NYC Airbnb Data Analysis

What Drives Price and Popularity?
Zachary Kruse

Introduction

Purpose of the Analysis:

To uncover key factors that influence Airbnb pricing and availability in New York City, using public data to guide hosts, travelers, and policymakers.

Why This Matters

Stakeholders Who Benefit:

- Hosts: Pricing competitively while maximizing earnings
- Travelers: Understanding how price and availability vary by location and room type
- Data Analysts: Practicing real-world data analysis with a rich dataset

Dataset & Tools

- Dataset: Inside Airbnb NYC (CSV)
- Unit: A single Airbnb listing in New York City
- Tools: Jupyter Notebook, Pandas, Seaborn, Matplotlib
- Data source: Inside Airbnb (http://insideairbnb.com)

Data Cleaning

Steps Taken:

- Removed duplicates and rows with missing values in key columns
- Cleaned price values by removing $ and , symbols and converting to float
- Standardized column names

Questions Answered

- 1. What factors influence the price of Airbnb listings in NYC?
- 2. Which neighborhoods have the highest and lowest average prices?
- 3. How does room type affect price and availability?
- 4. What is the relationship between the number of reviews and a listing's popularity?
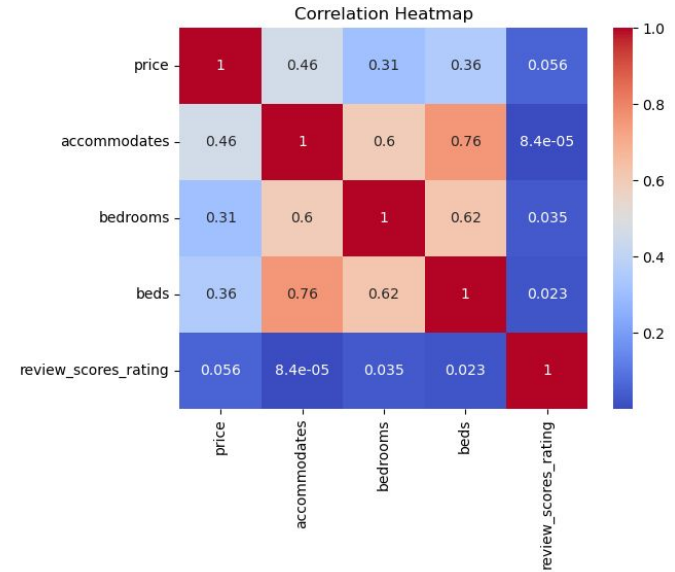
What factors influence Airbnb price?

Correlation Heatmap

Methods:

- Correlation analysis with price
- Grouped bar charts and heatmaps

Insight:

- The number of accommodations, bedrooms, and beds are positively correlated with price.
- Ratings have weak correlation, implying other qualitative factors may be at play.
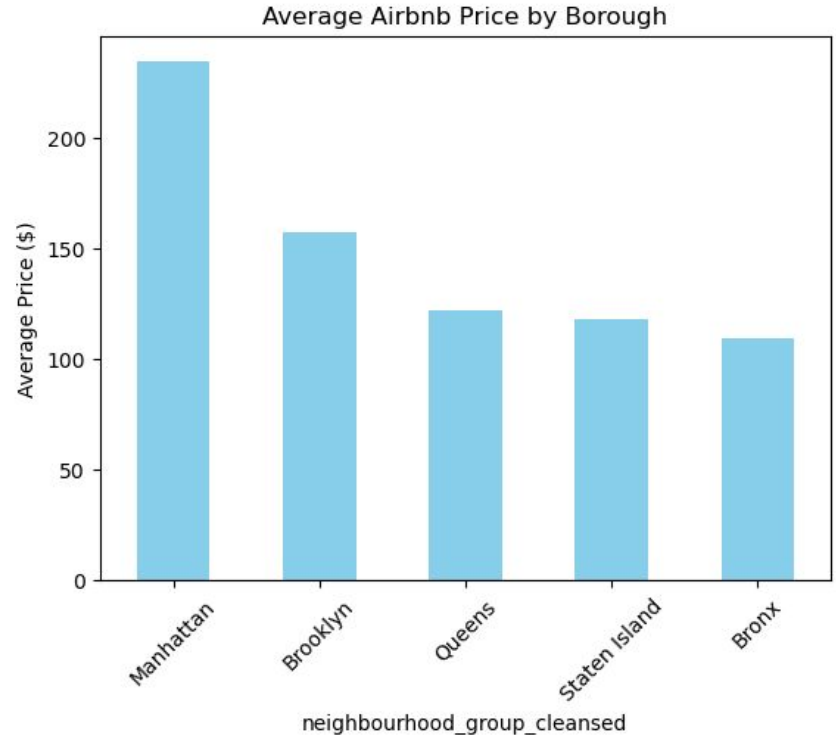
Which neighborhoods have the highest and lowest average prices?

Methods:

- Aggregated average price by neighbourhood and sorted

Insight:

- Manhattan has the highest average listing prices, followed by Brooklyn.
- The Bronx has the lowest.

Average Airbnb Price by Borough

How does room type affect price and availability?

Methods:

- Boxplots of availability_365 by room_type
- Violin plots or histograms for price by room_type

Insight:

- Hotel rooms are most expensive and moderately available.
- Shared rooms are cheaper and more widely available.

Average Price and Availability by Room Type

What is the relationship between the number of reviews and a listing's popularity?

Methods:

- Scatter plot of number_of_reviews vs. price
- Distribution of reviews

Insight:

- Most listings with high review counts are lower-priced.
- Popular listings tend to be budget-friendly, suggesting high turnover and value appeal.

Reviews vs. Price
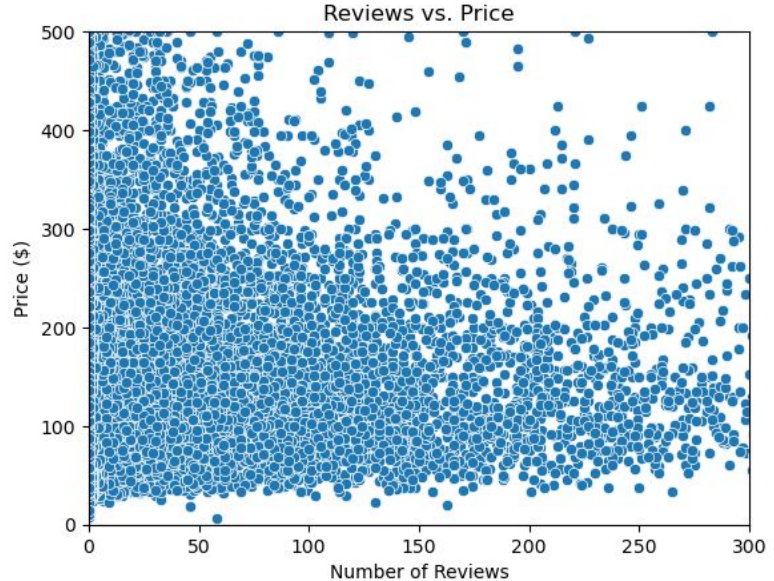
Future Work

- Conduct sentiment analysis on review comments.
- Apply machine learning to predict listing price or popularity.
- Analyze seasonal trends in pricing and availability

Appendix

- Data source: Inside Airbnb (http://insideairbnb.com)
- Tools: Jupyter, Python, Seaborn

Price Distribution by Room Type

Average Airbnb Price by Neighborhood Group