Airbnb NYC 2019 - EDA Documentation

# 1. Project Overview

This project involves an Exploratory Data Analysis (EDA) on the Airbnb NYC 2019 dataset. The aim is to uncover trends related to pricing, availability, room types, and host behaviors in New York City. The dataset contains ~49,000 listings with features such as neighbourhood, room type, price, and reviews.

# 2. Dataset Information

- Source: Inside Airbnb (http://insideairbnb.com)  
- Format: CSV  
- Rows: 48,895  
- Columns: 16  
- Key Features:  
 - `id`, `name`, `host\_id`, `host\_name`  
 - `neighbourhood\_group`, `neighbourhood`, `latitude`, `longitude`  
 - `room\_type`, `price`, `minimum\_nights`  
 - `number\_of\_reviews`, `last\_review`, `reviews\_per\_month`  
 - `calculated\_host\_listings\_count`, `availability\_365`

# 3. Data Cleaning & Preprocessing

- Removed rows with missing 'name' and 'host\_name'  
- Filled missing 'reviews\_per\_month' with 0  
- Converted 'last\_review' to datetime format  
- Removed outliers in 'price' (> $500) and 'minimum\_nights'

# 4. Analysis Performed

- Distribution of room types and popularity by borough  
- Price comparison by neighbourhood\_group and room\_type  
- Host listing counts and availability analysis  
- Review counts and monthly review trends  
- Correlation heatmap for numerical features

# 5. Tools & Technologies Used

- Python (Pandas, NumPy, Matplotlib, Seaborn)  
- Jupyter Notebook / Google Colab  
- PowerPoint for presentation  
- Word for reporting

# 6. Key Insights

- Manhattan listings are the most expensive and frequent.  
- Private rooms are more commonly booked than shared rooms.  
- Many hosts manage multiple listings.  
- Most listings are priced under $100 per night.  
- High review rates were observed in Brooklyn and Manhattan.

# 7. Conclusion

This EDA project provides valuable insights into NYC’s Airbnb market, highlighting trends in pricing, availability, and listing popularity. The cleaned and visualized data can guide further analysis or serve as a basis for predictive modeling and dashboard creation.