



airbnb

A circular image showing a close-up of a wooden interior, possibly a cabin or rustic setting. A large, white, semi-transparent infinity symbol watermark is overlaid on the left side of the circle. The word "airbnb" is written in a large, white, sans-serif font at the bottom of the circle.

Airbnb Data Analysis Project

Comprehensive Data Insights using Python



Airbnb Company Overview

- **Founded:** 2008 by Brian Chesky, Joe Gebbia, and Nathan Blecharczyk
- **Industry:** Online marketplace for accommodations and experiences
- **Global Presence:**
 - 8+ million active Listings
 - 100,000+ cities across 220 countries
- **Revenue:** \$3.73 billion
- **Mission:** To create a world where anyone can belong anywhere





Project Overview

- **Goal:** Conduct a detailed analysis of Airbnb listing data
- **Tools Used:**
 - Jupyter Notebook: Coding environment
 - Pandas: Data cleaning and manipulation
 - NumPy: Mathematical operations
 - Matplotlib & Seaborn: Data visualization
- **Project Breakdown:**
 - a. Check Missing Values
 - b. Remove Duplicates
 - c. Visualizations based on client requirements

Data Cleaning Process



STEP 1: CHECK FOR MISSING VALUES

Identified and handled missing values to ensure data accuracy.

STEP 2: REMOVE DUPLICATES

Cleaned duplicate records to avoid skewed analysis.



Full contribution

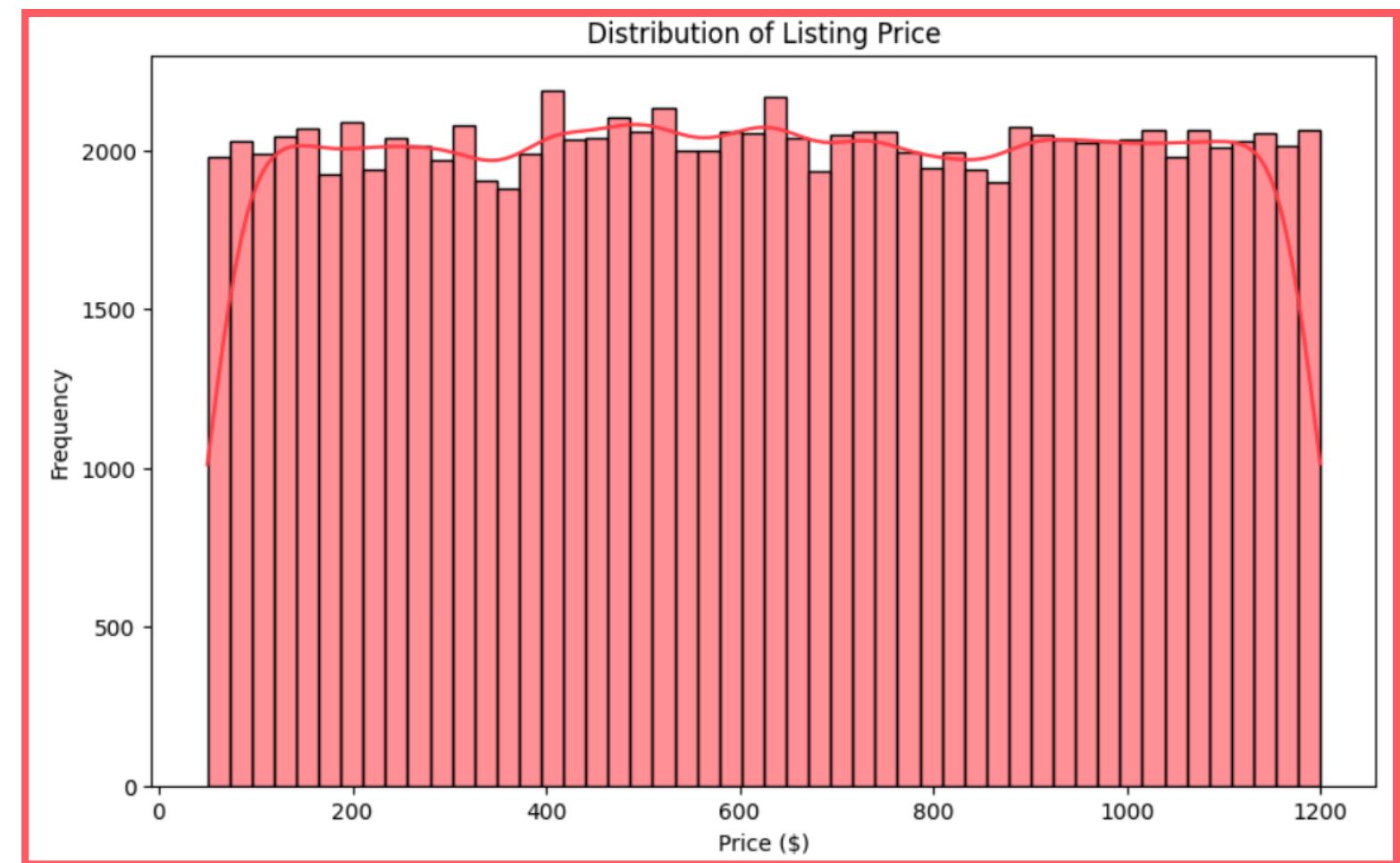
Clean, consistent, and reliable dataset



Distribution of Listing Prices



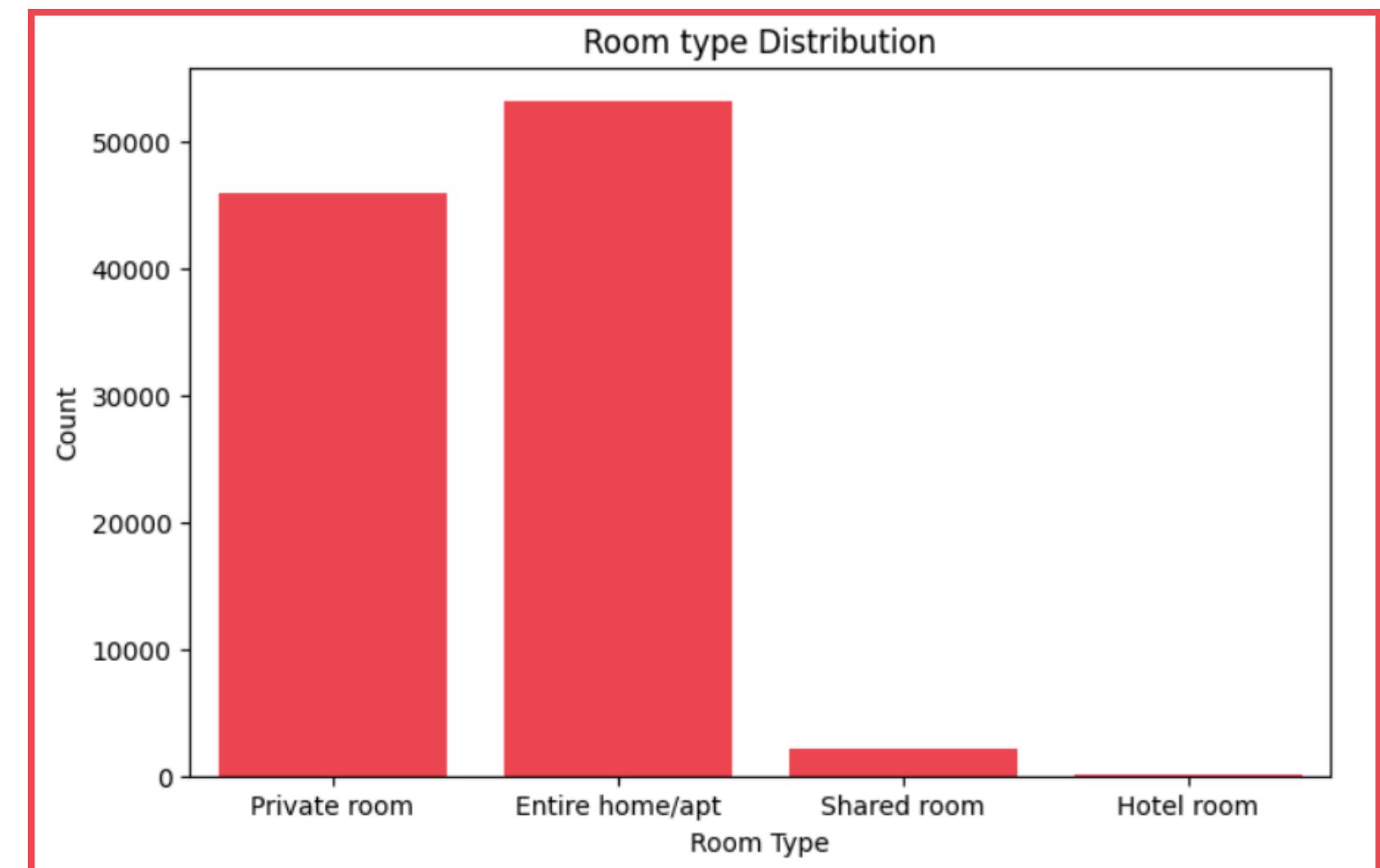
- **Question:** What is the distribution of listing prices?
- **Visualization:** Histogram with KDE (Kernel Density Estimation).
- **Insights:** Prices are evenly distributed across various ranges
 - No significant concentration in specific price brackets
 - Dataset showcases a wide variety of pricing



Room Type Distribution



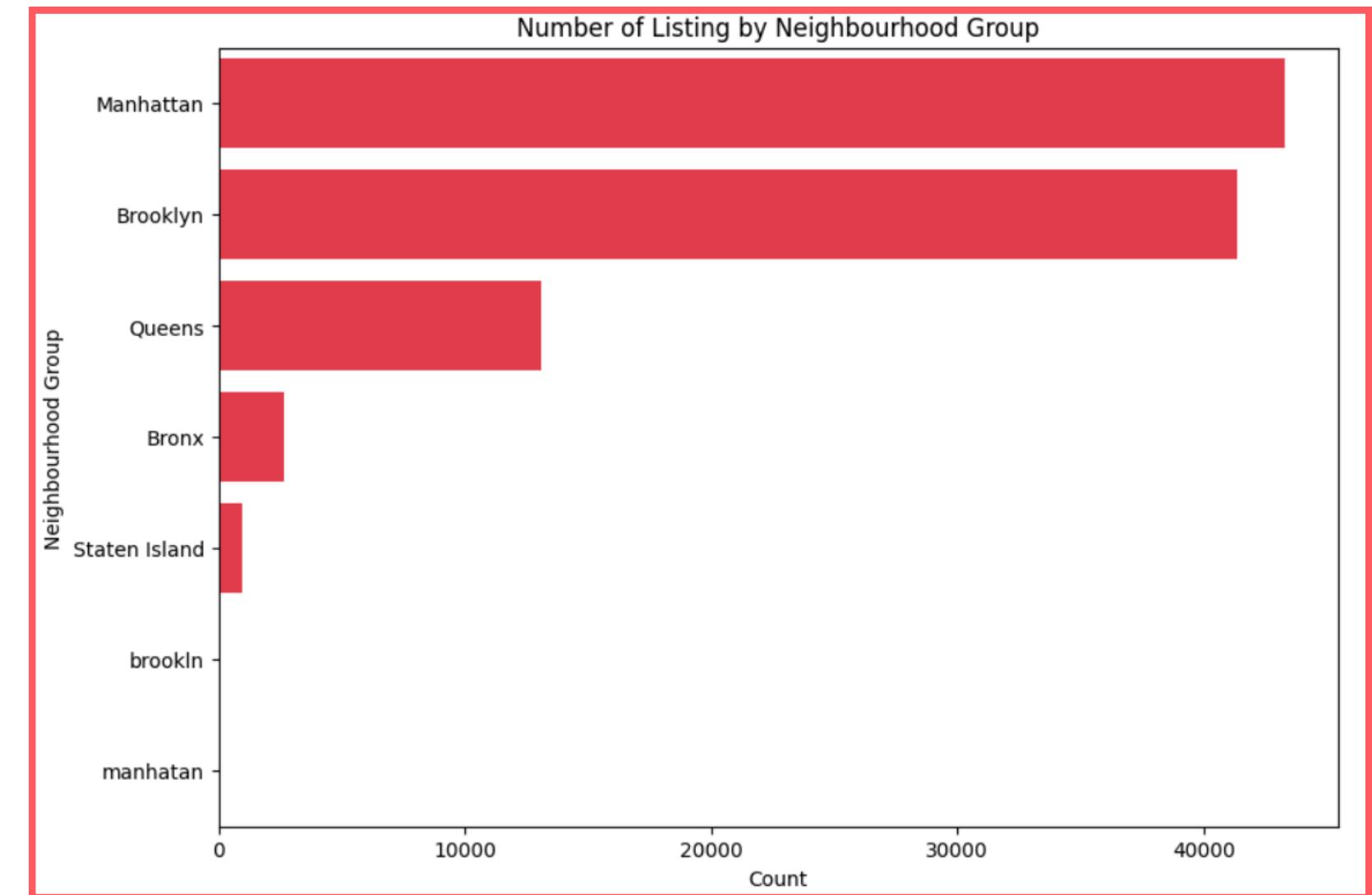
- **Question:** How are room types distributed?
- **Visualization:** Pie chart and bar plot
- **Insights:** Most preferred: Entire home/apt.
Next: Private room.
Followed by: Shared room and hotel room.
Entire home/apt has the highest share, followed by private rooms.



Neighborhood Distribution



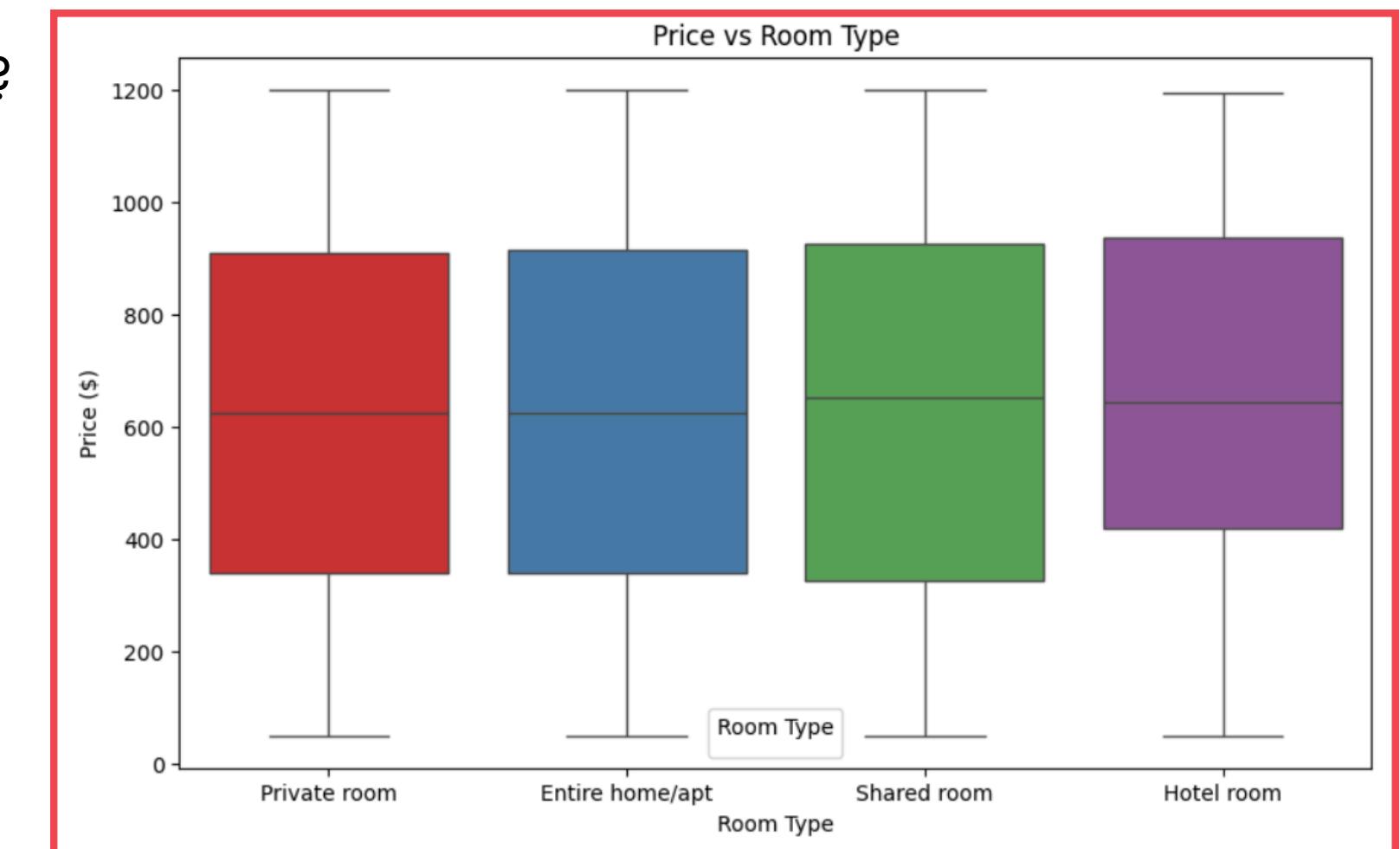
- **Question:** How are listings distributed across neighborhoods?
- **Visualization:** Bar Chart
- **Insights:**
 - Top neighborhoods:
 - Manhattan - highest number of listings.
 - Followed by Brooklyn and Queens.
 - Bronx and Staten Island have fewer listings
 - Data shows significant concentration in major boroughs.



Price vs. Room Type Relationship



- **Question:** What is the relationship between price and room type?
- **Visualization:** Box Plot
- **Insights:**
 - Similar price ranges for Private rooms. Entire homes/apt. and shared rooms : \$360 - \$900
 - Hotel rooms have slightly different price ranges: \$400 - \$920.
 - Indicates competitive pricing strategies across different room types.



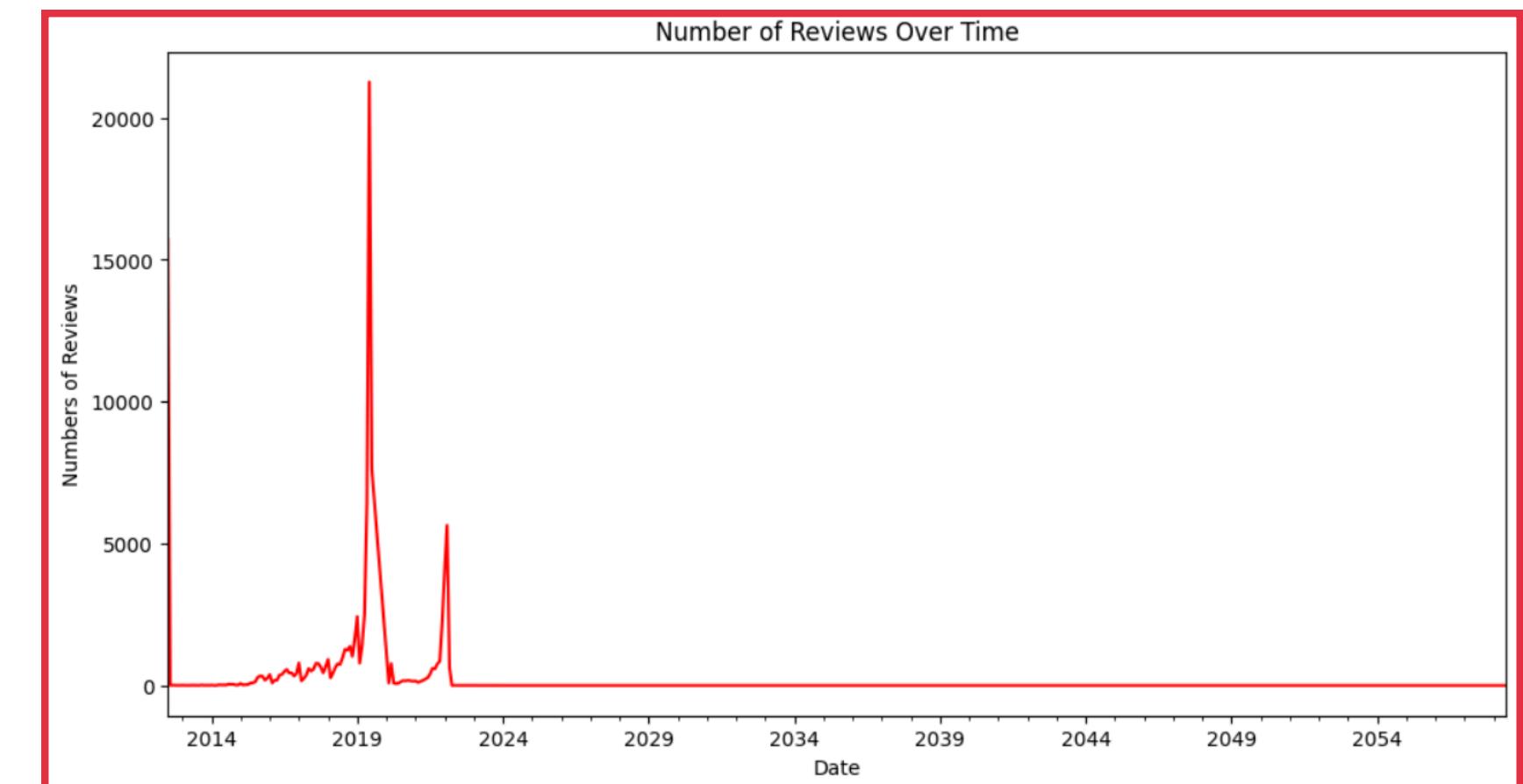
Reviews Over Time



- **Question:** How has the number of reviews changed over time ?

- **Visualization:** Line Chart

- **Insights:**
 - Top neighborhoods:
 - 2014 – 2019: Steady growth in reviews.
 - 2019 – 2024: Exponential growth in reviews.
 - Reflects increasing customer engagement and platform popularity.





Conclusion and Key Takeaways



KEY INSIGHTS:

- Airbnb listings are evenly distributed across price ranges.
- Entire home/ apt is the most preferred room type.
- Manhattan leads in listing concentration.
- Price ranges are consistent across most room types.
- Exponential growth in customer reviews post - 2019.

LEARNING OUTCOMES:

- Strengthened data cleaning, visualization, and analytical skills.
- Gained valuable insights into Airbnb's market trends.

LEARNING OUTCOMES:

- Further exploration of factors influencing pricing.
- Incorporating machine learning for predictive analysis.

A circular inset image in the bottom-left corner shows a traditional building with red-tiled roofs and wooden structures, partially obscured by red autumn leaves. A large white outline of the Airbnb logo is superimposed on this image.

Thank
You

airbnb

Jayita Mondal
jayitamondalofficial@gmail.com