

AIRBNB BOSTON DATA ANALYSIS USING AI

INTRODUCTION

- Problem Statement: The Airbnb market is highly competitive, and hosts need actionable insights to optimize pricing, improve guest satisfaction, and increase booking rates. Predicting these factors can help hosts make data-driven decisions to stay competitive in a dynamic market.
- Objective: This project aims to apply AI and machine learning techniques to analyze the Boston Airbnb market, focusing on three core challenges:
 1. Price Forecasting: Predicting listing prices based on property features.
 2. Guest Rating Prediction: Identifying factors contributing to high guest ratings.
 3. Booking Likelihood: Predicting locations with high booking probability and identifying common characteristics.
- Approach: Using models like Linear Regression for price prediction, Random Forest for guest ratings, and classifiers like SVM, Random Forest, and KNN for booking likelihood, the analysis integrates both numerical and categorical features, along with interactive geospatial tools, to provide comprehensive insights.

Forecasting Price Using AI

- Used Linear Regression to predict Airbnb listing prices based on features like property size, amenities, and ratings, achieving strong predictive performance with MAE: 66.25, RMSE: 151.69, and R²: 0.1961.



OPTIMAL
PRICING
STRATEGY

PREDICTING REVIEW RATINGS

Can we predict if an Airbnb listing will receive a high rating score & What features contribute the most to it.

- Used Random Forest to predict whether an Airbnb listing will receive a high rating, achieving Accuracy: 95.68%, Precision: 0.9592, Recall: 0.9971, and F1 Score: 0.9777.
- Feature Importance showed key contributing features included price, cleaning fee, security fee, and bathrooms.

FACTORS INFLUENCING BOOKINGS

What are the locations with the highest booking probability & What do these locations have in common?

- Tested SVM, Random Forest, and KNN classifiers, with Random Forest achieving the highest Accuracy: 99.58%. Key insights:
- Location: Roslindale neighborhood had the highest booking likelihood.
- Room Types: Most high-booking listings were Entire Home/Apt (Mean Price: \$209.22), followed by Private Room (\$96.99) and Shared Room (\$90.28).
- An interactive geospatial map was created to highlight high-demand listings.

CONCLUSION

- AI-driven Predictions: The project successfully leveraged AI and machine learning techniques, with Linear Regression and Random Forest models providing strong predictive capabilities for price forecasting and guest rating predictions.
- Insights for Hosts: The analysis revealed crucial insights, including the significance of price, cleaning fee, security fee, and bathrooms in predicting high ratings, as well as the Roslindale neighborhood and Entire Home/Apt listings being most likely to achieve high booking rates.
- Actionable Recommendations: Hosts can optimize pricing, improve amenities, and target high-demand neighborhoods to enhance guest satisfaction and booking probability.

Future Work: Future efforts could explore real-time prediction models, sentiment analysis from reviews, and incorporating external factors (e.g., events or local trends) to improve predictions and market competitiveness.