Airbnb Data Analysis (PYTHON)

Project Objective

The objective of this project is to analyze Airbnb listing data using Python to uncover meaningful insights related to pricing, availability, property types, and location-based trends. By applying data cleaning, exploratory data analysis (EDA), and data visualization techniques, the project aims to support data-driven decision-making for hosts, customers, and business stakeholders.

Problem Statements

The project focuses on answering the following key business and analytical questions:

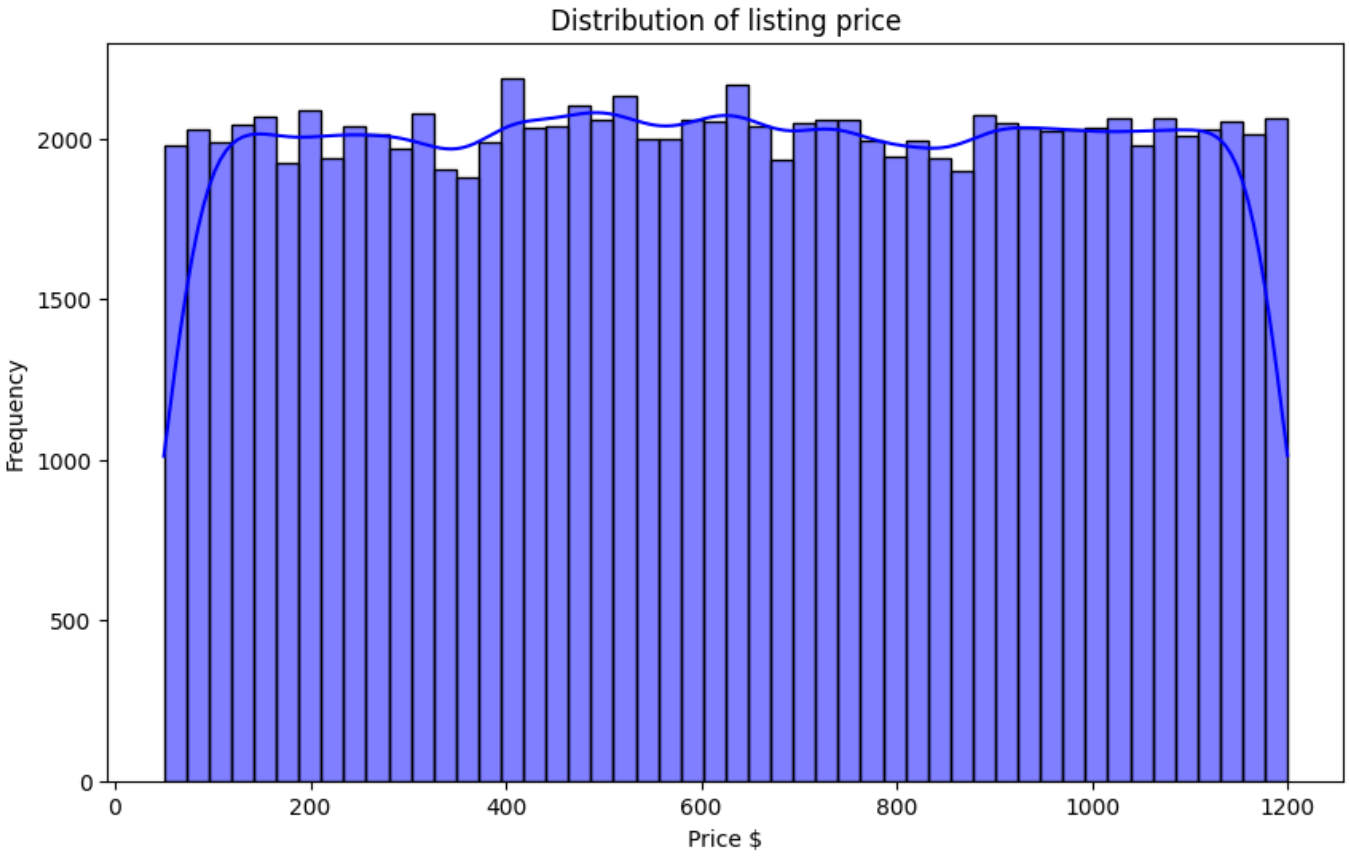
- What factors influence Airbnb listing prices?
- How does availability vary across different neighborhoods and property types?
- Which room types and locations are most popular among customers?
- Are there noticeable trends or patterns in pricing and reviews?
- How can hosts optimize pricing and availability based on data insights?

Data Visualization & Analysis

Various Python libraries such as Pandas, Matplotlib, and Seaborn were used to visualize and analyze the dataset.

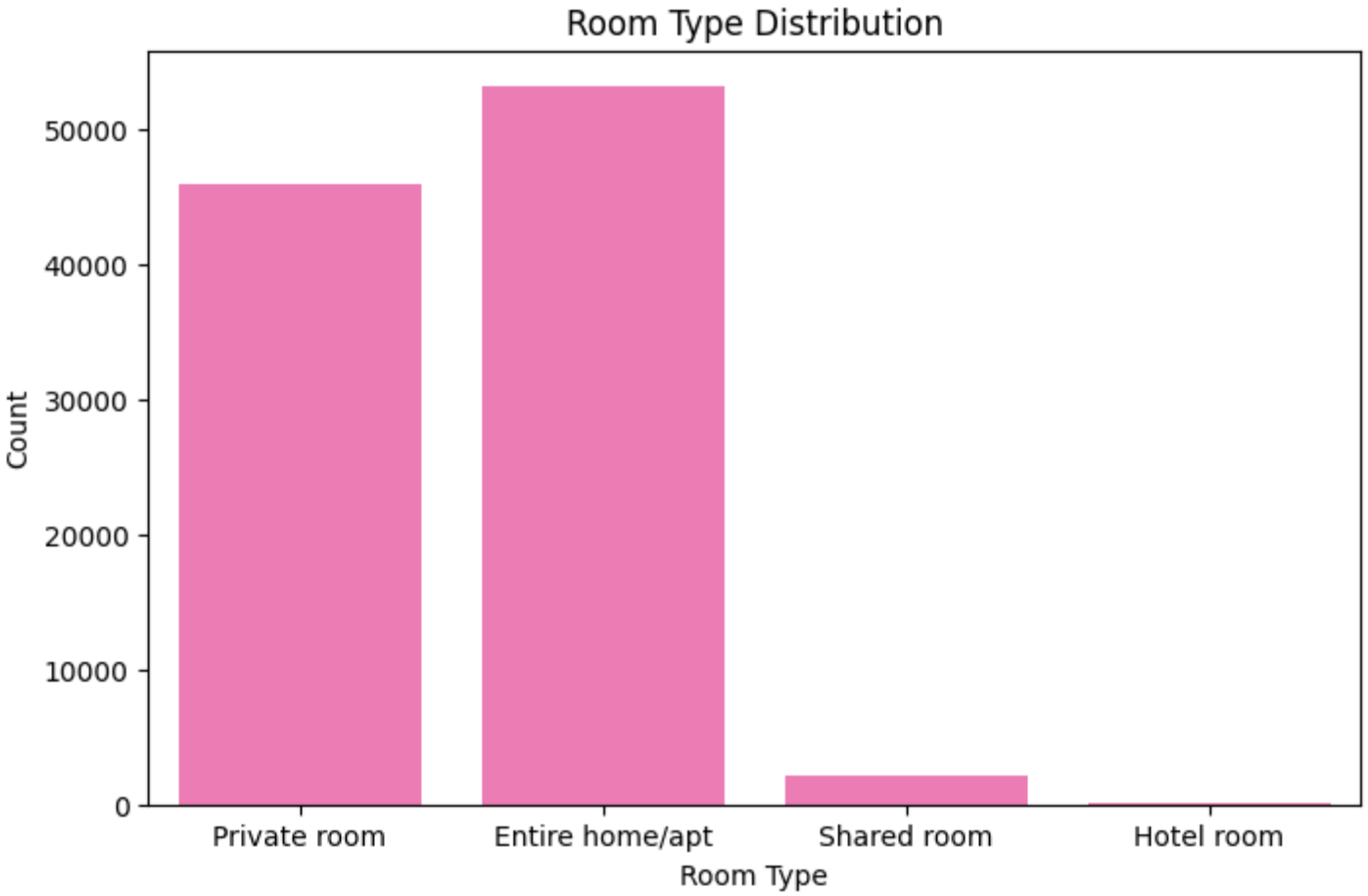
Key Visualizations Created:

Price Distribution Analysis



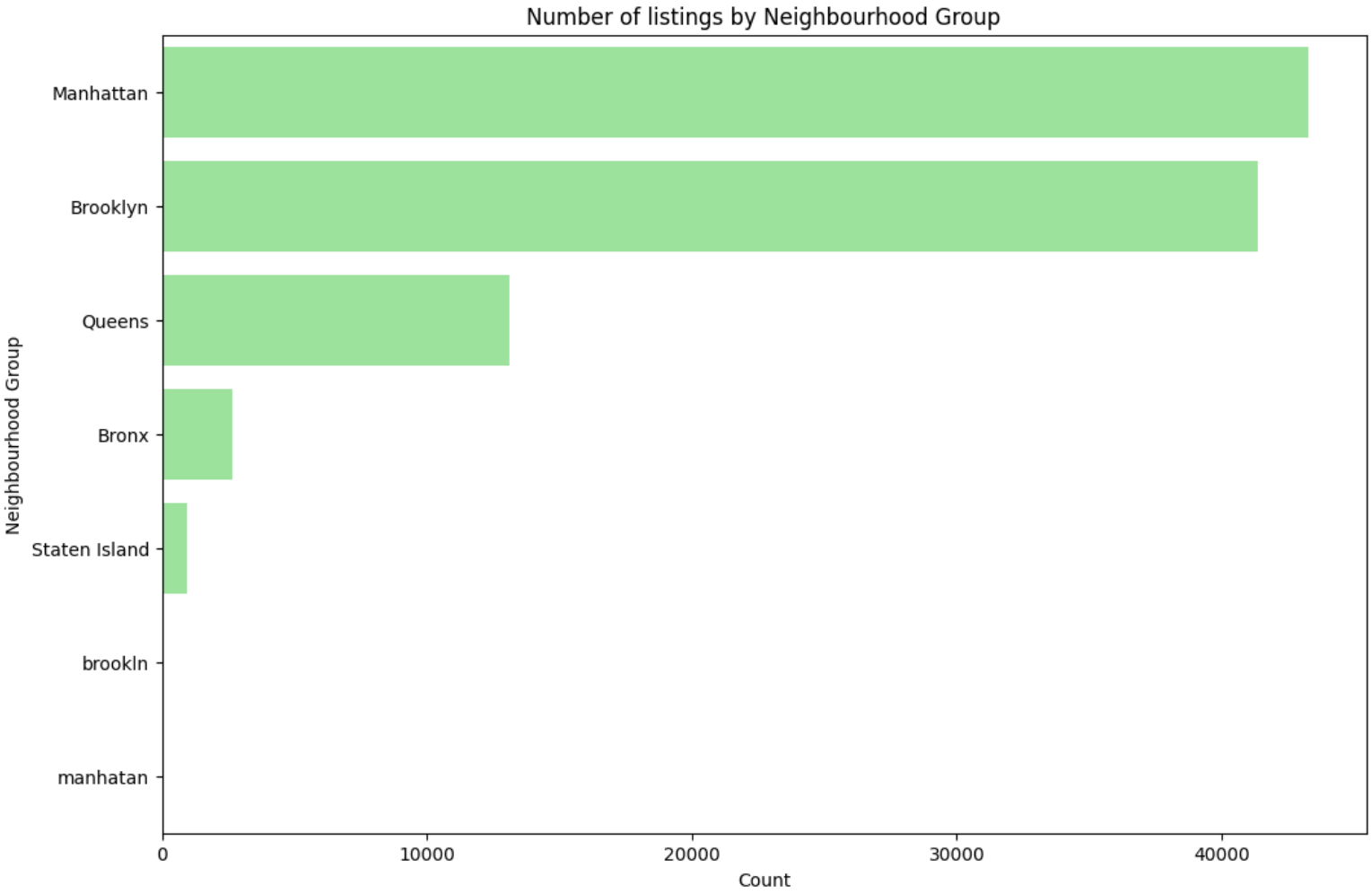
- Histograms and box plots to understand pricing spread and detect outliers.

Room Type Analysis



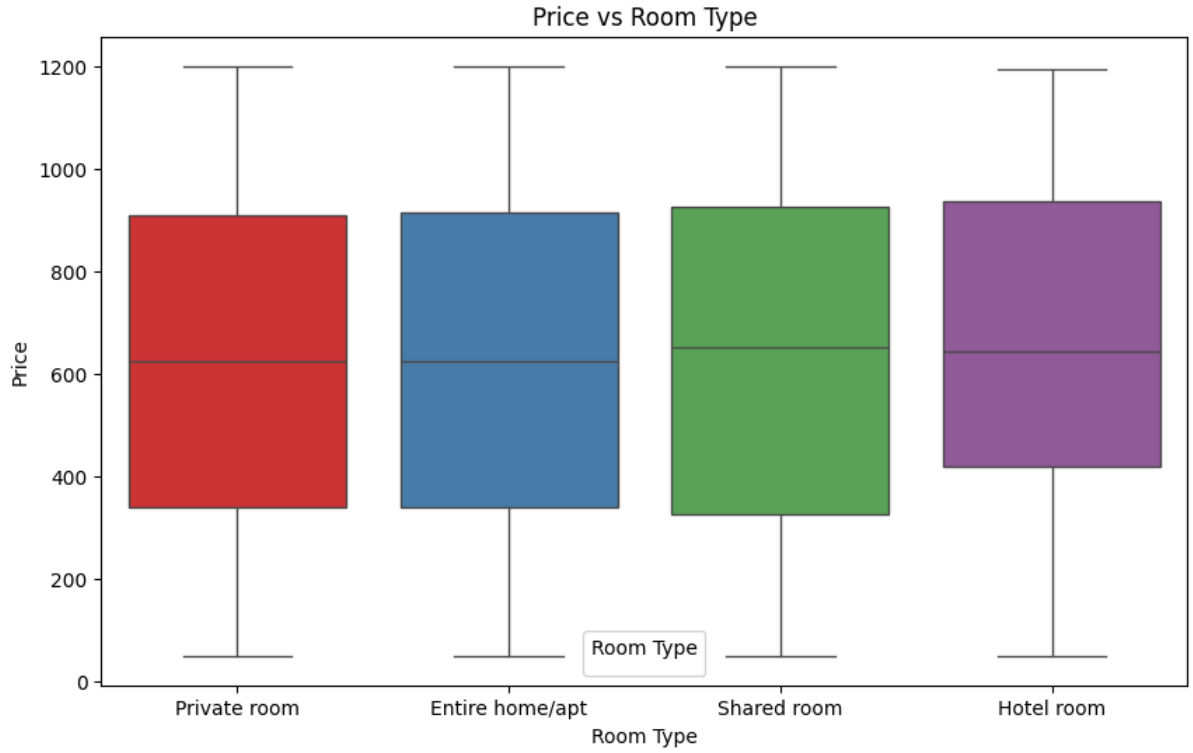
- Bar charts showing the count and average price of different room types.

Neighborhood-Based Analysis



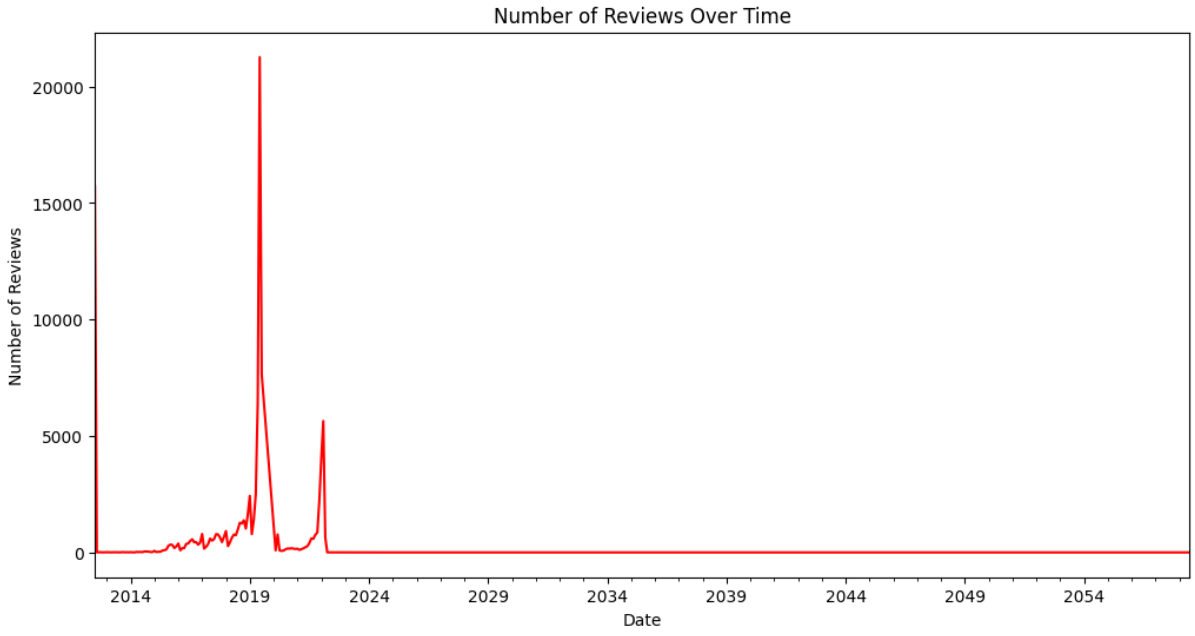
- Visual comparison of listings and prices across neighborhoods.

Availability Analysis



- Line plots and bar charts showing availability trends.

Correlation Heatmap



- To identify relationships between price, reviews, availability, and minimum nights.

Insights Gained:

- Entire homes/apartments are generally priced higher than private or shared rooms.
- Listings in prime locations have higher average prices.
- High availability does not always correlate with higher prices.
- A moderate correlation exists between price and minimum nights.

Conclusions

- Location and room type are the most significant factors affecting Airbnb prices.
- Budget-friendly options are mostly private and shared rooms.
- Listings with competitive pricing tend to have higher availability.
- Outliers in pricing can significantly impact average price calculations, making data cleaning essential.

Final Summary

This project demonstrates the effective use of Python for data analysis and visualization. By transforming raw Airbnb data into actionable insights, the analysis highlights how data-driven strategies can improve pricing decisions and customer satisfaction. The project also strengthens practical skills in data preprocessing, EDA, and visualization, making it highly relevant for real-world data analyst roles.

Tools & Technologies Used

- Python
- Pandas
- NumPy
- Matplotlib
- Seaborn
- Jupyter Notebook