

Airbnb Dynamic Pricing Recommendation Engine

Introduction: Setting the right price is crucial for success on platforms like Airbnb, where competition and customer expectations vary across locations and seasons. This project focuses on building a dynamic pricing recommendation engine that leverages historical Airbnb data to predict optimal listing prices. Using Python for data preprocessing and machine learning, along with Power BI and Excel for visualization and analysis, the project identifies key factors influencing pricing—such as location, room type, availability, and reviews. The goal is to help hosts make informed, data-driven pricing decisions that maximize both occupancy and revenue.

Abstract: This project builds a dynamic pricing engine for Airbnb using Python, Power BI, and Excel. It analyzes historical listing data to identify key pricing factors such as location, seasonality, and room type. A Random Forest regression model predicts optimal prices, with listings flagged as underpriced, overpriced, or fair. Insights are visualized in dashboards, enabling data-driven pricing decisions for hosts.

Tools Used: Excel, Python, PowerBI

Steps Involved:

1. Data Cleaning in Excel

- Removed irrelevant columns and obvious errors
- Handled missing values and standardized formats manually
- Saved the cleaned dataset for further analysis

2. Data Preprocessing in Python

- Converted data types
- Engineered new features
- Applied one-hot encoding to categorical variables
- Split the data into training and test sets
- Trained a Random Forest regression model to predict prices
- Flagged listings as **Overpriced**, **Underpriced**, or **Fair** based on prediction error
- Saved results to CSV files for visualization

3. Data Visualization in Power BI

- Imported model outputs and raw data
- Created interactive visuals:
 - **Map:** Shows listing distribution by neighbourhood group.
 - **Scatter Plot:** Compares actual vs predicted prices using Random Forest.
 - **Donut Chart:** Visualizes total price share by neighbourhood group.

- **Bar Chart:** Avg price by room type across neighborhoods and top hosts by total price generated.
- **KPI Cards:**
 - Total Number of Hosts-12.21K
 - Total Price Listing- 3M
- **Filters:** Interactive slicers by room type.

Key Insights:

- **Entire homes** have the highest price listing of 2M and 8682 hosts.
- **Eixample** has the highest number of hosts(1690) and has the highest total listing price of 1.17M (43.74%). It indicates a large number of active listings and/or higher pricing per night in this neighborhood.
- The host **Acomodis Apartments** has the highest total price listing of 177k.

Conclusion:

This project demonstrates a data-driven approach to optimizing Airbnb pricing using historical listing data. By cleaning and preparing the dataset in Excel and Python, applying regression models, and visualizing insights through Power BI, we identified key pricing predictors such as room type, neighborhood group, and availability. Neighborhoods like **Eixample** emerged with the highest total prices, reflecting strong market activity. However, pricing alone doesn't imply profitability, as it excludes factors like occupancy and operational costs. The final dashboard equips hosts and analysts with actionable insights to make smarter, location-specific pricing decisions.