

## Project Title

### End-to-End Airbnb Listings Analysis using Python, MySQL & Power BI

---

#### Project Objective

The objective of this project is to analyze Airbnb listing data to uncover insights related to **pricing trends, availability, host performance, and location-based patterns**. The project demonstrates an end-to-end data analytics workflow — from raw data extraction and cleaning to database management and interactive dashboard visualization.

---

#### Tech Stack Used

Layer	Tools
Data Processing	Python (Pandas, NumPy)
Database	MySQL
Visualization	Power BI
Data Source	Airbnb Listings Dataset (CSV/JSON)

---

#### End-to-End Project Workflow

##### 1 Data Collection

- Airbnb listings dataset containing information such as:
    - Listing ID
    - Host details
    - Location (city, neighborhood)
    - Price
    - Room type
    - Availability
    - Reviews & ratings
- 

##### 2 Data Cleaning & Preprocessing (Python)

Using **Python (Pandas & NumPy)**:

- Removed duplicate listings
- Handled missing values (price, reviews, availability)

- Converted price fields from string to numeric format
- Standardized categorical columns (room type, neighborhood)
- Performed basic feature engineering such as:
  - Price per night
  - Availability classification
  - Host activity indicators

📌 **Outcome:** Clean, analysis-ready dataset

---

### 3 Data Storage & Management (MySQL)

- Created structured tables in **MySQL**
- Defined appropriate data types and constraints
- Loaded cleaned data from Python into MySQL
- Performed SQL queries for:
  - Average price by location
  - Room type distribution
  - Top performing hosts
  - Availability trends
  - Review-based performance analysis

📌 **Outcome:** Optimized relational database for analytics

---

### 4 Data Analysis (SQL + Python)

- Used **SQL** for aggregations, filtering, and joins
- Used **Python** for exploratory data analysis (EDA):
  - Price distribution
  - Room type vs price
  - Availability patterns
  - Host listing counts







📌 **Outcome:** Business-ready insights extracted from raw data

---

### 5 Data Visualization (Power BI)

Built an **interactive Power BI dashboard** connected to MySQL.

#### Dashboard Features:

-  Average Price by Neighborhood
-  Room Type Distribution
-  Availability Trends
-  Reviews vs Pricing
-  Top Hosts by Listings
-  Location-wise Insights

#### Power BI Enhancements:

- Slicers for city, room type, and price range
- KPI cards for total listings, avg price, and availability
- Interactive drill-downs

✦ **Outcome:** Actionable, decision-ready dashboard

---

#### Key Insights Derived

- Certain neighborhoods consistently command higher prices
  - Entire homes are priced significantly higher than private rooms
  - High availability does not always correlate with higher pricing
  - Hosts with multiple listings dominate high-revenue segments
  - Listings with better reviews tend to have optimized pricing
- 

#### Business Impact

- Helps hosts optimize pricing strategies
  - Enables investors to identify profitable locations
  - Assists travelers in understanding market trends
  - Demonstrates real-world data analytics pipeline skills
- 

#### Skills Demonstrated

- ✓ Data Cleaning & Transformation
- ✓ SQL Query Optimization
- ✓ Relational Database Design

- ✓ Exploratory Data Analysis
- ✓ Power BI Dashboard Design
- ✓ End-to-End Data Analytics Workflow