

Airbnb Dynamic Pricing Recommendation Engine

Introduction

Short-term rental markets like Airbnb experience frequent fluctuations in demand due to seasonality, guest behaviour, and listing quality. Many hosts struggle to set optimal nightly rates. This project uses machine learning to build a dynamic pricing recommendation engine that predicts ideal listing prices based on data from major U.S. cities.

Project Objective

The goal is to develop a machine learning model that predicts the log-transformed nightly price of an Airbnb listing based on its features (location, room type, reviews, amenities, etc.), helping hosts improve occupancy and revenue.

Tools Used

- Python (Pandas, Scikit-learn, XGBoost, Seaborn)
- Google Colab
- Excel, Tableau or Power BI (dashboard design)
- Microsoft Word (reporting)

Dataset Overview

~43,000 listings with 29 columns including:

- log_price (target), city, room_type, accommodates
- review_scores_rating, number_of_reviews, amenities
- host_identity_verified, availability, bedrooms, beds

Methodology

- ✓ Data Cleaning – Handled nulls, removed incomplete rows
- ✓ EDA – Analyzed trends across cities, review scores, amenities
- ✓ Feature Engineering – Encoded categorical data, transformed text-based columns
- ✓ Modeling – Trained Linear Regression, Random Forest, and XGBoost
- ✓ Evaluation – Used MAE, RMSE, and R^2 to assess performance

Results & Insights

XGBoost delivered the best predictive performance with $R^2 = 0.71$.

Key factors influencing price:

- Room Type: Entire homes priced ~40–70% higher than shared/private rooms
- City: Listings in NY and SF showed significantly higher base prices
- Review Scores: Strong correlation between high scores and pricing power
- Amenities: Listings with Wi-Fi, kitchen, and AC priced more competitively

Model Performance:

Model	MAE	RMSE	R ² Score
XGBoost Regressor	0.2804	0.3869	0.7086
Random Forest	0.2819	0.3915	0.7016
Linear Regression	0.3597	0.4830	0.5458

Recommendations

- Base pricing should vary by city and neighborhood demand
- Boost price by 10–20% for listings with `review_scores_rating > 90`
- Highlight high-demand amenities to justify premium pricing
- Use calendar availability and seasonal patterns to adjust rates
- Retrain the model periodically to reflect current trends
- Integrate predictions into a dashboard for real-time suggestions

Conclusion

This project demonstrates how machine learning can support Airbnb hosts with competitive, data-driven pricing. By leveraging key listing attributes, the model recommends nightly rates that reflect real market dynamics. With dashboard integration and automated retraining, this engine can serve as a practical pricing tool for short-term rentals.