

Software Requirements Specification (SRS)

Project Title: Airbnb NYC Case Study

Prepared by: Vaibhav Negi

Document Date: 18-05-2025

1. Introduction

1.1 Purpose

The purpose of this document is to define the software requirements for the Airbnb NYC case study project. It aims to provide data-driven insights into Airbnb listings across New York City for both business and data analytics purposes. The insights will support stakeholders in understanding market dynamics, listing performance, and host behaviour.

1.2 Scope

The project covers data analysis and dashboard development using Power BI and Python. Key tasks include:

- Data cleaning using Python
- Exploratory data analysis (EDA)
- Development of interactive Power BI dashboards
- Business documentation (BRD, DRD, SRS, etc.)

1.3 Audience

- Business Analysts
 - Data Analysts
 - Product Managers
 - Data Science Teams
 - Airbnb Strategy & Operations Teams
-

2. System Overview

The system is a **dashboard-based data analysis tool** created with cleaned Airbnb NYC data. It helps monitor:

- Room type distribution
- Price dynamics
- Host behavior
- Review trends
- Availability patterns

Tools Used:

- **Python (for data cleaning and EDA)**
 - **Power BI (for interactive dashboarding)**
-

3. Functional Requirements

3.1 Data Ingestion

- The system should accept a .csv dataset.

- It should allow preprocessing in Python (handling nulls, removing duplicates, type conversion).

3.2 Data Cleaning

- Drop unnecessary columns.
- Fill or remove missing values.
- Convert relevant columns to appropriate data types (e.g., date for reviews, integers for price).

3.3 Data Visualization (Python)

Include charts such as:

- Average Price by Room Type
- Listings by Neighbourhood Group
- Availability vs Reviews (Scatter)
- Room Type Distribution (Pie)
- Correlation Heatmap
- Confusion Matrix (Optional if classification is applied)

3.4 Dashboard Features (Power BI)

Dashboard 1: "Airbnb Market Overview NYC"

- Bar Chart: Average Price by Room Type
- Pie Chart: Room Type Distribution
- TreeMap: Revenue Potential by Area
- Map: Price/Review distribution by Latitude/Longitude
- Cards: Total Listings, Average Price, Total Availability

Dashboard 2: "Airbnb Room Trends NYC"

- Line Chart: Review Activity Trend
- Bar Chart: Superhost vs Regular Host
- Scatter Chart: Host Listing Count vs. Reviews
- KPI Cards: Active Hosts, Occupancy Rate, Average Review Score

4. Non-Functional Requirements

Requirement Type Description

Performance	Dashboard must load under 5 seconds for basic filters
Scalability	Should be extendable to other cities or newer datasets
Usability	Interface should be intuitive for business users
Reliability	Accurate KPIs and data visualizations
Maintainability	Easy to update with new data or visuals

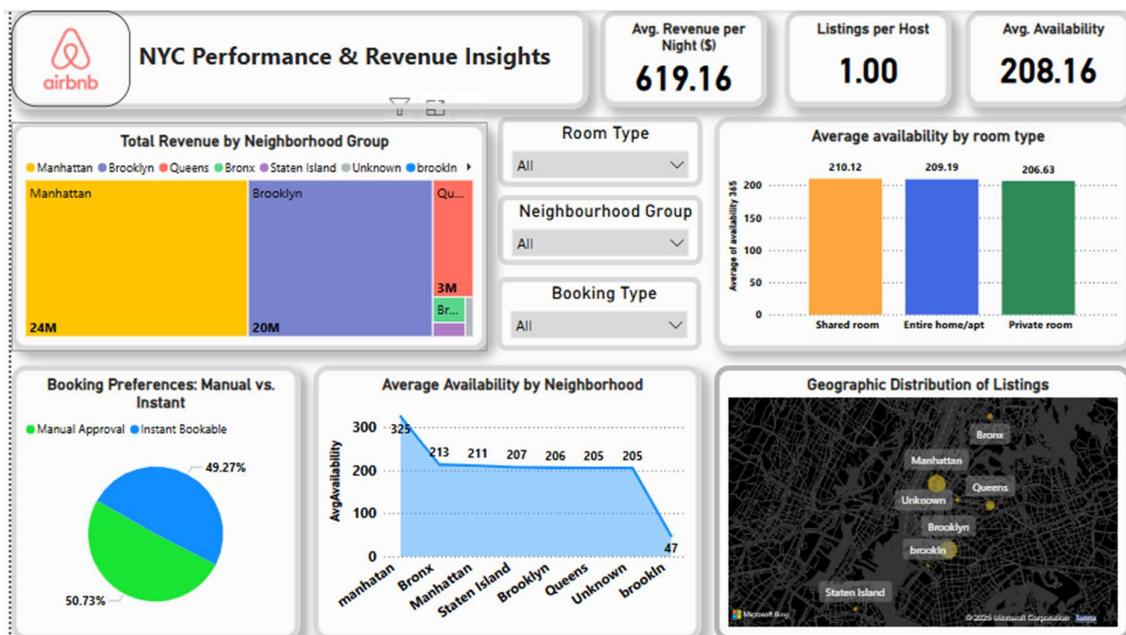
5. Dataset Description

See DRD for full list of columns used. Key ones include:
room_type, neighbourhood_group, price, number_of_reviews, availability_365,
review_rate_number, instant_bookable, etc.

6. Visuals (Insert Locations)

You can paste screenshots of your Power BI dashboards here for better clarity:

- Dashboard 1



- Dashboard 2



7. Success Criteria

- At least **6 visualizations** in Python
- **2 full dashboards** in Power BI
- Clear and concise business documents (BRD, DRD, SRS)
- Insightful conclusions and possible recommendations