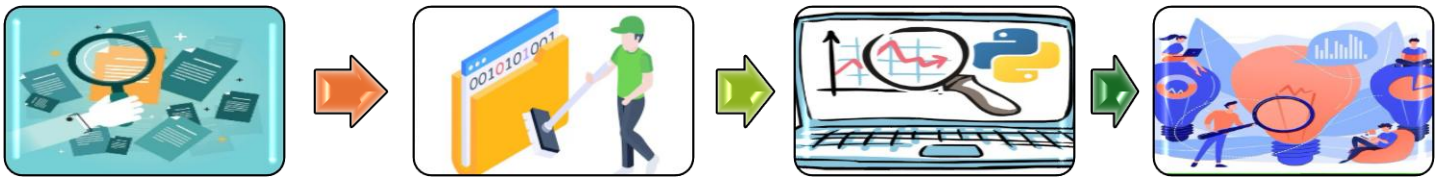
ATLIQ HOTELS DATA ANALYSIS PROJECT

Overview

The project is titled "**AtliQ Hotels Data Analysis Project**", indicating that the analysis focuses on understanding and deriving insights for AtliQ Hotels.



Import and Exploration | **Cleaning and Preparation** | **Analysis and Visualization** | **Insights and Recommendations**

Step 1: Data Import and Exploration

(A) Importing Library

The Python library **pandas** is imported to enable data manipulation and analysis:

```
import pandas as pd
```

This step ensures the required tools are ready for handling data in subsequent steps.

(B) Description of Datasets

Listed the datasets used in the project:

- **dim_date.csv**: Contains date-related information.
- **dim_hotels.csv**: Stores hotel details.
- **dim_rooms.csv**: Provides room details.
- **fact_aggregated_bookings.csv**: Summarized booking data.
- **fact_bookings.csv**: Detailed booking records.

(C) Reading Bookings Data

A markdown cell introduced the task of reading booking data into a DataFrame:

(D) Loading Data into DataFrames

- Loaded data from the CSV files into individual pandas DataFrames using commands like:
- date_df = pd.read_csv('dim_date.csv')
- hotels_df = pd.read_csv('dim_hotels.csv')
- rooms_df = pd.read_csv('dim_rooms.csv')
- agg_bookings_df = pd.read_csv('fact_aggregated_bookings.csv')
- bookings_df = pd.read_csv('fact_bookings.csv')

This step ensured the data was available for analysis.

(E) Exploratory Data Analysis (EDA)

Conducted exploratory analysis using:

a. **head**() to view the first few rows of each dataset.

b. **info**() to check structure and data types.

c. **isnull**(),.**sum**() to identify null values.

Analyzed unique values for categorical columns.

Step 2: Data Cleaning

Steps Performed:

(A) Cleaned the data through:

- Handling missing values (filling or dropping nulls).
- Renaming columns for consistency.
- Removing duplicates for data integrity.

Example command:

```
bookings_df.drop_duplicates(inplace=True)
```

(B) Data Merging and Transformation

Steps Performed:

- Merged datasets using common keys such as hotel IDs, date fields, and room IDs:

```
merged_df = bookings_df.merge(hotels_df, on='hotel_id').merge(date_df, on='date_id')
```

- Performed transformations to create new columns, like revenue per booking or total occupancy.

Step 3: Visualization and Insights

Steps Performed:

- **Generated visualizations to highlight key metrics:**
  - **Revenue Trends**: Line charts for monthly or yearly revenue.
  - **Occupancy Rates**: Bar charts for occupancy comparisons across hotels or time periods.
  - **Top-performing Hotels**: Pie charts or bar graphs.
- Used tools like **matplotlib** and **seaborn:**

```
import matplotlib.pyplot as plt
import seaborn as sns
```

Example Command: sns.barplot(x='hotel_name', y='revenue', data=merged_df)

plt.show()

Step 4: Insights and Recommendations

- **Derived insights such as:**
  - Identifying high-revenue hotels.
  - Recognizing underperforming properties.
  - Seasonal trends in bookings and revenue.
- Provided actionable recommendations to improve hotel performance and customer satisfaction.

Conclusion

This notebook systematically addresses AtliQ Hotels' data analysis, from importing datasets to cleaning, merging, analysing, and visualizing. Each step ensures clarity and provides actionable insights to drive informed decision-making.