Pillow Palooza NYC Short-Term Rental Insights

By Lex Romo ***June 4th, 2023***

Introduction

The real estate industry is a highly competitive and dynamic market, demanding continuous adaptation and innovation to thrive. As a Data Analyst at Pillow Palooza, I am dedicated to extracting valuable insights from the highly competitive and dynamic real estate market, specifically the short-term rental sector in New York City. My role involves collecting, cleansing, and analyzing data to unlock growth opportunities and drive success for Pillow Palooza in this challenging landscape.

In response to the Head of Data's request, I undertook a comprehensive project that involved gathering Airbnb listing data from various sources and employing rigorous data cleaning techniques to ensure accuracy. The objective was to develop a deep understanding of the market by uncovering key trends related to popular neighborhoods, rental prices, property types, length of stay, and demand patterns over time.

The analysis conducted as part of this project holds immense potential in providing invaluable insights that can guide Pillow Palooza's business decisions. By identifying emerging trends and patterns, we can pinpoint profitable neighborhoods for investment, determine the most sought-after property types, and strategically price our rentals to remain competitive.

In the forthcoming report, I will present the findings of the analysis, emphasizing the key trends and insights uncovered. By examining various facets of the short-term rental market, our aim is to provide a comprehensive view that enables Pillow Palooza to make informed and strategic choices. With a forward-looking perspective, this project lays the foundation for Pillow Palooza's success in the dynamic real estate landscape of New York City.

Data Collection and Preparation

Part 1:

In this project, multiple datasets were collected, cleaned, and merged to gain insights into the short-term rental market in New York City. The datasets included the Price Dataset, Room Type Dataset, and Reviews Dataset.

The Price Dataset consisted of three columns: listing_id, price, and nbhood_full. Initially, listings with a price of 0 were removed. A new column named price_per_month was created by multiplying the price by 365 and dividing it by 12 to represent the monthly price. The nbhood_full column contained both the borough and neighborhood information, which was split into separate columns named Borough and Neighborhood. Last;y, 7 Listings with a price of $0.00 were removed.

The Room Type Dataset contained three columns: listing_id, description, and room_type. The values in the room_type column were converted to lowercase, and the data type was changed from text to category for better analysis.

Part 1 Cont'd:

The Reviews Dataset included three columns: listing_id, host name, and last_review. The last_review column's data type was changed from text to datetime to enable date-based analysis.

Using pandas, all the datasets were merged into a single dataset. A new column named price_range was created based on the Price column. The values in the price_range column were assigned as follows: if the price was less than or equal to 69, it was labeled as "Budget"; if it was less than or equal to 175, it was labeled as "Average"; if it was less than or equal to 350, it was labeled as "Expensive"; otherwise, it was labeled as "Extravagant". This jupyter notebook can be found under Project/data_cleaning.ipynb.

The resulting merged dataset included the following columns: listing_id, price, Borough, Neighborhood, price_per_month, description, room_type, host_name, last_review, and price_range. The final merged dataset was saved in folder Part_1_Datasets under "all_data_cleaned.csv".

Part 2:

Additional data was later gathered which resulted in columns such as Minimum_Nights, Number_Of_Reviews, Reviews_Per_Month, Availability_365, and Booked_Days_365. The three same datasets, prices, room_types & reviews, were merged into one dataset using SQL and saved under Project/Part_2_Datasets/pillow_final.csv. The SQL queries can be found under Project/SQL/queries.txt. **Please note, the dataset from Part 1 which was used for cleaning is different from the dataset in Part 2.**

The final dataset, enriched with various columns, provides a comprehensive view of the short-term rental market in New York City. This dataset serves as the foundation for further analysis and exploration of trends, patterns, and correlations within the market.

Key Findings

<u>Listings:</u>

In our dataset, we discovered a total of 25,202 unique listings. To provide a comprehensive overview, we have organized these listings based on their price range and borough.

Brooklyn has the highest share of listings, accounting for 41.50%, closely followed by Manhattan with 40.96%. Queens comprises only 13.71% of the listings, while the combined share of the Bronx and Staten Island is 3.83%.

	Borough				
price_range	Bronx	Brooklyn	Manhattan	Queens	Staten Island
Average	286	5,534	5,289	1,505	123
Budget	381	3,201	1,150	1,632	124
Expensive	25	1,466	3,073	291	20
Extravagant	5	259	810	28	

Price:

The dataset provides various statistical measures for analyzing the prices of listings:

The average price of listings is $141.82, serving as a measure of the central tendency of prices in the dataset. The median price, $105.00, represents the middle value and gives an indication of the typical price, less affected by extreme values. The maximum price observed is $7500.0, indicating the highest price among all listings. The mode, $150.00, reveals the most common price category in the dataset.

The price range is $7490.0, representing the difference between the maximum and minimum prices and indicating the spread of prices across the dataset. The interquartile range (IQR) for all data is 106.00, capturing the range within which the central 50% of prices fall. The standard deviation, 147.35, reflects a relatively large spread of prices from the mean.

Additionally, the dataset includes 1326 outliers, which are values significantly deviating from the rest of the data. These outliers may represent unusual or extreme price points deserving further examination.

Booked Days 365:

Booked Days and Price Correlation: The correlation coefficient between Booked_Days_365 and Price is -0.07964. This correlation coefficient suggests a weak negative relationship between the length of stay and the price of the listings in the dataset.

The analysis of the review dates as an approximation for bookings throughout the year reveals interesting insights. The "Budget" price range category emerges as the most popular category overall, consistently attracting a high number of bookings throughout the year. In particular, June experiences a peak with a total of 163,381 booked days, indicating a strong demand for affordable accommodations during that month.

December stands out as another noteworthy period, with a significant increase in bookings, totaling 59,196 booked days. This surge in bookings can be attributed to the holiday season, as many tourists flock to New York City to celebrate the New Year in iconic locations like Times Square.

Booked Days 365 Cont'd:

The price range categories of "Budget," "Average," and "Expensive" all exhibit a similar pattern, surpassing 10,000 booked days starting in early May and continuing to rise until the end of June. However, after July 5, there is a lack of further information as it marks the last review timestamp available in the dataset.

Overall, the analysis provides valuable insights into the booking trends across different price ranges throughout the year. It highlights the popularity of the "Budget" category, the impact of holiday seasons on bookings, and the time frame in which bookings tend to peak.

Room Types:

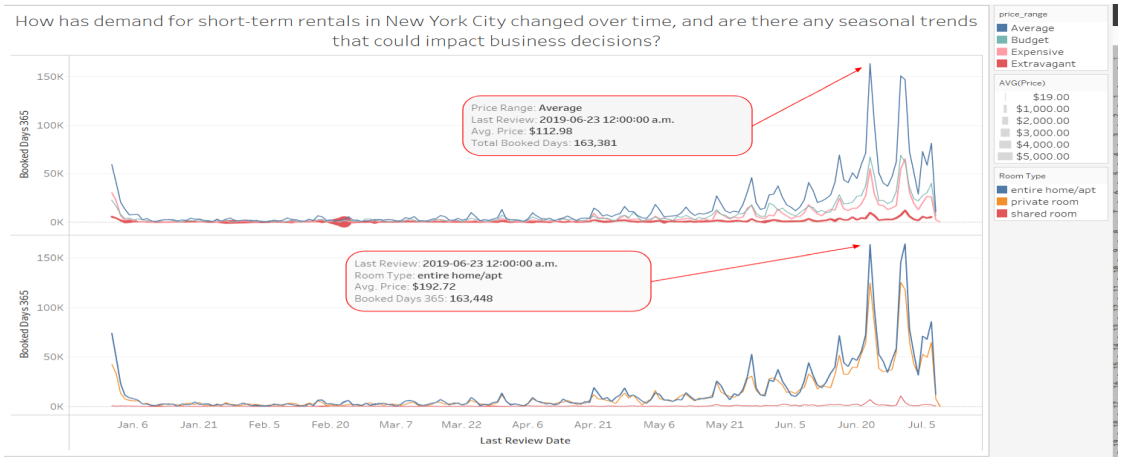
The analysis of room types reveals interesting patterns and preferences among Airbnb listings. The three main categories, namely Entire Home/Apt, Private Room, and Shared Room, exhibit distinct characteristics in terms of the number of listings and average prices.

Among these categories, Entire Home/Apt stands out as the most prevalent, with a total of 13,266 listings and an average price of $197.00. It is followed closely by Private Room, which boasts 11,351 listings and an average price of $82.00. Although Shared Room has the lowest number of listings, with only 585, it offers the most affordable option, with an average price of $54.00.

Furthermore, when considering the review dates as an approximation for bookings throughout the year, it becomes evident that Entire Home/Apt is the most popular choice among travelers. It experiences a peak with a total of 164,162 booked days, indicating a strong preference for having an entire home or apartment to themselves. On the other hand, Shared Room has the lowest number of booked days, peaking at only 10,659.

These findings emphasize the dominance of Entire Home/Apt listings in terms of popularity and demand. It is worth noting that Private Room closely follows Entire Home/Apt, indicating that many guests still value the option of having their own private space within a larger property. Conversely, the Shared Room category appears to be less popular, potentially due to its communal nature and lower number of available listings.

Overall, this analysis sheds light on the distribution of room types, their respective average prices, and the booking trends associated with each category. These insights can inform strategic decision-making for hosts and help them cater to the preferences and needs of their target audience.



Boroughs & Neighbourhoods:

The top 10 neighborhoods, including East Village, Williamsburg, and East Harlem, demonstrate high average booked days ranging from 210 to 256 days, along with average prices per stay ranging from $104 to $190.

Brooklyn stands out as the leading borough in terms of the number of listings, with a total of 10,460 listings, accounting for 41.50% of the dataset. With an average monthly price of $3,710.00, these listings generate a substantial annual revenue of $279,130,240.00.

Manhattan closely trails behind with 10,322 listings, representing 40.96% of the total. The annual revenue generated by Manhattan listings amounts to $393,405,670.00, with an average monthly price of $5,597.00.

In third place is Queens, with a total of 3,456 listings, making up 13.71% of the dataset. These listings contribute to an annual revenue of $58,404,083.00.

The Bronx and Staten Island, although accounting for less than 3% of the total listings, still generate a combined annual revenue of $12,768,099.00.

Hosts:

Reviews:

Brooklyn takes the lead in terms of total reviews, particularly in the "Average" price range, with a substantial count of 254,489 reviews. Following closely behind is the "Budget" price range, which receives 112,421 reviews.

Moving to Manhattan, the "Average" price range receives the highest number of reviews with a total of 229,668. In second place is the "Expensive" price range, accumulating 101,529 reviews.

In Queens, the "Budget" price range is the most reviewed category, garnering a total of 72,163 reviews. Not far behind is the "Average" price range, with 62,943 reviews, making it a close contender for the top spot.

Recommendations

Based on the analysis of the Airbnb listing data, the following recommendations can be made:

Focus on the "Entire Home/Apt" Category: Given its popularity and desirability among renters, it is recommended to prioritize and expand the availability of entire homes or apartments in your rental portfolio. This category has consistently shown higher demand and booking rates, indicating a strong preference among travelers.

Target Top Neighborhoods: Concentrate your efforts and resources on the top-performing neighborhoods such as East Village, Williamsburg, and East Harlem. These neighborhoods exhibit high average booked days and attractive average prices per stay, indicating a favorable market for short-term rentals. Investing in these areas and optimizing your listings can lead to increased occupancy and revenue.

Pay Attention to Brooklyn and Manhattan: These two boroughs, particularly Brooklyn and Manhattan, are the most promising areas for your rental business. They account for a significant portion of the total average booked days, suggesting a higher potential for bookings and profitability. Allocate resources and marketing efforts to maximize your presence in these boroughs.

Consider Seasonal Demand: Take advantage of the seasonal booking patterns observed in the data. June, with its peak in bookings, presents an opportunity to capitalize on the demand for affordable accommodations. Additionally, December experiences a surge in bookings, likely due to the holiday season. Tailor your pricing and marketing strategies to cater to these seasonal peaks and adjust your inventory accordingly.

Continuously Collect and Analyze Data: To stay updated on the evolving market trends and patterns, it is recommended to regularly collect and analyze data beyond the available dataset. This will provide more accurate and up-to-date insights, allowing you to make informed decisions and adapt your strategies accordingly.

By implementing these recommendations, you can enhance your competitiveness, optimize occupancy rates, and maximize revenue in the dynamic short-term rental market of New York City. Stay agile, monitor market changes, and align your offerings with the preferences and demands of your target audience to ensure sustained growth and success in the industry.

Conclusion

The analysis of Airbnb data provides valuable insights for making informed decisions in the short-term rental market. It is crucial to prioritize the "Entire Home/Apt" category, which consistently attracts higher popularity among users.

Brooklyn and Manhattan stand out as the top boroughs, with a significant share of average booked days. Focusing efforts on these areas can significantly increase rental occupancy and revenue.

Examining review dates reveals notable trends. The "Budget" price range consistently generates bookings, with a peak in June. Additionally, December experiences increased demand, likely due to the holiday season.

It's important to note that the dataset lacks information beyond July 5, which limits a comprehensive analysis of booking trends throughout the year. Collecting additional data would provide a more comprehensive understanding.

These insights serve as a foundation for strategic decision-making in pricing, investment, and marketing. By leveraging these findings, Pillow Palooza can optimize their offerings, adapt to customer demands, and thrive in New York City's dynamic short-term rental market.

Appendix

Tableau Visualizations