PROJECT TITLE:

AIRBNB LISTING ANALYSIS – NEW YORK CITY (2019)

AUTHORS:

- **ZUERA ALABI**
- **AMADASUN SAMUEL**

DATE:

8TH MARCH, 2025

DATA ANALYSIS PROJECT

PROJECT TITLE: Airbnb Listings Analysis – New York City (2019)
DATE: [3/8/2025]

1. Introduction:

This report presents an in-depth analysis of Airbnb listings in New York City using the 2019 dataset from Kaggle. The objective of this analysis is to identify pricing trends, neighborhood popularity, and factors influencing listing performance. To achieve this, we utilized Python for data cleaning and transformation, Excel for normalization, and Power BI for data visualization and reporting.

2. Data Collection & Preparation:

Dataset Description:

- Source: Kaggle (New York City Airbnb Listings 2019)
- Number of records: Approximately 49,000 listings
- Key variables: Listing ID, Host ID, Neighborhood, Name, Reviews per months, Host Name, Neighborhood_group, Room Type, Price, Availability, Number of Reviews

Data Cleaning & Preprocessing:

- Removed missing and duplicate records.
- Standardized price and availability data.
- Handled outliers in pricing using seaborn method.
- Categorized neighborhoods for better analysis.

3. Data Analysis & Methodology:

The AIRBNB data was analyzed using the following methodologies:

- **Descriptive Analysis:** Perform descriptive analysis to understand the distribution of listings by **room type, price, and location**.
- **Trend Analysis:** Examined pricing trends across different neighborhood and room types.
- **Correlation Analysis:** Identified relationships between price, availability, and customer reviews

4. Pricing Analysis:

- Identify **key factors affecting listing prices** such as location, room type, reviews, and availability.
- Compare pricing patterns across different boroughs (**Manhattan, Brooklyn, Queens, etc.**).
- Detect outliers in price distribution.

5. Host Activity & Availability Analysis:

- Determine the **top 10 hosts with the highest listings**.
- Analyze **host types** (individual vs. commercial hosts)

6. Demand & Customer Preference Analysis:

- Identify neighborhoods with the **highest demand** based on the number of reviews and availability.
- Understand how **number of reviews** and **review scores** impact listing popularity.
- Detect areas popular for **budget travelers** vs. **luxury stays**.

7. Findings & Insights

Key Insights:

1. **Price Variations:** The median price for entire homes is significantly higher than private rooms and shared room with Manhattan being the most expensive borough.
2. **Neighborhood Popularity:** Listings in Manhattan and Brooklyn receive the highest number of reviews, indicating higher demand.
3. **Availability Trends:** Superhosts tend to have higher occupancy rates, showing a strong correlation between host reputation and bookings.

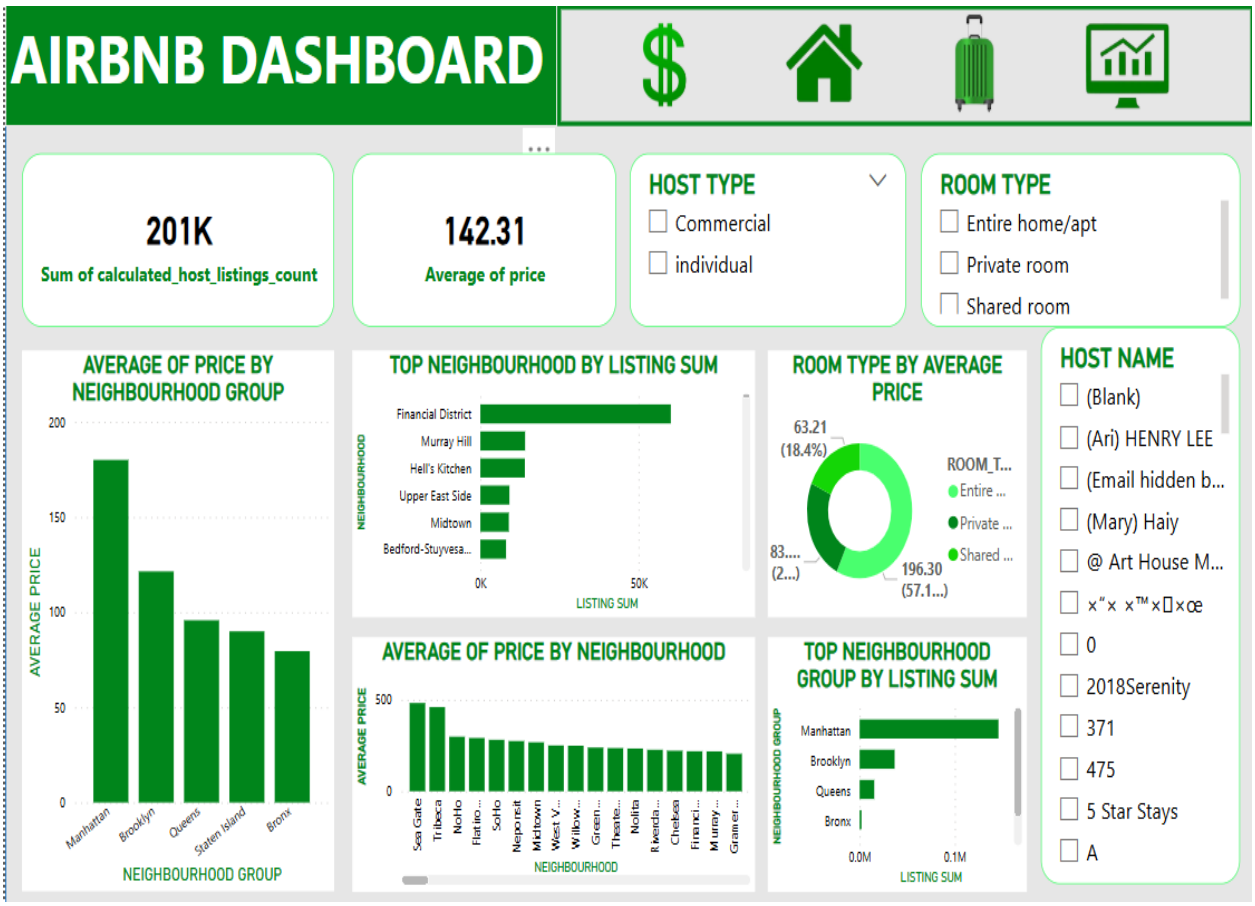
8. Visualization & Dashboard Overview:

The interactive Power BI dashboard provides a comprehensive view of:

- Average price distribution across boroughs.
- Availability trends by neighborhood.
- Average price by neighbourhood group
- Room type by average price
- Top neighbourhood by list sum

Recommendations:

1. **Hosts should optimize pricing** based on neighborhood trends to maximize revenue.
2. **S**eem price outliers.
3. Ensuring accurate geospatial mapping.
4. Visualizing complex relationships in Power BI.



9. Conclusion:

This analysis provides valuable insights into Airbnb listings in New York City, helping hosts and potential investors make informed decisions based on pricing, demand, and neighborhood popularity. Future analysis could incorporate additional datasets, such as seasonal trends and customer demographics, to enhance decision-making further.