

Detailed Analysis Report: Airbnb in NYC

1. Introduction

This report presents a comprehensive analysis of Airbnb listings in New York City. The dataset, sourced from Huggingface Hub, includes over 20,000 listings, providing insights into various aspects such as room prices, reviews, availability, and host information.

2. Methodology

The analysis involved several stages, including data cleaning and preprocessing, exploratory data analysis (EDA), geographical analysis, and host analysis. Tools used included Python with libraries such as Pandas, Seaborn, and Folium for data manipulation and visualization.

3. Data Cleaning and Preprocessing

- **Missing Values:** Identified and addressed missing values in 'name' and 'host_name' columns.
- **Outliers:** Capped extreme values in 'price' and 'minimum_nights' at their 99th percentiles to mitigate their impact.

4. Exploratory Data Analysis

- **Price and Reviews Distribution:** Prices mostly ranged between \$50 and \$200, with many listings having a small number of reviews.
- **Room Types:** 'Entire home/apt' was the most common listing type, but 'Private rooms' received slightly more reviews on average.

5. Geographical Analysis

- **Heatmap Visualization:** Displayed the concentration of listings across NYC.
- **Neighborhood Analysis:** Manhattan emerged as the most expensive neighborhood group. Notable differences in average prices were observed across neighborhoods.
- **Deeper Geographic Analysis Findings:**
 - Top 5 Neighborhoods by Average Price:

- **Fort Wadsworth:** \$799 per night.
- **Woodrow:** \$700 per night.
- **Tribeca:** \$355 per night.
- **Sea Gate:** \$316 per night.
- **Flatiron District:** \$299 per night.
- **Top 5 Neighborhoods by Average Availability:**
 - **Fort Wadsworth:** 365 days per year.
 - **Co-op City:** 364 days per year.
 - **Willowbrook:** 351 days per year.
 - **Eastchester:** 333 days per year.
 - **Richmondton:** 300 days per year.
- **Top 5 Neighborhoods by Average Number of Reviews:**
 - **Silver Lake:** 118.5 reviews per listing.
 - **East Elmhurst:** 81.66 reviews per listing.
 - **Richmondton:** 79 reviews per listing.
 - **Eltingville:** 76 reviews per listing.
 - **Mount Eden:** 70 reviews per listing.
- **Top 5 Neighborhoods by Average Number of Reviews:**
 - **Silver Lake:** 118.5 reviews per listing.
 - **East Elmhurst:** 81.66 reviews per listing.
 - **Richmondton:** 79 reviews per listing.
 - **Eltingville:** 76 reviews per listing.
 - **Mount Eden:** 70 reviews per listing.
- **Average Price by Neighbourhood Group:**
 - **Manhattan:** Highest average price at approximately \$183 per night.
 - **Brooklyn:** Next, with an average price of about \$119 per night.
 - **Staten Island:** About \$102 per night.
 - **Queens:** Around \$96 per night.
 - **Bronx:** The lowest average price at approximately \$86 per night.
- **Top 10 Neighborhoods by Average Price:**
 - **Fort Wadsworth:** \$799 per night (note: this might be influenced by a limited number of listings).
 - **Woodrow:** \$700 per night.
 - **Tribeca:** \$355 per night.
 - **Sea Gate:** Around \$316 per night.
 - **Flatiron District:** Approximately \$299 per night.
 - **Prince's Bay:** Around \$297 per night.
 - **Riverdale:** Approximately \$287 per night.
 - **NoHo:** Around \$283 per night.
 - **Neponsit:** About \$275 per night.
 - **Midtown:** Approximately \$271 per night.
- **Insights:**
 - Manhattan stands out as the most expensive neighborhood group, which aligns with its reputation as a prime location in NYC.

- Within individual neighborhoods, there's a wide range in average prices, with some areas like Fort Wadsworth and Woodrow showing exceptionally high averages. This could be due to a small number of luxury listings.

6. Host Analysis

- Top Hosts by Number of Listings:

- The top hosts in terms of the number of listings they have are identified by their `host_id`. The host with ID 219517861 has the highest number of listings at 327, followed by 107434423 with 232 listings, and so on.

- Impact on Prices and Reviews:

- **Price:** The average prices for these top hosts' listings vary. For instance, the host with the most listings (219517861) has an average price of approximately \$253 per night, whereas 137358866 (with 103 listings) has a lower average price at around \$44 per night.

- **Number of Reviews:** The average number of reviews per listing also varies among these hosts. Notably, hosts with a high number of listings tend to have fewer reviews per listing on average. For example, 107434423 (232 listings) has an average of only 0.125 reviews per listing.

- Insights:

- **Price Range:** There's a wide range in average prices among hosts with many listings. Some are priced in the mid to high range, while others are more affordable.

- **Reviews:** Hosts with a large number of listings generally have fewer reviews per listing. This could be due to the listings being newer, less personal interaction with guests, or the nature of the listings (e.g., more business-oriented).

- **Listing Management:** Managing a high number of listings might impact the overall guest experience, as indicated by the lower average number of reviews.

7. Conclusions

The analysis provided valuable insights into the Airbnb market in NYC. Key findings include:

- A wide range of prices and availability across different neighborhoods.
- The number of listings a host manages can significantly influence the price and the number of reviews of the listings.

8. Recommendations

- **For Hosts:** Understanding neighborhood trends and pricing strategies can enhance competitiveness and occupancy rates.

- **For Guests:** Awareness of price variations and popularity trends across neighborhoods can inform better booking decisions.

This report encapsulates the breadth of analysis conducted. For a more detailed breakdown, including data visualizations and specific data points, refer to the accompanying visual materials and data files.