Airbnb Data Analysis Report

Prepared in collaboration with Elevate Labs

1. Introduction

In today's fast-evolving rental market, understanding what drives property performance and guest satisfaction is more important than ever. This project set out to explore the dynamics of Airbnb listings—digging into what makes properties successful, how pricing strategies can be improved, and which factors boost customer reviews. The analysis drew on tools like Excel, Power BI, and Python to deliver both deep insights and practical recommendations.

2. Dataset Overview

The data for this study comes from Airbnb's publicly available listing records on Kaggle, with thousands of entries covering various cities and property types. Each listing shares details such as:

The kind of property and room being offered

How many guests it can accommodate

Bedroom and bathroom counts

Pricing and cleaning fees

Guest reviews plus overall ratings

City location and availability info

Having this rich dataset allowed for a comprehensive look at trends across geographies and property categories.

3. Data Cleaning & Transformation

To ensure that insights were trustworthy, the dataset first went through a thorough cleaning stage in Excel. This involved:

Removing duplicate entries and filling in or eliminating missing values

Converting log-transformed price data back to actual figures

Standardizing property and room type labels for consistency

Identifying and treating outliers, especially in price points and review counts

These steps ensured that the following analysis reflected the true patterns in the data, rather than quirks or inconsistencies.

4. Power BI Dashboard

A key part of this project was turning numbers into stories through interactive visuals and summary metrics in Power BI. Some dashboard highlights include:

KPIs–see at a glance how many listings there are, what the average price is, how much total revenue is generated, and how guests rate their experiences

Bar and column charts depicting which cities have the most listings and the price differences among room types

Donut charts showing the share of each property type

Trend lines linking price to the number of guests accommodated

Filters and slicers—make it easy to explore specific cities, room types, ranges of prices, or guest capacities

This dashboard lets users instantly pinpoint patterns and drill into details that matter to them.

5. Predictive Modeling (Python)

To forecast pricing and identify what drives listing value, a linear regression model was built using Python and Scikit-learn. By feeding in features like property type, room type, capacity, reviews, ratings, and city, the model predicts what price a listing is likely to command. Model performance will be evaluated by:

Root Mean Squared Error (RMSE)—measuring average prediction error

R^2 Score—how well key features explain price variation

6. Key Insights

What emerged from the analysis?

Whole homes and apartments form the bulk of Airbnb's market share.

Prices do rise with guest capacity—but not in a predictable linear way.

A handful of cities contribute most to overall revenue, showing strong local demand.

Properties with higher ratings reliably command better prices, suggesting hosts' effort in service quality pays off.

7. Tools Used

This end-to-end process leveraged:

Excel for data cleaning and transformation

Power BI for dashboards and KPI visualizations

Python for predictive modeling and deeper statistical analysis

Kaggle's Airbnb dataset as the foundational data source

8. Conclusion

Working on this project demonstrated how systematic analysis can turn raw data into practical advice for hosts, managers, and property investors. By cleaning, visualizing, and modeling the data, it's possible to identify factors that matter most for pricing, bookings, and guest satisfaction. These insights have the power to help stakeholders optimize their offerings and maximize returns as the Airbnb ecosystem continues to grow.