

Business Requirements Document (BRD)

Project Title: Airbnb NYC Data Analysis Project

Prepared by: Vaibhav Neg

Date: 18-05-2025

1. Executive Summary

This BRD outlines the business context, goals, and analytical deliverables for a data analysis project based on Airbnb listings data in New York City. The objective is to uncover actionable insights related to pricing, availability, room types, and host performance. This project combines the responsibilities of a Business Analyst and Data Analyst, and is supported by Power BI dashboards and Python-based visualizations.

2. Project Objective

To enable data-driven decision-making for Airbnb stakeholders by analyzing New York City listings data. The analysis will support:

- Strategic pricing decisions
 - Room type performance evaluations
 - Host engagement metrics
 - Neighborhood-based revenue potential
-

3. Stakeholders

- Product Managers
- Data Analysts
- Business Intelligence Teams
- Airbnb Hosts
- City Planners

4. Scope of Work

In Scope:

- Cleaning and preprocessing NYC Airbnb dataset using Python
- Creating exploratory and statistical visualizations in Jupyter Notebook
- Building two interactive dashboards in Power BI
- Documenting findings in BRD, DRD, and SRS formats

Out of Scope:

- Predictive modeling or advanced ML algorithms
- Real-time data streaming

5. Business Requirements

1. Identify high-performing room types in terms of pricing and reviews.
2. Analyze neighborhood-level revenue contribution.
3. Evaluate host-level metrics like availability, number of reviews, and revenue.
4. Understand booking behavior preferences (manual vs. instant).

6. Data Sources

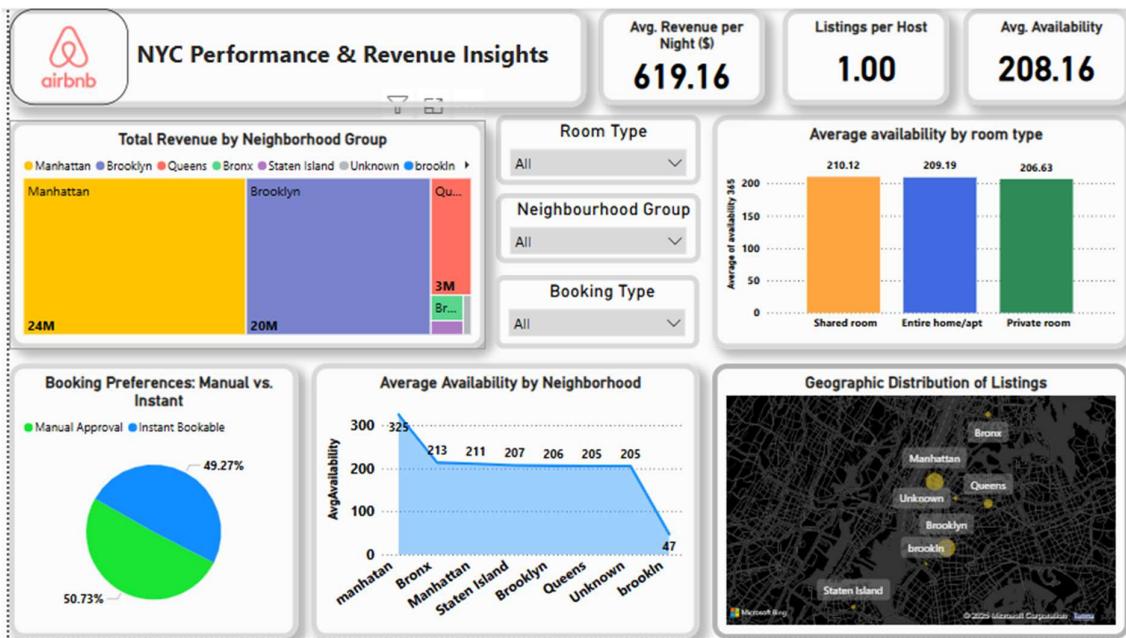
- Dataset: NYC Airbnb Open Data (CSV format)
- Source: Inside Airbnb public repository or Kaggle
- Cleaning & Preprocessing: Done in Python (Jupyter Notebook)

7. Visuals & Descriptions

Power BI Dashboards:

📍 Dashboard 1: Airbnb Room Trends NYC

- KPI Cards: Total Listings, Avg Price per Listing, Total Revenue
- Bar Chart: Price by Location
- Bar Chart: Average Price by Room Type
- Bar Chart: Room Type Count
- Bar Chart: Average Review Rating by Room Type
- Map: Listing Location Map



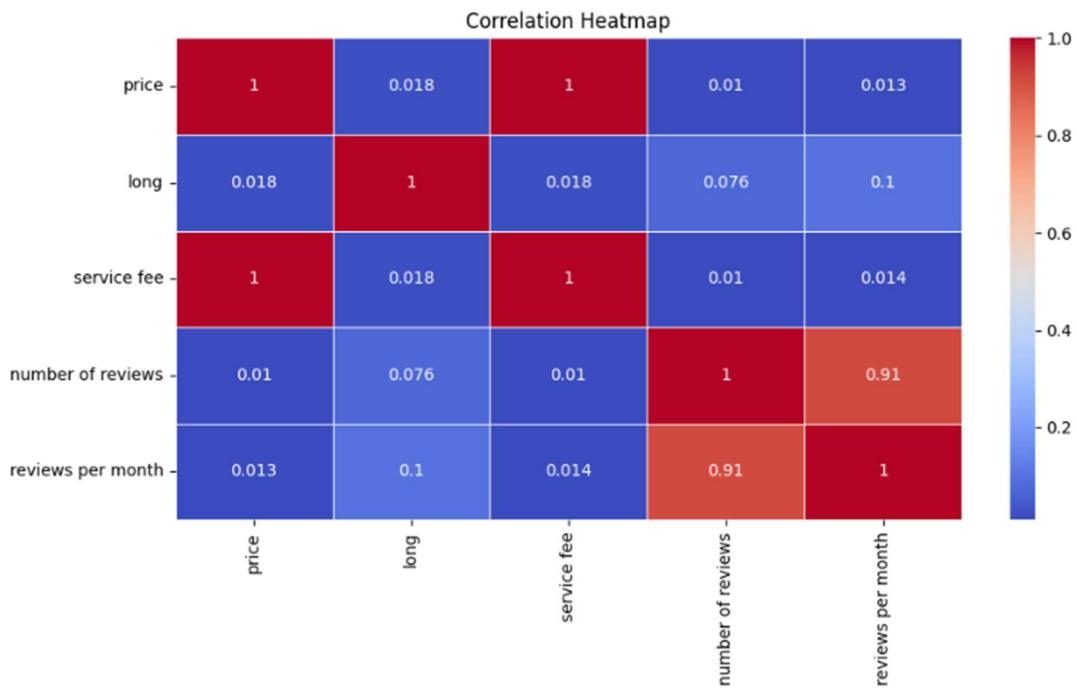
📍 Dashboard 2: NYC Performance & Revenue Insights

- KPI Cards: Revenue per Night, Listings per Host, Avg Availability
- Treemap: Total Revenue by Neighborhood Group
- Pie Chart: Booking Preferences: Manual vs. Instant
- Bar Chart: Average Availability by Room Type
- Line Chart: Average Availability by Neighborhood
- Map: Geographic Distribution of Listings

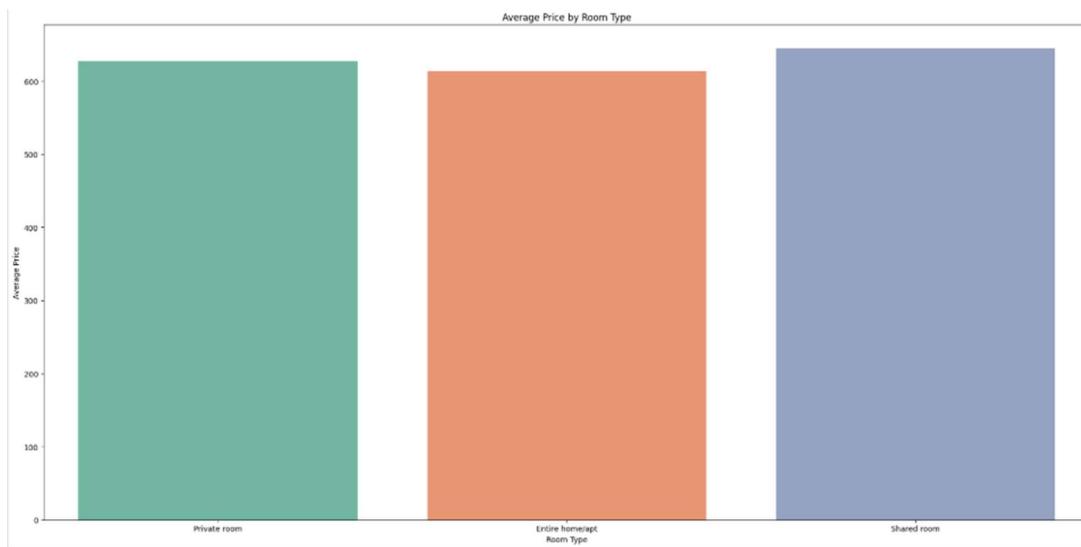


Python Visualizations (Jupyter Notebook):

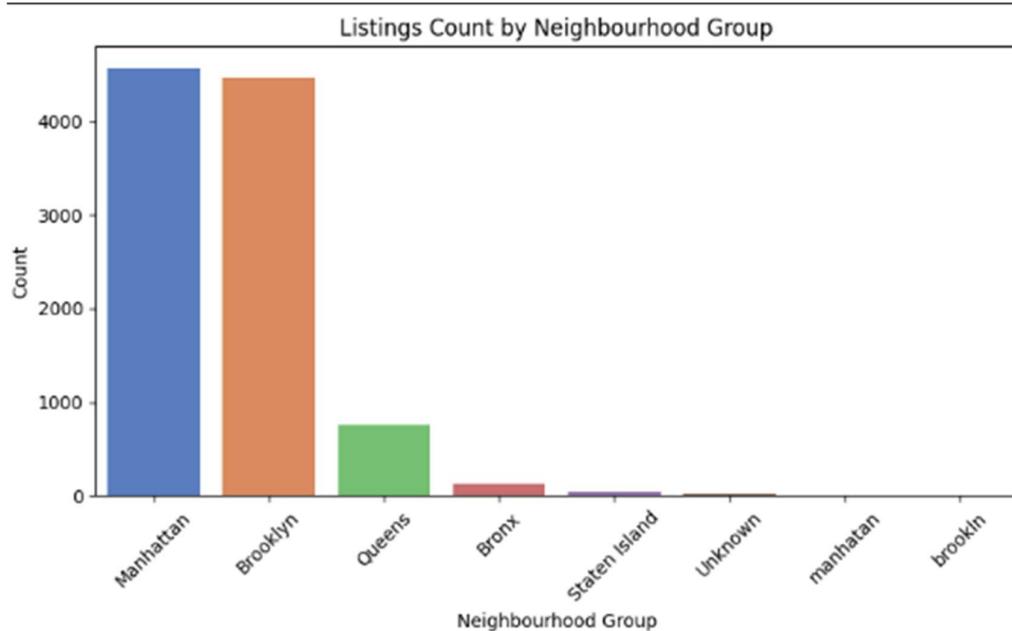
1. Confusion Matrix — Classification result evaluation.



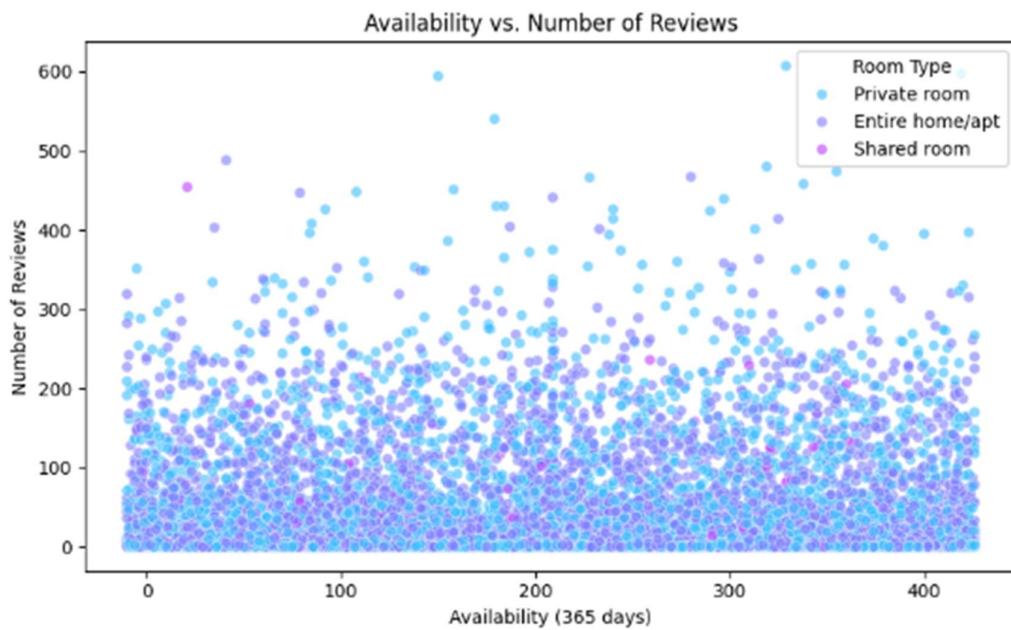
2. Bar Plot: Average Price by Room Type



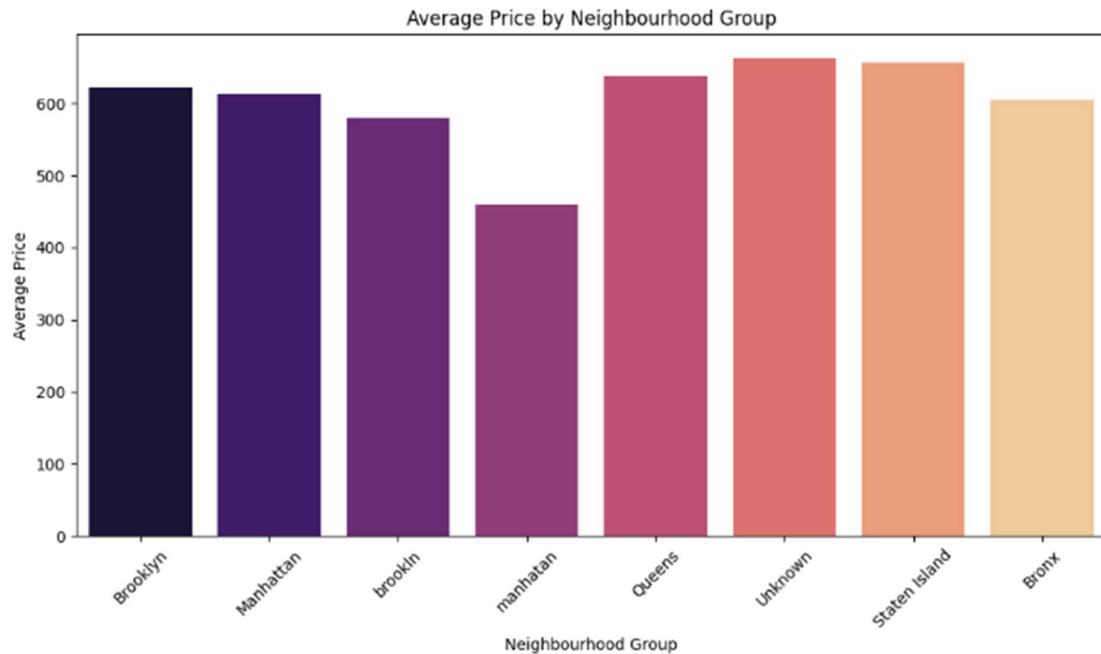
3. Count Plot: Listings Distribution by Neighborhood Group



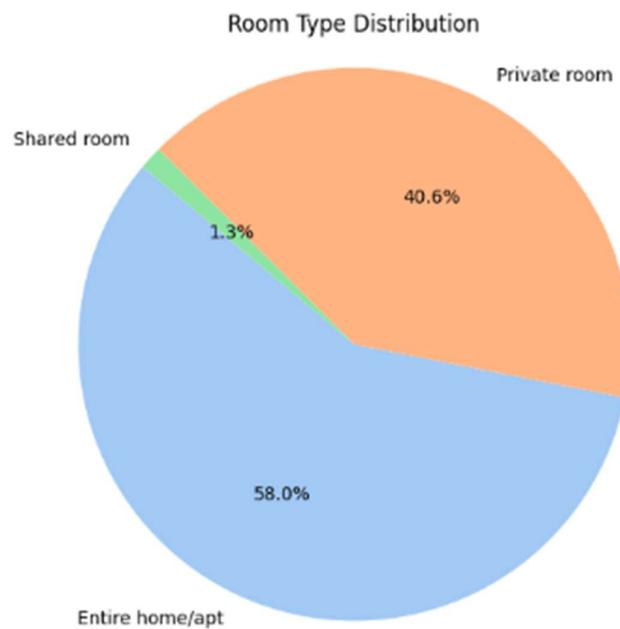
4. Scatter Plot: Availability vs. Number of Reviews



5. Bar Plot: Average Price by Neighborhood Group



6. Pie Chart: Room Type Distribution



8. Assumptions

- Dataset is representative of typical Airbnb operations in NYC.
 - Data cleaning was correctly performed.
 - Visualizations are assumed to be accurate at the time of analysis.
-

9. Constraints

- Limited to static data (no live updates)
 - Potential missing values in original dataset
-

10. Success Criteria

- Clear, interpretable visualizations
 - Accurate documentation (BRD, DRD, SRS)
 - Dashboards used to derive meaningful, actionable insights
-

11. Appendix

- Data dictionary (see DRD)
- Technical steps (see SRS)