Introduction

Airbnb pricing is highly dynamic and influenced by various factors such as location, seasonality, availability, and guest reviews. The goal of this project is to analyze historical Airbnb listing data from Amsterdam and build a price recommendation engine to help hosts set competitive prices for their properties. The model predicts the optimal price based on room type, season, reviews, and neighbourhood characteristics.

Abstract

This project enables data-driven pricing for Airbnb listings using Python and Tableau. It involves analyzing key pricing factors, building regression models, and creating a price recommendation engine. The final output includes a trained model, an interactive dashboard, and a tool that suggests personalized optimal prices.

Tools Used

- Python (Pandas, Scikit-learn, Seaborn, Matplotlib)
- Tableau Public (Data Visualization)
- Google Colab (Coding Environment)
- Excel (Basic Data Checks)
- Joblib (Model Saving & Loading)

Steps involved in Building the Project

1. Data Preprocessing

- Loaded Airbnb Amsterdam data
- Cleaned missing & irrelevant fields
- One-hot encoded categorical features like room type, season, and neighbourhood

2. Exploratory Data Analysis (EDA)

- Boxplots, histograms, and heatmaps to understand relationships
- Identified trends like higher prices in Centrum-West and during summer

3. Feature Engineering

- Extracted season from review date
- Calculated review_frequency
- Reconstructed room_type and neighbourhood for interpretability

4. Modeling Phase

- <u>Target:</u> price, Features: room type, availability, reviews, etc.
- Trained multiple models: Linear Regression, Decision Tree, Random Forest & XG Boost
- Evaluated with MAE, RMSE, R^2
- Selected Linear Regression as the best-performing model

5. Recommendation Engine

- Built recommend_price() function in Python
- Accepts inputs (room type, season, reviews, availability, neighbourhood)
- Returns predicted price using trained model

6. Tableau Dashboard

- Loaded cleaned CSV
- <u>Filters:</u> Room Type, Season, Neighbourhood
- <u>Charts:</u> Price Distribution, Reviews vs Price
- Sliders for availability and reviews_per_month to explore pricing patterns

Conclusion

This project successfully delivered a complete dynamic pricing recommendation system for Airbnb listings using machine learning and data visualization. The pricing engine offers personalized price predictions based on user inputs, while the dashboard provides hosts with insights on market trends. With minimal input, a host can determine a competitive listing price-helping improve bookings and revenue.

Suggestions

Based on data analysis and model outcomes, the following pricing recommendations are proposed:

Room Type Pricing: Hotel rooms yield the highest average prices. Private and shared rooms should be priced more competitively to attract budget-conscious travelers.

Neighbourhood Strategy: Premium pricing is suitable for areas like De Baarsjes – Oud-West, Centrum-West, and De Pijp – Rivierenbuurt. Lower-demand areas should focus on affordability and value offerings.

Seasonal Pricing: Spring shows higher price trends, followed by Summer. Apply dynamic pricing with discounts during Winter to boost bookings.