```

Approach:

1. **MongoDB Connection and Data Retrieval:** Establish a connection to the MongoDB Atlas database and retrieve the Airbnb dataset. Perform queries and data retrieval operations to extract the necessary information for your analysis.

2. **Data Cleaning and Preparation:** Clean the Airbnb dataset by handling missing values, removing duplicates, and transforming data types as necessary. Prepare the dataset for EDA and visualization tasks, ensuring data integrity and consistency.

3. **Geospatial Visualization:** Develop a streamlit web application that utilizes the geospatial data from the Airbnb dataset to create interactive maps.

Visualize the distribution of listings across different locations, allowing users to explore prices, ratings, and other relevant factors.

4. **Price Analysis and Visualization:** Use the cleaned data to analyze and visualize how prices vary across different locations, property types, and seasons. Create dynamic plots and charts that enable users to explore price trends, outliers, and correlations with other variables.

5. **Availability Analysis by Season:** Analyze the availability of Airbnb listings based on seasonal variations. Visualize the occupancy rates, booking patterns, and demand fluctuations throughout the year using line charts, heatmaps, or other suitable visualizations.

6. **Location-Based Insights:** Investigate how the price of listings varies across different locations. Use MongoDB queries and data aggregation techniques to extract relevant information for specific regions or neighborhoods. Visualize these insights on interactive maps or create dashboards in tools like Tableau or Power BI.

7. **Interactive Visualizations:** Develop dynamic and interactive visualizations that allow users to filter and drill down into the data based on their preferences. Enable users to interact with the visualizations to explore specific regions, property types, or time periods of interest.

8. **Dashboard Creation**: Utilize Tableau or Power BI to create a comprehensive dashboard that presents key insights from your analysis. Combine different visualizations, such as maps, charts, and tables, to provide a holistic view of the Airbnb dataset and its patterns.

Reference Documents:

📄 Topic: MongoDB, EDA, SQL, PANDAS - Skill Enhancement Session

The learning outcomes of this project are:

1.**MongoDB Atlas:** Gain proficiency in working with MongoDB Atlas to store and retrieve the Airbnb dataset, developing skills in data management with a NoSQL database technology.

2.**Streamlit Web Application:** Build a user-friendly web application using Streamlit, enhancing skills in web application development for interactive data exploration and visualization.

3.**Python Data Analysis:** Utilize Python for data cleaning, analysis, and visualization tasks, developing expertise in Python libraries such as Pandas and NumPy for data manipulation.

4.**Geospatial Analysis:** Leverage geospatial libraries like GeoPandas or Folium for geospatial data processing and visualization, gaining knowledge and skills in analyzing and visualizing spatial patterns.

5.**Tableau or Power BI:** Create interactive dashboards using tools like Tableau or Power BI, refining skills in data visualization and dashboard creation for comprehensive data presentation.

6.**Data Cleaning and Preparation:** Develop proficiency in cleaning and preparing the Airbnb dataset, including handling missing values, duplicates, and data type conversions, ensuring data quality and consistency.

7.**Data Visualization Techniques:** Master data visualization techniques to effectively communicate insights, developing skills in creating visually appealing and informative charts, maps, and plots**.**

8.**Problem-Solving Skills:** Apply analytical skills to analyze pricing dynamics, availability patterns, and other factors, developing problem-solving abilities in extracting valuable insights from data.

9.**Data-Driven Decision Making:** Enhance decision-making skills by enabling stakeholders to make informed choices based on the insights and visualizations provided by the project.

10.**Collaboration and Project Management:** Strengthen collaboration and project management skills through the end-to-end development of the project, including task planning, coordination, and timely delivery of project milestones.

Project Evaluation metrics:
- You are supposed to write a code in a modular fashion (**in functional blocks**)
- Maintainable: It can be maintained, even as your codebase grows.
- Portable: It works the same in every environment (operating system)
- You have to maintain your code on **GitHub**.(Mandatory)

- You have to keep your **GitHub** repo public so that anyone can check your code.(Mandatory)
- Proper readme file you have to maintain for any project development(Mandatory)
- You should include basic workflow and execution of the entire project in the readme file on **GitHub** (Mandatory)
- Follow the coding standards: https://www.python.org/dev/peps/pep-0008/
- You need to Create a Demo video of your working model and post in **LinkedIn(Mandatory)**

PROJECT DOUBT CLARIFICATION SESSION (PROJECT AND CLASS DOUBTS)

About Session: The Project Doubt Clarification Session is a helpful resource for resolving questions and concerns about projects and class topics. It provides support in understanding project requirements, addressing code issues, and clarifying class concepts. The session aims to enhance comprehension and provide guidance to overcome challenges effectively.
Note: Book the slot at least before 12:00 Pm on the same day

Timing: Tuesday, Thursday, Saturday (5:00PM to 7:00PM)

Booking link :https://forms.gle/XC553oSbMJ2Gcfug9

LIVE EVALUATION SESSION (CAPSTONE AND FINAL PROJECT)

About Session: The Live Evaluation Session for Capstone and Final Projects allows participants to showcase their projects and receive real-time feedback for improvement. It assesses project quality and provides an opportunity for discussion and evaluation.
Note: This form will Open on Saturday and Sunday Only on Every Week

Timing: Monday-Saturday (12:00PM to 1:00PM)

Booking link : https://forms.gle/1m2Gsro41fLtZurRA