TITLE : AIRBNB NYC CASE STUDY

subtitle : Data + Business Analysis using Power BI & Python

Presented By:
Vaibhav Negi
Atul Kumar

PROBLEM STATEMENT

- Airbnb struggles with maintaining consistent profitability across major cities.
- Host churn and guest dissatisfaction have emerged due to inconsistent experiences.
- Listing quality and engagement levels vary significantly by neighborhood.
- Lack of actionable insight into underperforming regions and room types.
- Need to identify key operational inefficiencies affecting revenue and satisfaction.
- **Aim:** Use NYC listing-level data to uncover issues and recommend strategic improvements.

Objectives

- Analyze NYC Airbnb listings to uncover pricing trends, availability patterns, and guest engagement behavior.
- Identify the top-performing room types and neighborhoods based on average price and review volume.
- Detect underperforming listings using availability, review activity, and low rating indicators.
- Study host behavior and scale by examining the number of listings per host.
- Explore guest activity through reviews per month and total reviews as engagement signals.
- Provide data-driven recommendations to improve pricing strategy, occupancy, and platform performance.

Dataset Overview

- **Source: Airbnb NYC Open Data (from Kaggle / public repository)**
- **Format: CSV (Cleaned using Python before Power BI import)**
- **Total Columns:** 23 **|** **Total Rows:** 9992
- **Key Features Analyzed:**
  room_type, price, availability_365, review_scores_rating
  neighbourhood_group, host_id, instant_bookable, reviews_per_month
- **Calculated Fields:**
  total_cost = price + service fee
  occupancy_rate, revenue_potential, review activity categories
- **Cleaning Done:**
  Null handling, type conversion, filtering out price/availability outliers

Tools & Methodology

Tools Used:

- **Python (Jupyter Notebook) –** for data cleaning, preprocessing & visualizations
- **Power BI –** for building interactive dashboards
- **Excel / CSV –** as the data source file
- **GitHub –** for version control and portfolio hosting
- **MS Word / Google Docs –** for business documentation (BRD, DRD, SRS, FRD)

Methodology Followed:

1. Data Understanding
Explored the dataset structure and feature meanings

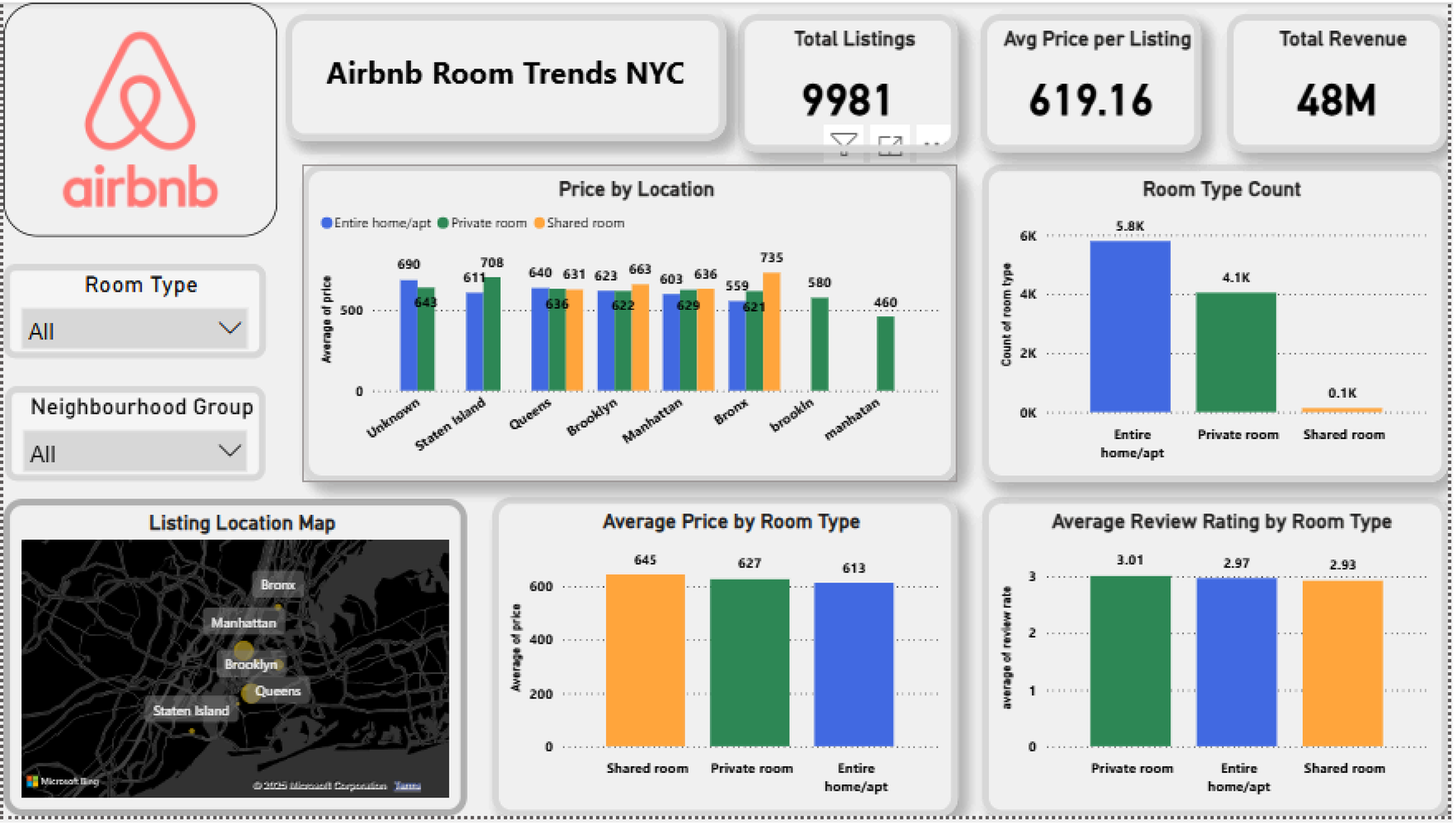
2. Data Cleaning (Python)
Handled missing values, removed duplicates, converted types

3. Exploratory Data Analysis
Used charts in Python to explore patterns and correlations

4. Dashboard Creation (Power BI)
Built 2 dashboards showing room trends, revenue, host activity

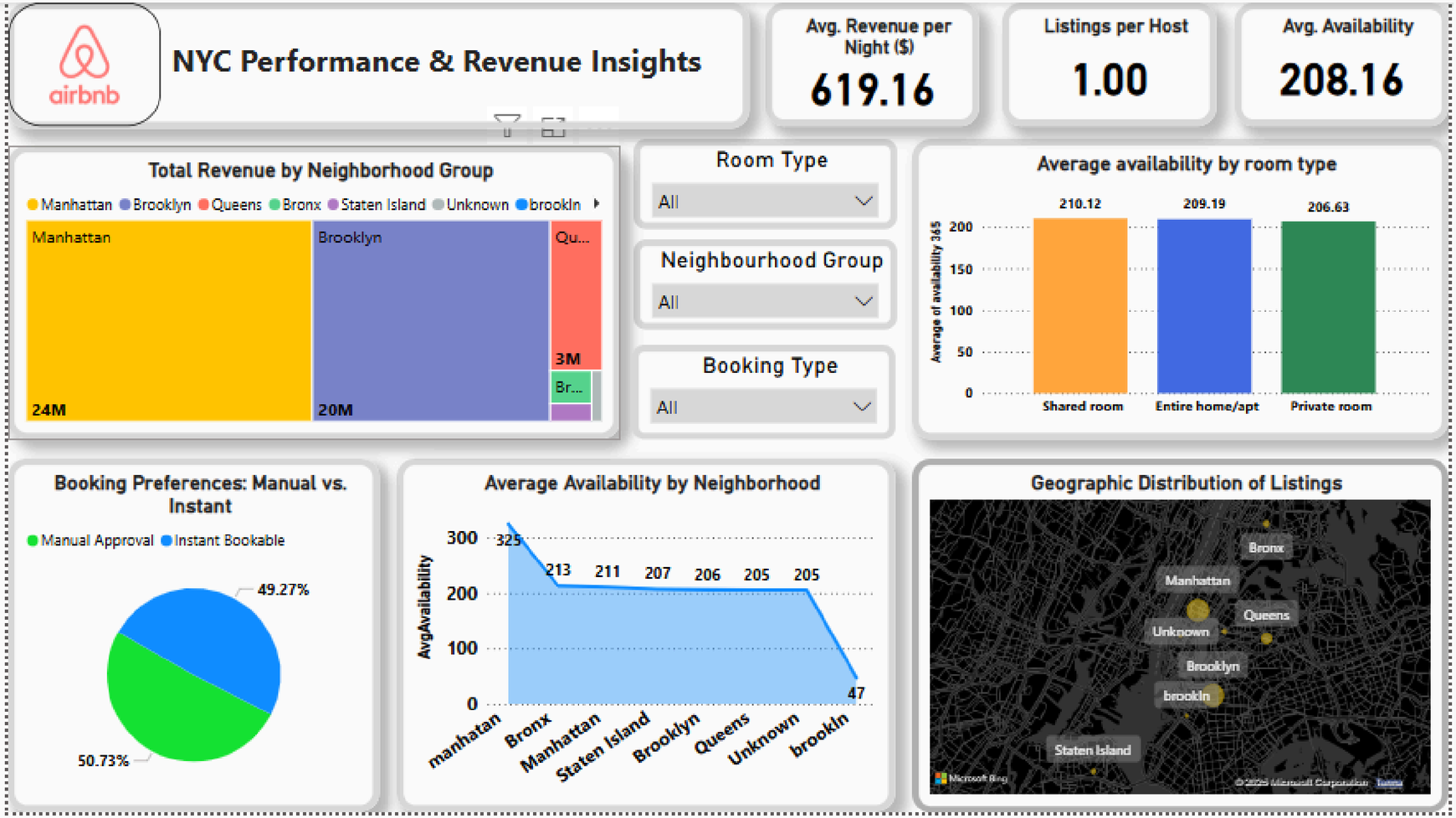
5. Documentation & Reporting
Created business-focused docs to support insights and recommendations

Power BI Dashboard 1



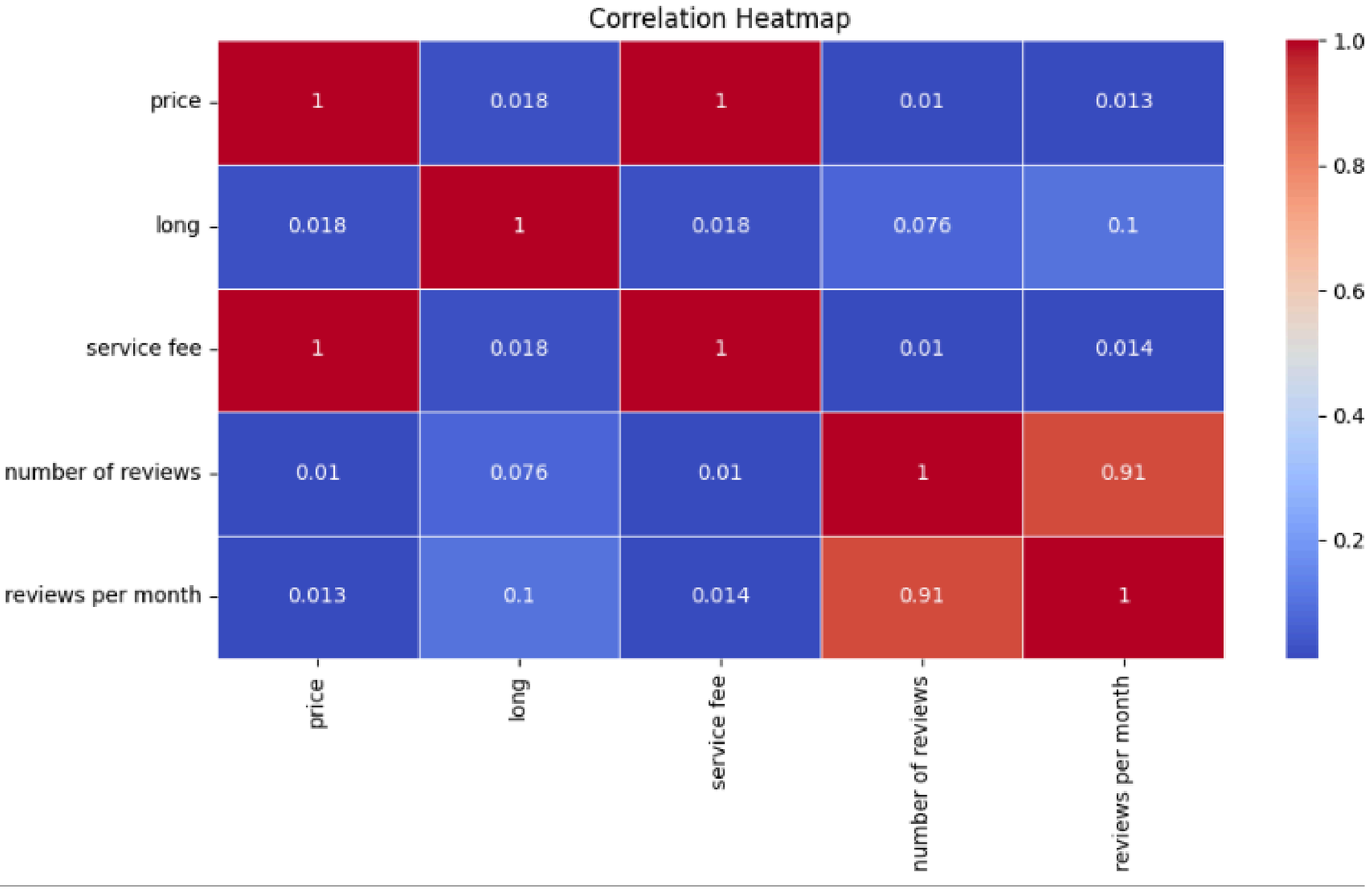
- Entire home/apt listings are priced the highest on average, making them a major source of revenue.
- Private rooms dominate the listing count, especially in Brooklyn and Queens.
- Revenue potential is highest for entire homes, while shared rooms contribute the least.
- Bubble map reveals listing clusters in Manhattan and Brooklyn, the densest Airbnb zones.
- Average review ratings are fairly consistent across room types, indicating uniform guest satisfaction.

Power BI Dashboard 2



- Manhattan and Brooklyn generate the highest total revenue, indicating strong demand and higher pricing.
- Average availability is consistent across room types, around 206–210 days per year.
- Listing density is highest in central NYC areas — especially Manhattan, Brooklyn, and Queens.
- Manual approval and instant bookings are nearly equal, showing a balance in host control vs. convenience.
- Staten Island and ‘Unknown’ neighborhoods have low availability, suggesting lower listing activity or inactive hosts.

Python Analysis Highlights



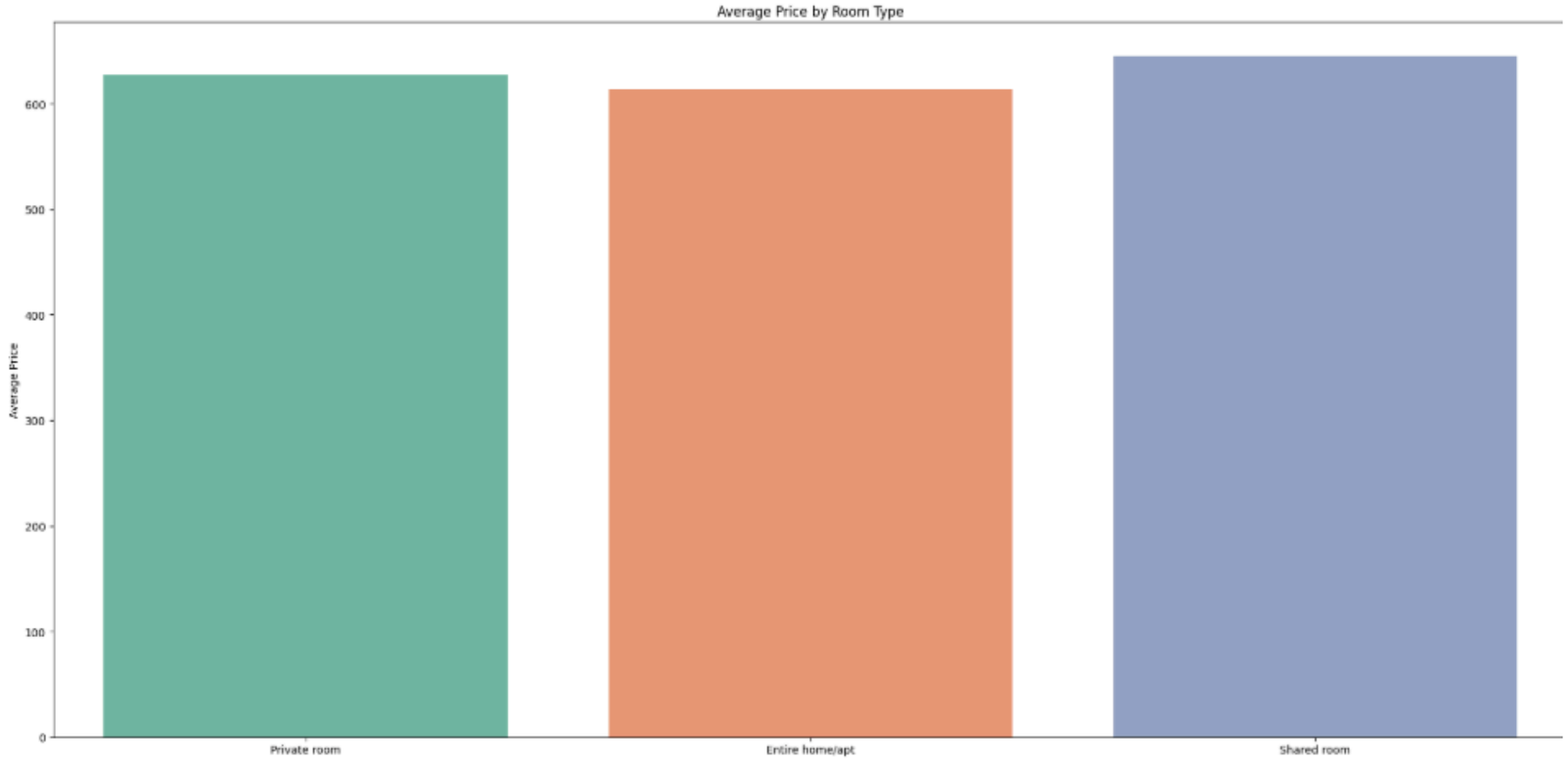
1. Correlation Heatmap

- Strong positive correlation (0.91) between reviews per month and number of reviews, indicating consistent guest engagement.
- Price has little correlation with review activity, meaning high cost doesn't guarantee popularity.
- Service fee and price are nearly identical, suggesting fees are price-derived.

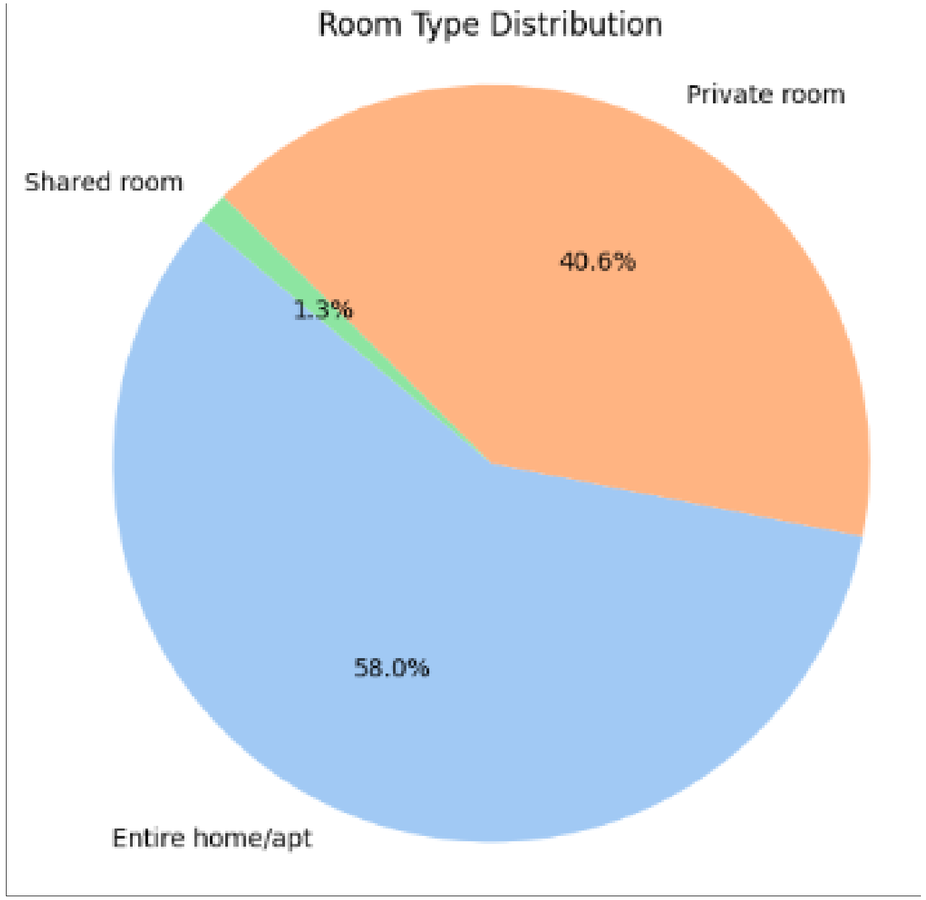
Python Analysis Highlights

2. Avg Price by Room Type

- **Entire home/apt listings have the highest average price, indicating premium positioning.**
- **Shared rooms are the cheapest, making them more accessible but less profitable.**
- **Private rooms fall in the middle, appealing to budget-conscious travelers.**



Python Analysis Highlights



3. Room Type Share (Pie Chart)

- **Private rooms dominate the NYC Airbnb market, showing high host participation and guest preference for affordable stays.**
- **Entire homes make up a smaller portion, likely due to higher cost and regulations.**
- **Shared rooms have minimal share, suggesting low demand and listing frequency.**

Business Insights

- **Manhattan and Brooklyn generate the highest revenue, making them ideal areas for expansion and premium listing focus.**
- **Private rooms dominate listing volume, but earn lower average revenue per listing than entire homes.**
- **Shared rooms show minimal presence and lowest pricing, indicating weak guest demand.**
- **Listings with frequent availability (high availability_365) tend to attract more reviews, showing stronger engagement.**
- **Areas like Staten Island and ‘Unknown’ regions show low activity, hinting at underutilized zones.**

Recommendations

- Increase focus on entire home listings in top-performing areas like Manhattan and Brooklyn for better ROI.
- Optimize pricing for private rooms to balance affordability with revenue generation.
- Promote Instant Bookable listings to streamline guest booking and increase conversions.
- Target inactive areas with host engagement campaigns, discounts, or platform support.
- Review low-performing listings for quality control or strategic improvement.

Conclusion

This project provided a deep, data-driven analysis of Airbnb listings in New York City using both Python and Power BI. Through exploration of pricing patterns, listing types, host activity, and availability, we uncovered actionable insights that highlight key opportunities for strategic growth. The analysis revealed that while private rooms dominate in volume, entire homes yield higher revenue, with Manhattan and Brooklyn emerging as the most profitable regions. By combining interactive dashboards with strong documentation and meaningful recommendations, this project demonstrates how raw data can be transformed into business strategy — bridging the gap between analytics and real-world decision-making.

THANKS FOR WATCHING

“ तमसो मा ज्योतिर्गमय ”

“From darkness, lead me to light.”

91+ 7895160343
vaibhavn2024@gmail.com

91+ 8603174522
atulsingh845101@gmail.com