

Storytelling Case Study: Airbnb, NYC

Debanik, Pratyusha and
Abhishek



CONTENTS

- 01 Objective
- 02 Problem Statement
- 03 Insights
- 04 Recommendations
- 05 Appendix – Methodology Document

OBJECTIVE

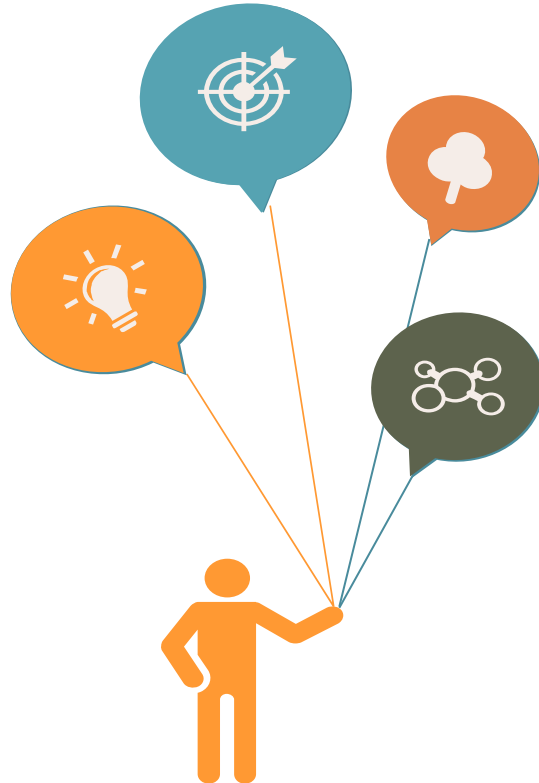
Data Analysis

1. Get a better understanding about Airbnb listings with respect to various parameters.
2. Understand the customer preferences.
3. Understand the customer booking trend.

Exploratory Data Analysis

To understand some important insights, we have explored the following questions:

1. How are the Airbnb listings spread out in NYC?
2. What type of rooms do customers prefer?
3. What could be the ideal number of minimum nights to increase customer bookings?



Data Presentation

Based on customer review:

1. Most preferred neighborhood
2. Most preferred room type
3. Who are the Hosts who have the highest listings w.r.t Neighborhood?

Methodology

- > The data was analyzed through univariate and bivariate analysis.
- > The analysis and visualizations were done using Tableau considering various parameters.
- > The main parameters that have been taken into account for analysis are –
 1. Geography based bookings
 2. Bookings based on room type
 3. Number of reviews
 4. Minimum number of nights

PROBLEM STATEMENT

For the past few months, Airbnb has seen a major decline in revenue.

Now that the restrictions have started lifting and people have started to travel more, Airbnb wants to make sure that it is fully prepared for this change.



The different leaders at Airbnb want to understand some important insights based on various attributes in the dataset so as to increase the revenue

Prepare for the next best steps that Airbnb needs to take as a business to increase revenue.

DATA ANALYSIS

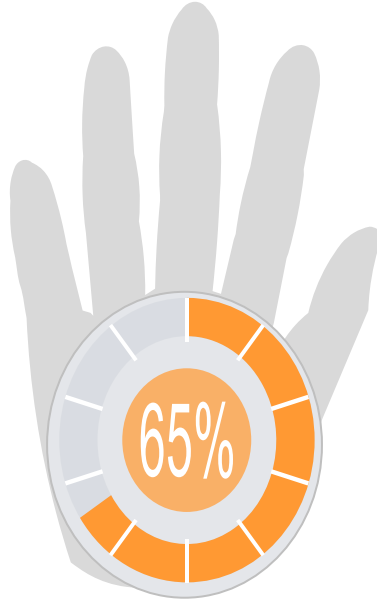
Read Data

Data is read using python and tableau.



EDA

Analyze and clean data and create bins for continuous variables



Recommendations

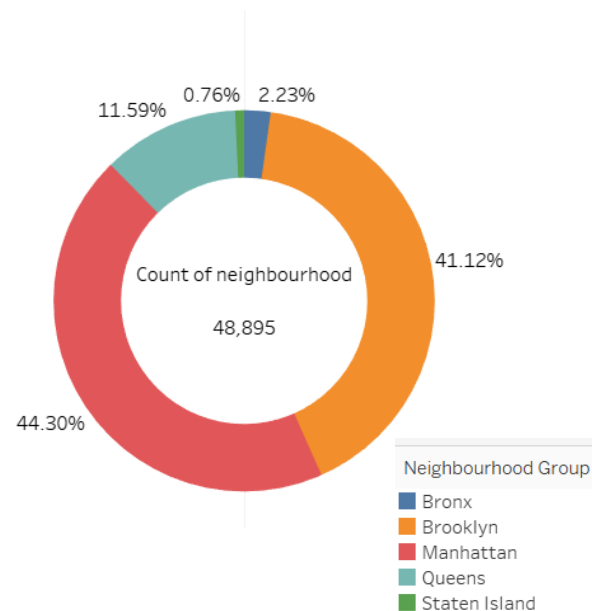
Perform univariate, bivariate analysis and correlations to capture insights



Visualizations

Create visualizations using Tableau for further insights

Insights – Neighborhood

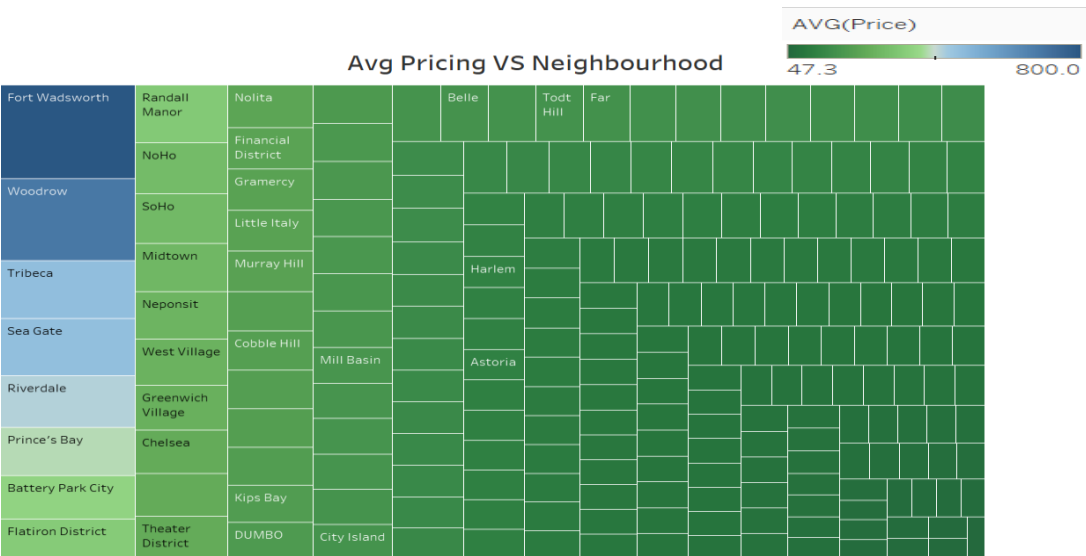


neighbourhood_group

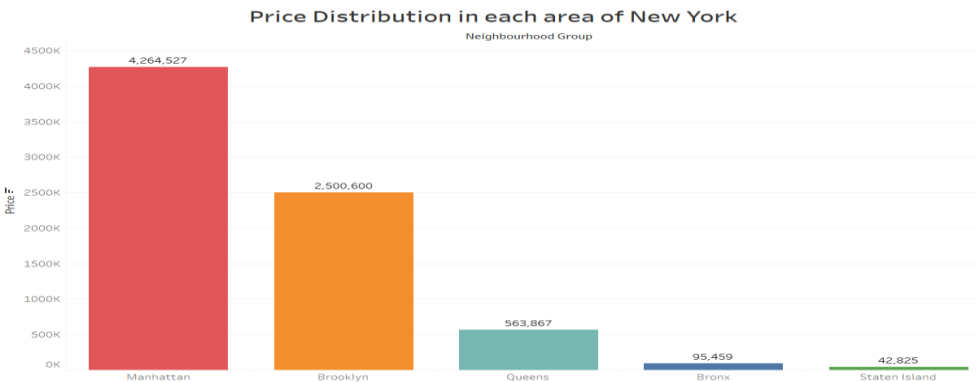
Manhattan	44.301053
Brooklyn	41.116679
Queens	11.588097
Bronx	2.231312
Staten Island	0.762859

Observation:

85% of the listing are Manhattan and Brooklyn neighbourhood_group



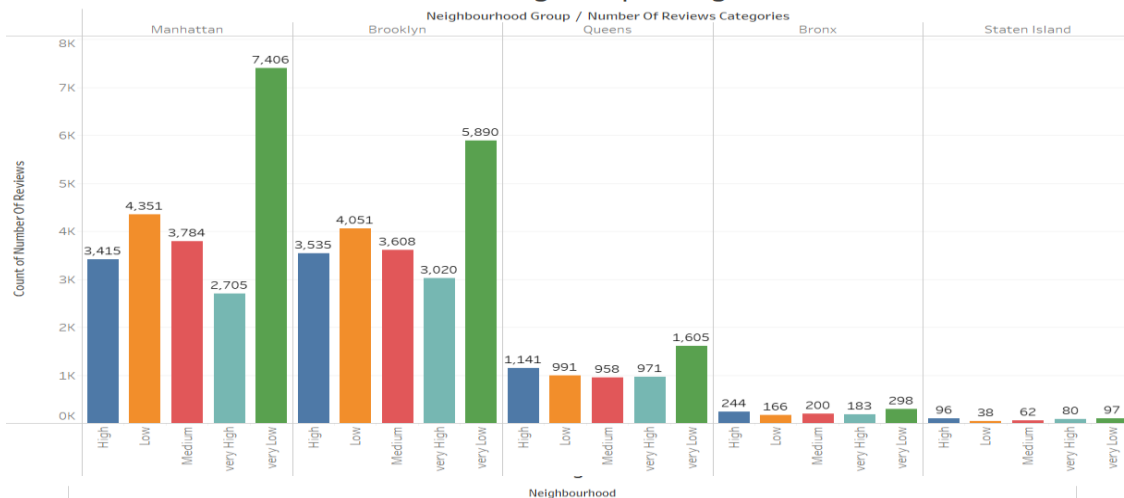
Average pricing is higher than average in only 6 neighborhoods



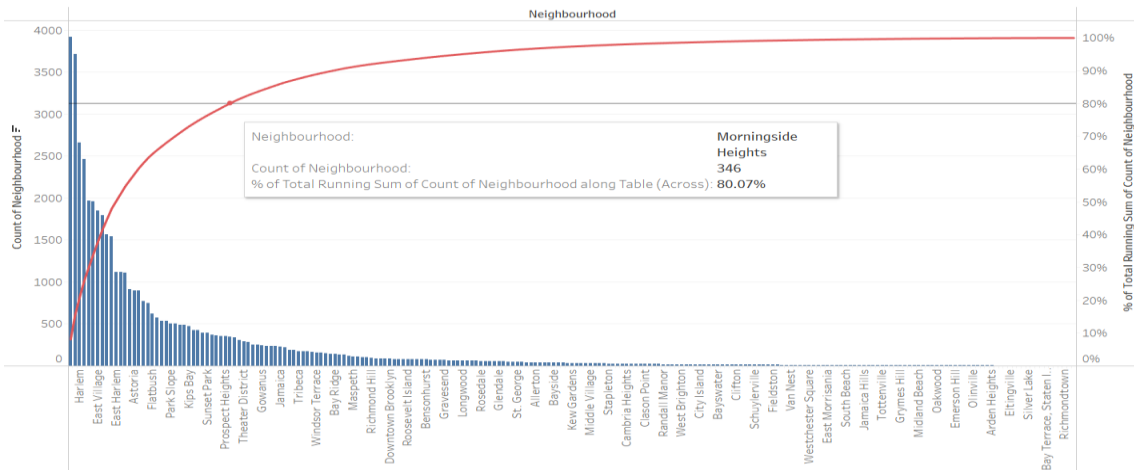
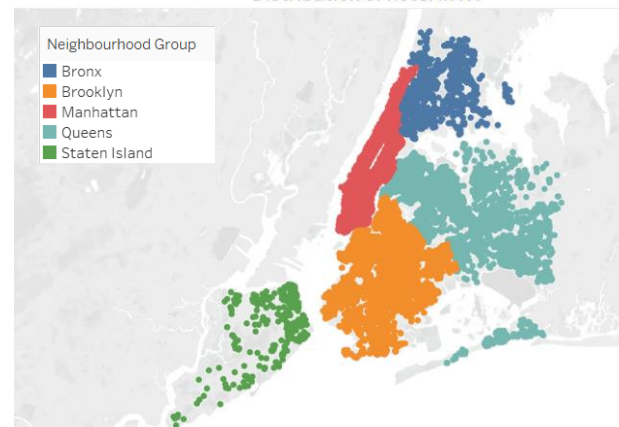
Manhattan and Brooklyn are more expensive than other neighborhood groups

Insights – Neighborhood Contd.

Count of review categories per neighbourhood

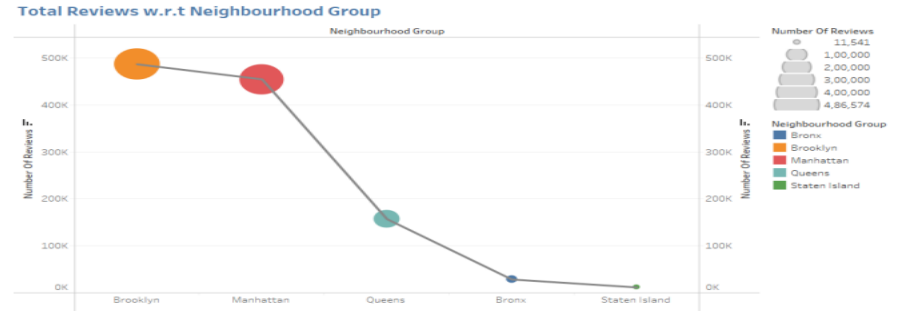
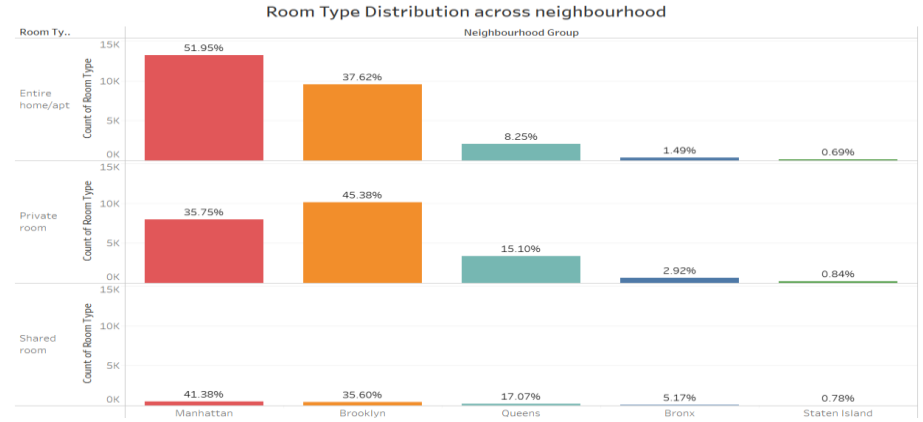
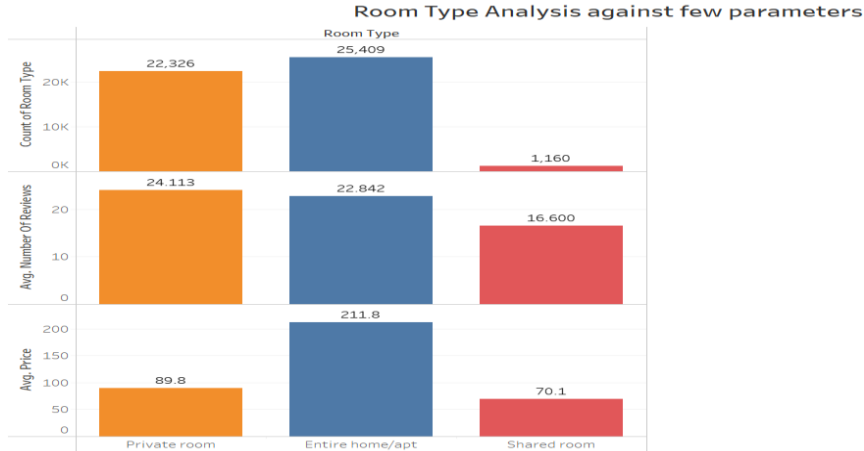
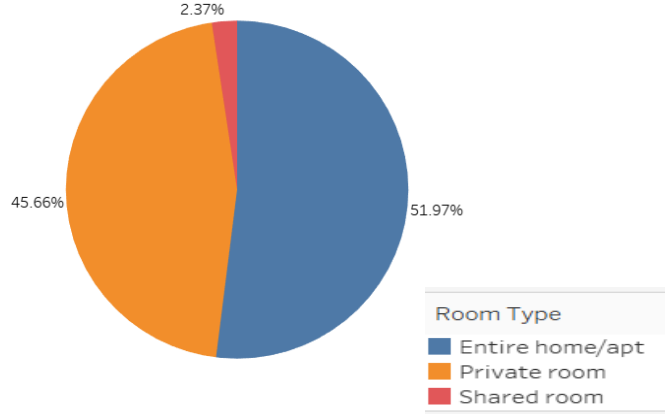


Distribution of hotel in NY



- 346 neighborhoods make up 80% of the entire neighborhoods
- Lower reviews are observed across Manhattan and Brooklyn
- Airbnb's footprint appears dominant in Manhattan, Brooklyn, and Queens. These boroughs boast the highest number of listings, likely due to their high population density and central roles as NYC's financial and tourism hubs.
- Conversely, Staten Island sees minimal Airbnb presence. With its lower population density and fewer tourist attractions, listings dwindle to around 1%, reflecting a less vibrant market for short-term rentals.

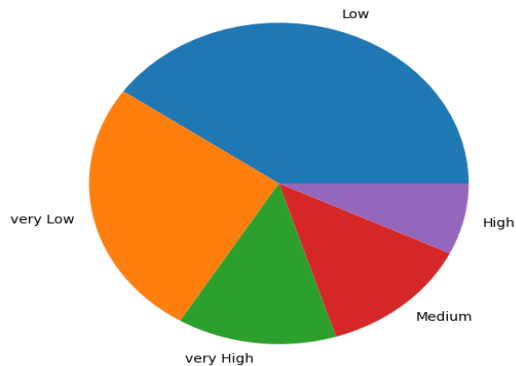
Analysis – Room Types



- Private rooms and Entire home/apt take up 97% of the room types
- Manhattan and Brooklyn have more Pvt rooms and Entire home/apt
- Pvt rooms and Entire home/apt have more rooms, reviews and are pricier than shared rooms
- The popularity of Manhattan and Brooklyn for bookings is reinforced by positive customer reviews, highlighting their appeal.

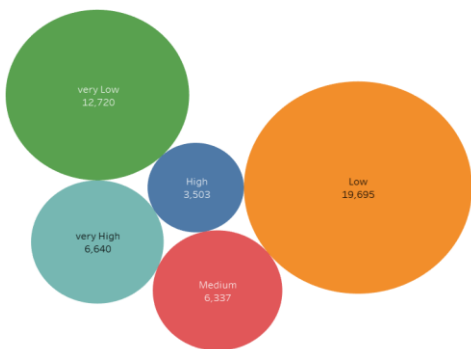
Analysis – Minimum Night Categories

Minimum night categories



```
minimum_night_categories
Low          40.280192
very Low     26.014930
very High    13.580121
Medium       12.960425
High         7.164332
```

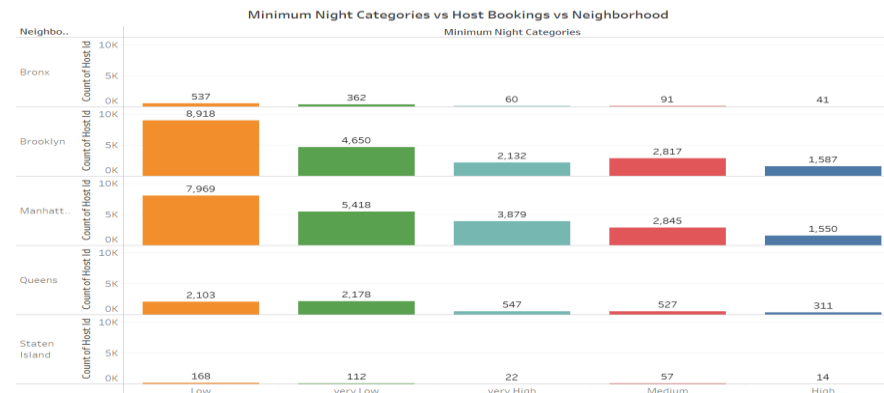
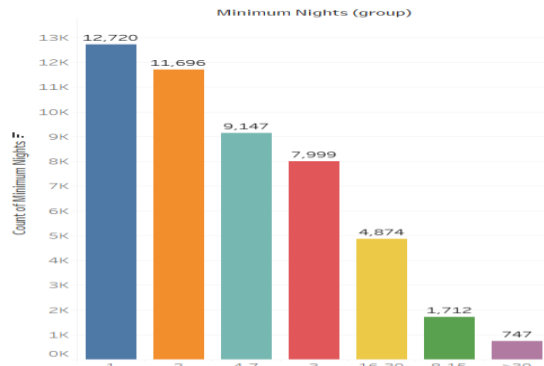
Minimum Night Categories vs Number of Reviews



Minimum Night Categor...

- High
- Low
- Medium
- very High
- very Low

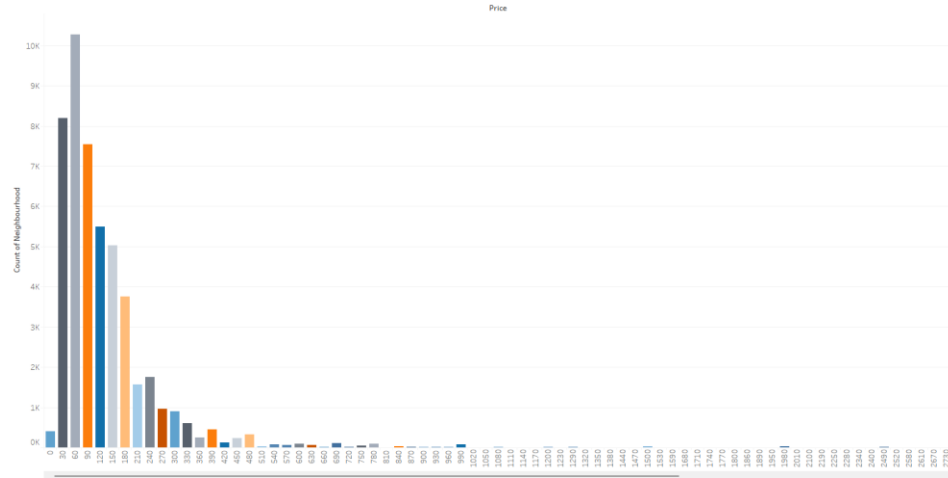
Minimum night group count



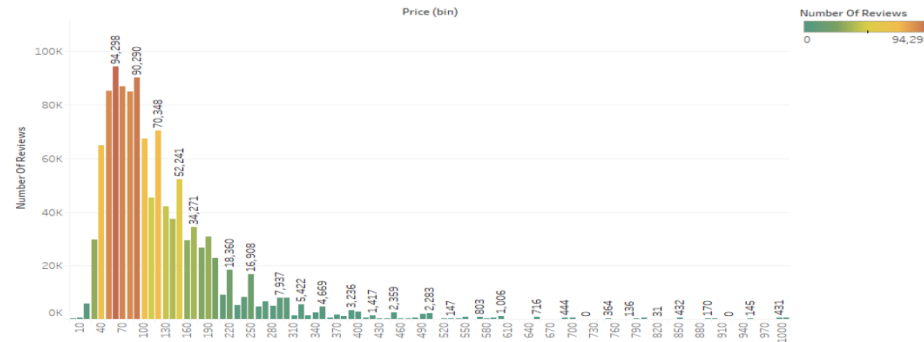
- Lower category (1-6 days) minimum nights take up 66% of the bookings, suggesting a preference among customers for shorter stays.
- Reviews for lower category minimum nights is more
- There is a noticeable surge in bookings at the 30-day mark (13% for very high category), indicating a trend of customers renting on a monthly basis.
- Manhattan and Brooklyn stand out with a higher number of 30-day bookings compared to other areas. Possible explanations include tourists opting for extended stays or mid-level employees choosing budget-friendly options for company visits.

Analysis – Price

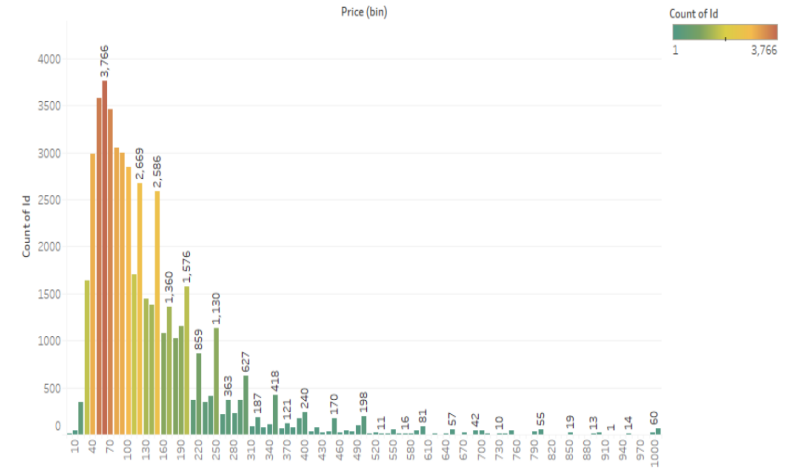
Price Distribution VS Count of Neighbourhood



Price Variation wrt Reviews



Preferred Price By customers



Pricing Preference:

The analysis considers two parameters: volume of bookings in a price range and the number of reviews in that range.

Both graphs indicate that the most favorable price range for customers is \$40 - \$190. This range aligns with the majority of bookings and also receives a significant number of positive reviews, indicating high customer satisfaction. Recommendation:

Expansion and Acquisition Strategy:

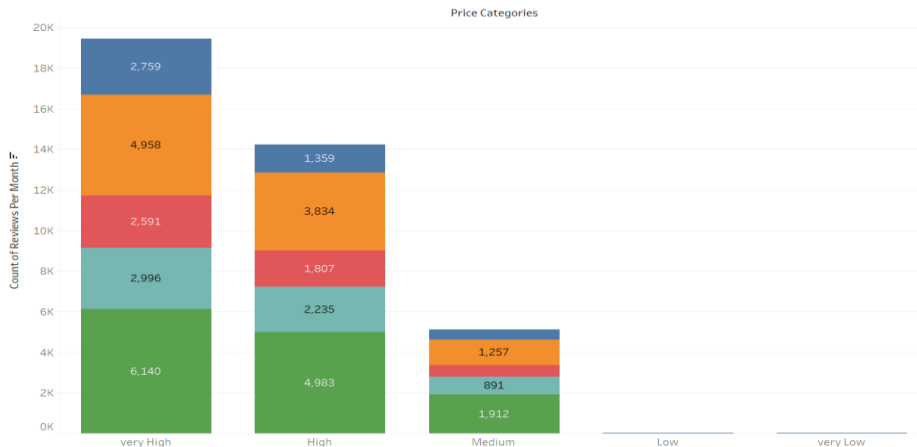
It is recommended to focus on new acquisitions and expansion within the price range of \$40 - \$190.

This range demonstrates a balance between attracting a substantial volume of customer traffic and ensuring high customer satisfaction.

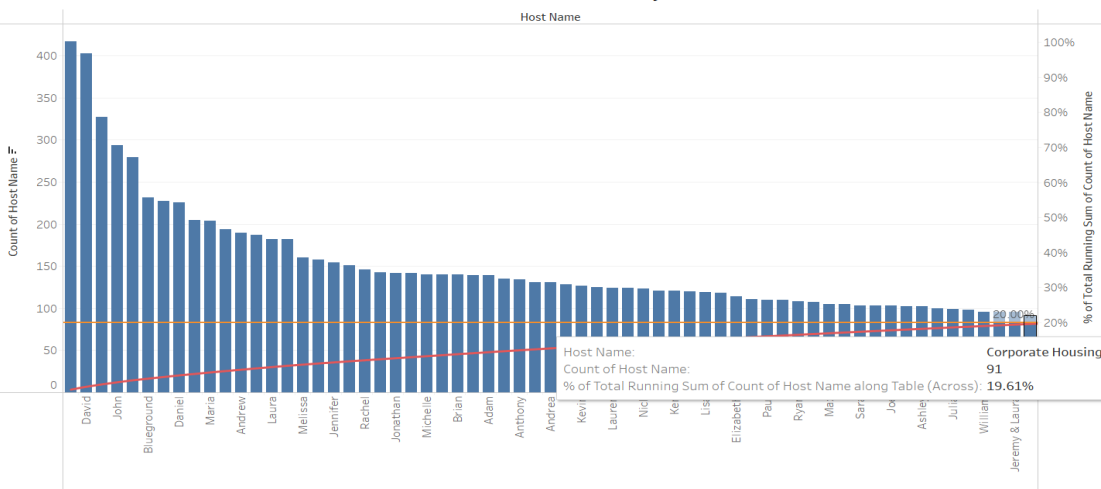
Investing in this price range is likely to yield positive results in terms of both booking volume and customer experience, contributing to overall business success.

Insights – Remaining Features

Analysis - Availability Categories vs Price Categories vs No. of Reviews per month



Pareto Chart for Host Analysis



- If the combination of availability and price is very high, reviews per month will be low on average.
- Very high availability and very low price are likely to get more reviews.
- Lower categories for availability and price will get lesser reviews.
- Top 91 hosts make up only 20% of the bookings

RECOMMENDATIONS

Neighborhood

01

- 85% of the listings are in Manhattan and Brooklyn
- Average pricing is higher than average in only 6 neighborhoods
- Manhattan and Brooklyn have more expensive offerings
- 346 neighborhoods make up 80% of the entire neighborhoods

Room Types

02

- Private rooms and Entire home/apt take up 97% of the room types
- Manhattan and Brooklyn have more Pvt rooms and Entire home/apt
- Initiate targeted promotional campaigns and offer discounts to boost bookings for shared rooms by highlighting affordability.
- Prioritize acquiring more hosts and listings offering monthly rental durations, especially in Manhattan and Brooklyn where there is a higher demand for such extended stays.
- Tap into the market of customers requiring short-term rentals for quarantine purposes. Consider offering weekly or bi-weekly rentals to accommodate individuals stranded in NYC

Minimum Night Categories

03

- Lower category minimum nights take up 66% of the bookings
- Reviews for lower category minimum nights is more
- Airbnb should promote lowering of minimum nights bookings to increase its revenue

04

Additional Observations

- Top 91 hosts make up only 20% of the bookings so all guests should be given equal importance
- Plan new acquisitions in the price range of \$40 - \$190. This range, identified through analysis of both booking volume and customer reviews, signifies a sweet spot that attracts substantial customer traffic while ensuring high satisfaction levels.
- Consider acquiring more 'private rooms' in Manhattan and Brooklyn, and 'entire homes' in Bronx and Queens. This targeted approach aims to meet the specific preferences and demands of customers in different regions.
- With an average price of \$124 and potential for growth, prioritize expansion efforts in Brooklyn. This recommendation takes into account the existing saturation in Manhattan and identifies Brooklyn as a viable location for acquiring new listings and accommodating diverse customer preferences.
- Increase acquisitions in coastal regions to diversify the property portfolio and attract customers seeking accommodations with scenic views or proximity to waterfronts.

APPENDIX – Data Methodology

Data Analysis

Data loading and analysis using Python

EDA

Univariate, Bivariate Analysis and Correlations identified using Python

Binning of continuous variables

Categorization of numerical columns for EDA

Missing Value Analysis

Missing value analysis using Python

Data Visualizations

Used Tableau for building graphs for analysis and identifying relationships and correlations between variables



APPENDIX – Data Methodology Contd.

Column	Description
id	listing ID
name	name of the listing
host_id	host ID
host_name	name of the host
neighbourhood_group	location
neighbourhood	area
latitude	latitude coordinates
longitude	longitude coordinates
room_type	listing space type
price	
minimum_nights	amount of nights minimum
number_of_reviews	number of reviews
last_review	latest review
reviews_per_month	number of reviews per month
calculated_host_listings_count	amount of listing per host
availability_365	number of days when listing is available for booking

Dataset Description

Categorical Variables:

- room_type
- neighbourhood_group
- neighbourhood

Continuous Variables(Numerical):

- Price
- minimum_nights
- number_of_reviews
- reviews_per_month
- calculated_host_listings_count
- availability_365
- Continuous Variables could be binned in to groups too

Location Variables:

- latitude
- longitude

Time Variable:

- last_review

Variable Categories

THANK YOU

