



IIITB & UpGrad

# Impact of Customer Preferences on Airbnb strategy

For Data Analytics Team

# Agenda

- Objective
- Background
- Key Findings
  - Host Analysis
  - Customer Analysis
- Recommendation
- Appendix
  - Data Sources
  - Data Methodology
  - Data Assumptions

# Objective

- Analyze properties in given 5 neighborhood groups.
- Analyze customer preferences.
- Provide data recommendations in order to improve the analysis quality.
- Provide recommendations in order to increase revenue.

# Background

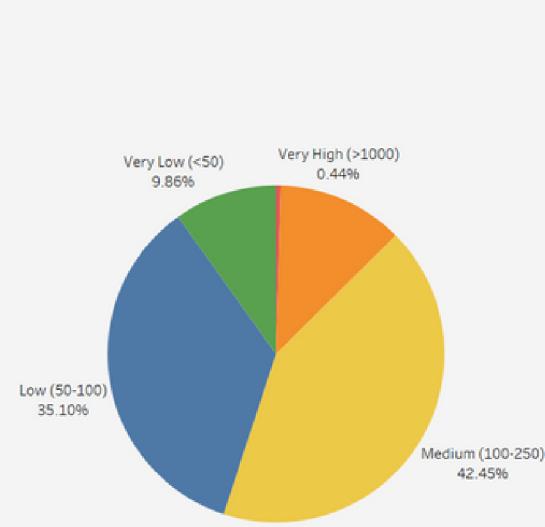
- The hospitality industry, including has significantly impacted by the COVID-19 pandemic.
- Governments around the world implemented strict lockdowns and travel restrictions, which resulted in a sharp decrease in travel demand and bookings. Due to which revenue of the Airbnb New York has declined significantly.

# Key Findings

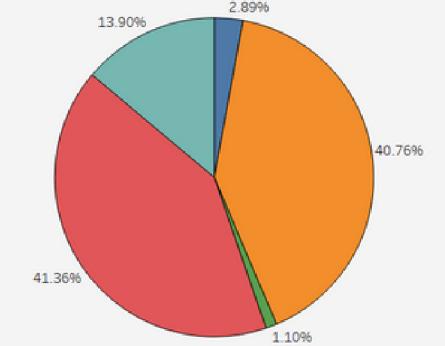
# Host Analysis

- Around 97% of properties listed are in Manhattan, Brooklyn and Queens.
- Majority of the hosts avail their properties at a price range of \$100-\$250.
- 3928 properties have been inactive for 1-4 years.

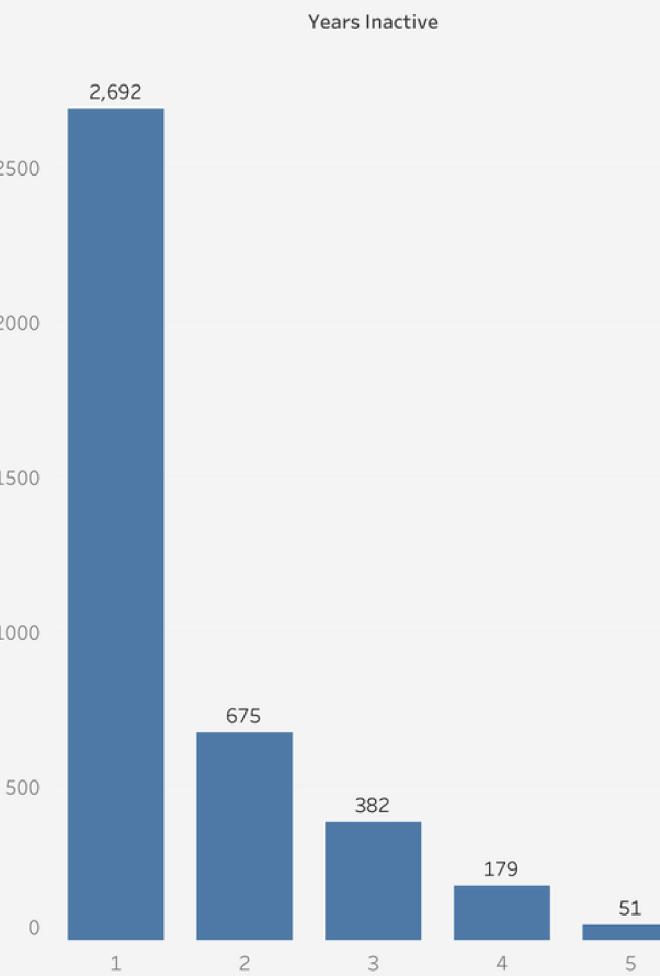
Prices in New York



Percentage of Listings by Neighbourhood Groups

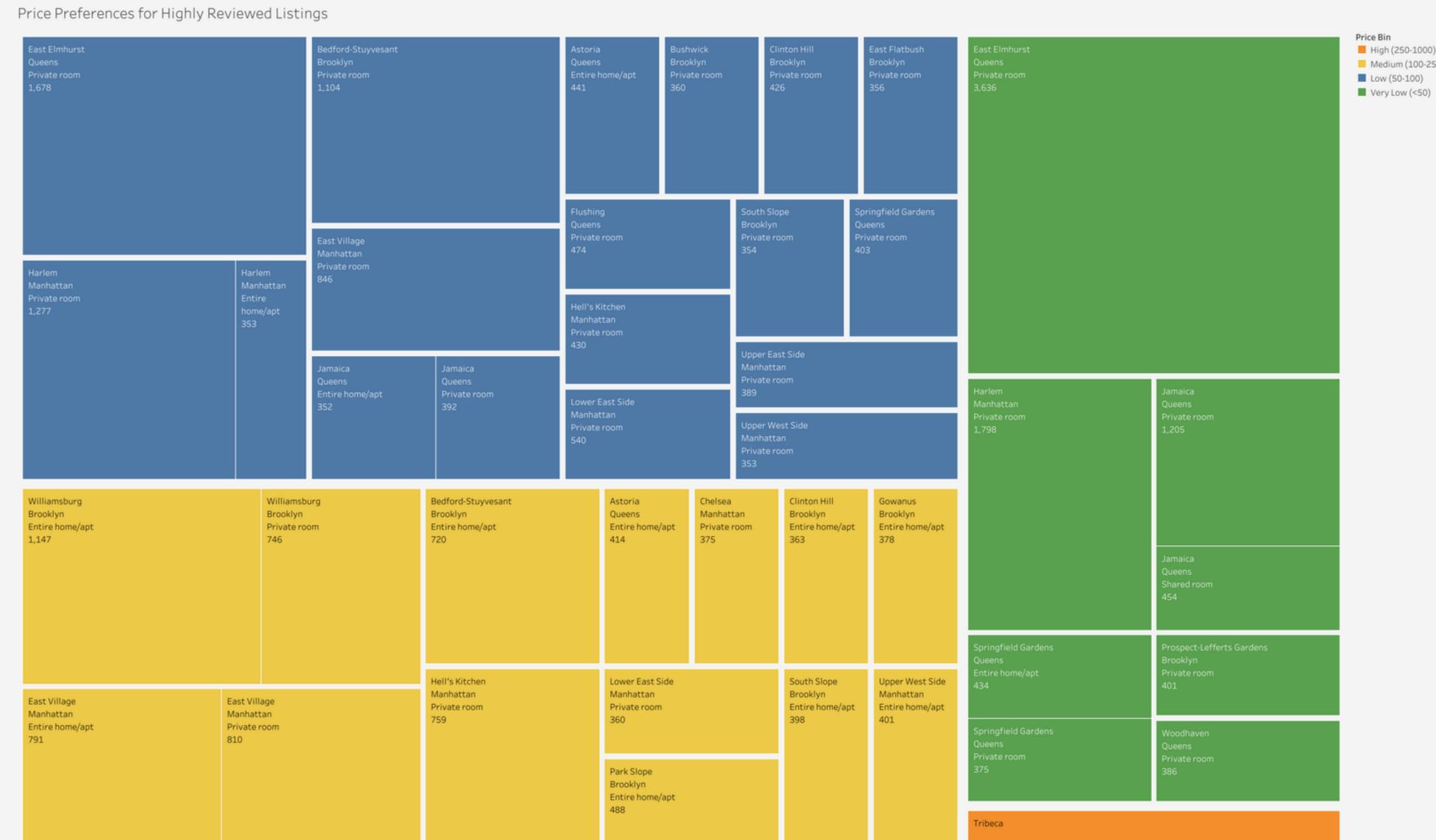


Inactive Properties



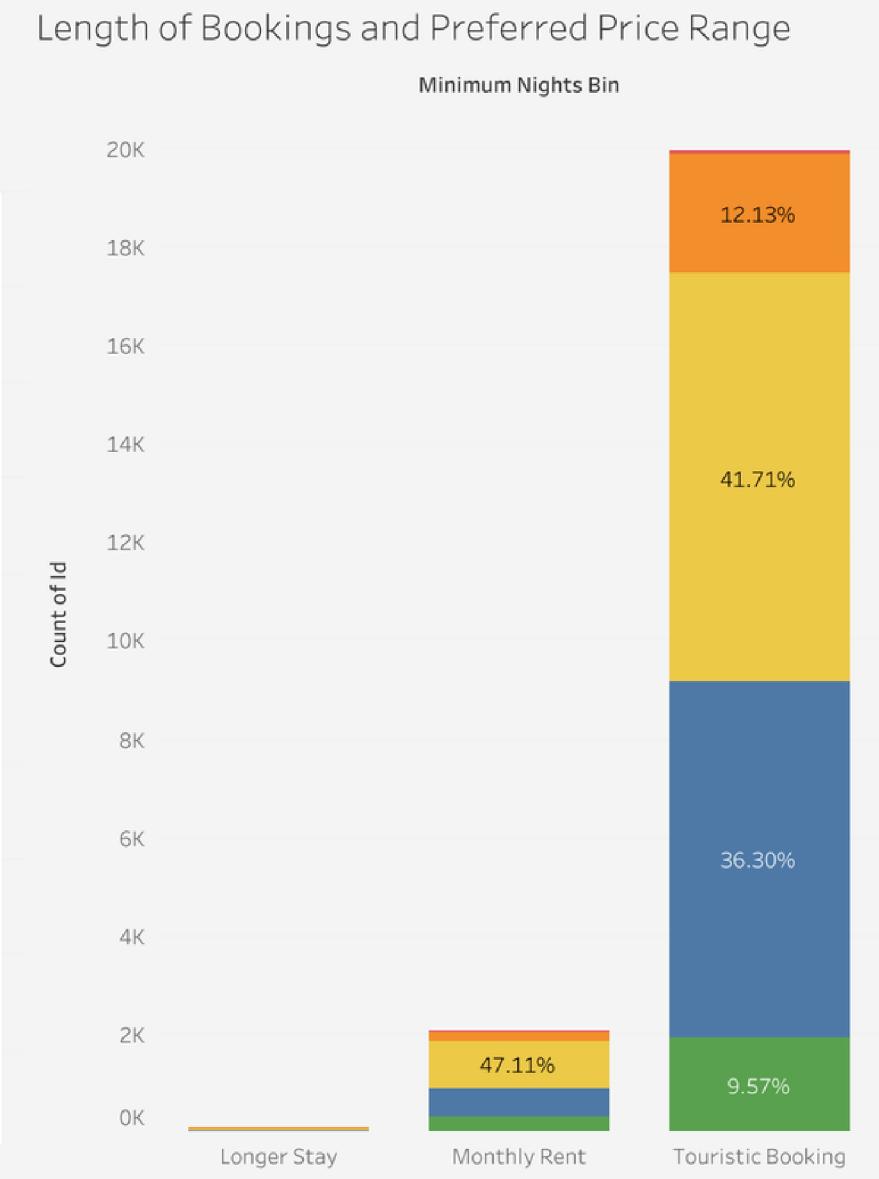
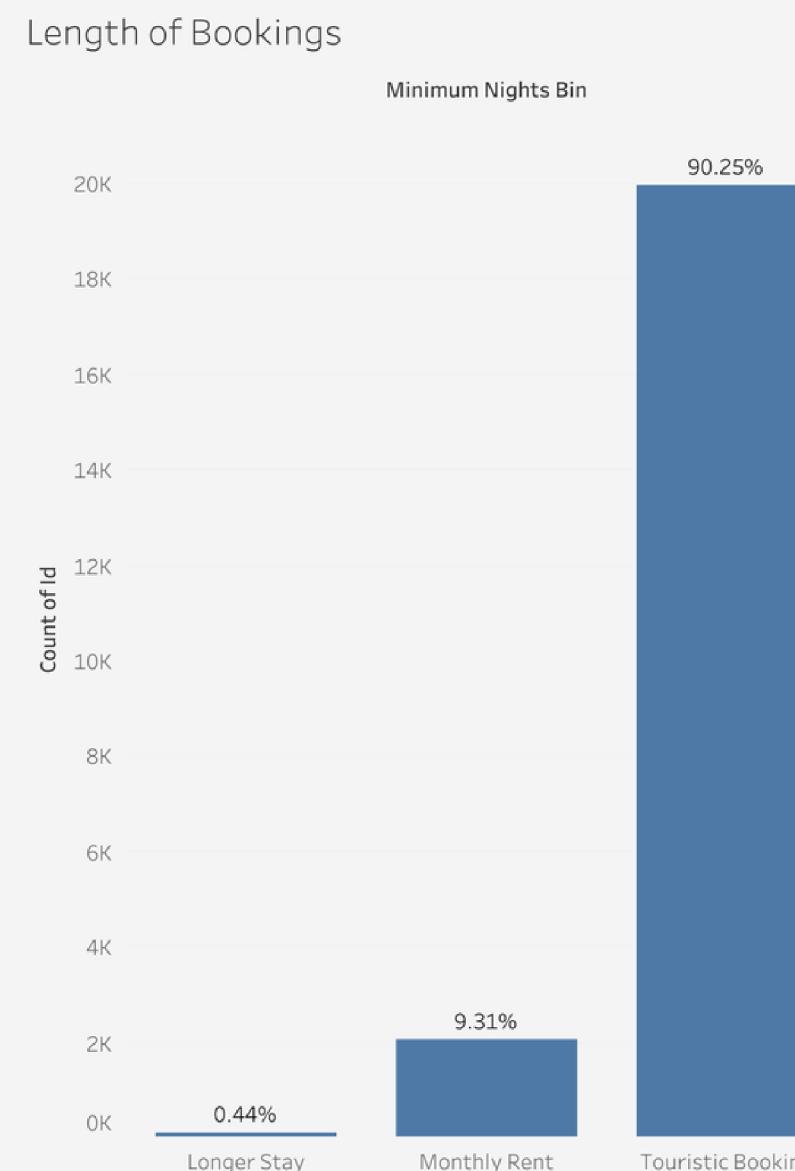
# Customer Preferences

- The preferred room types are : **Entire Home and Apartment and Private Rooms.**
- The price range preferred for **Entire Home** is \$100-\$250 per night. The price preferred for a **Private Room** is less than \$50.
- Shared room** is comparatively less preferred choice, however, such rooms booked at price of less than \$50 have high reviews in Queens.



# Customer Preferences

- 90.3% of the active properties were booked for 1-8 nights.
- These customers can be categorized as tourists.
- Price range preferred is \$50-\$250 per night.



Price Bin

- Very High (>1000)
- High (250-1000)
- Medium (100-250)
- Low (50-100)
- Very Low (<50)

# Recommendations

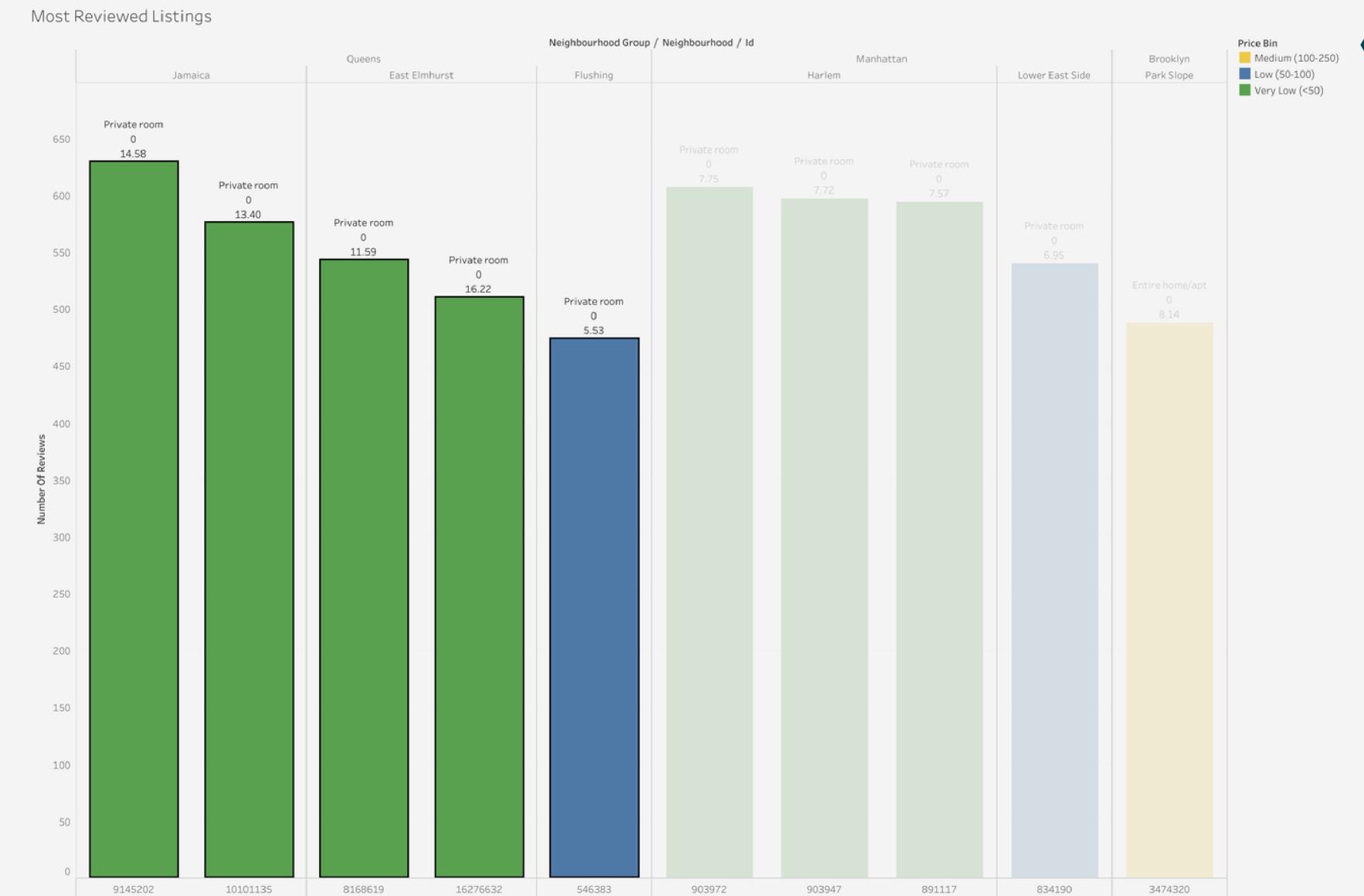
# **Additional data from hosts and customers will improve analysis quality and further recommendations**

- **Customers' ratings** of different property features will help improve customers satisfaction.
- **Hosts' ratings** of Airbnb services will enrich company's relationship with the hosts, which will attract more of them.
- Occasional **surveys** can be conducted for the platform user experience improvement.
- Survey needs to be done in order to understand the root cause of hosts' inactivity.



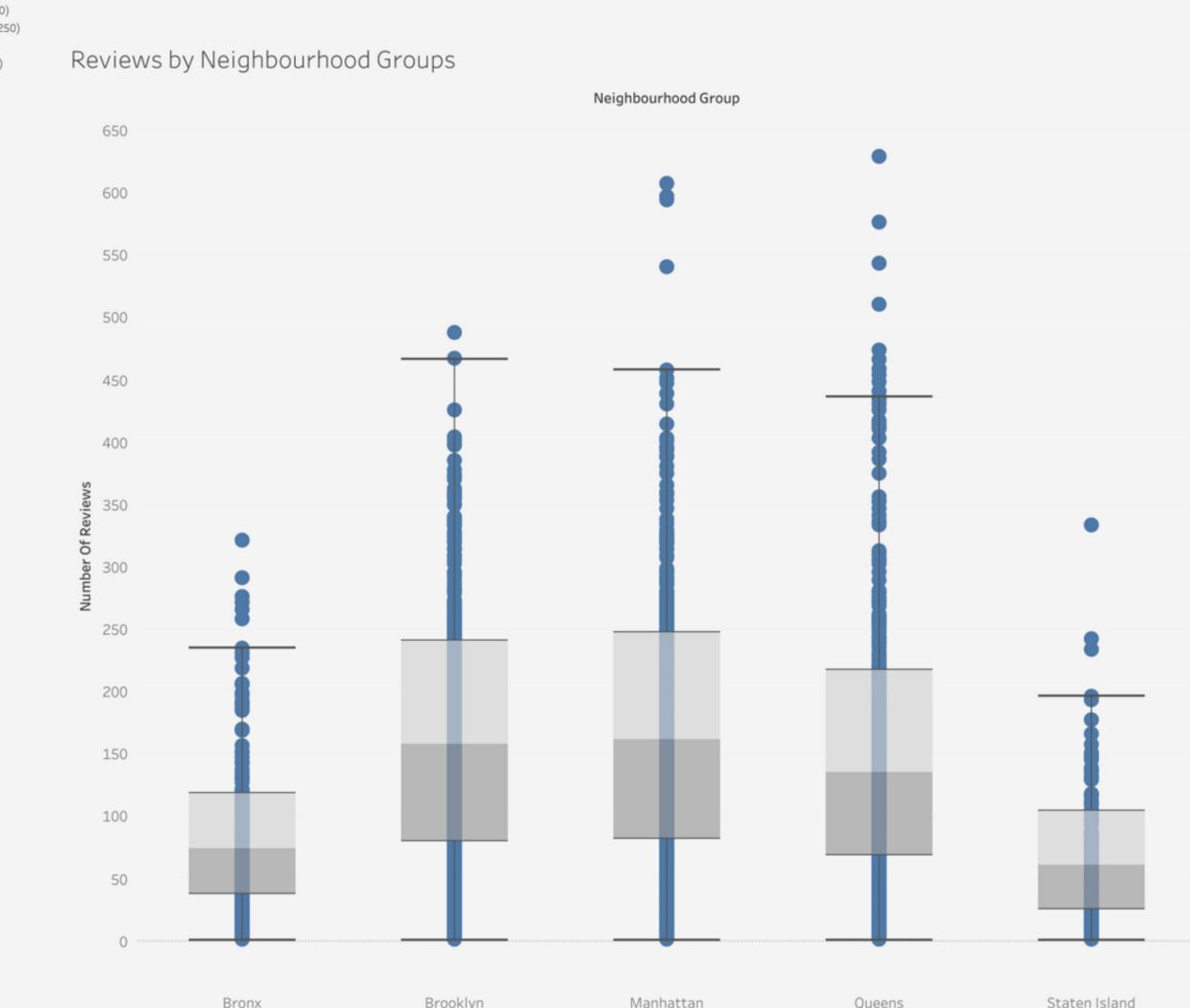
# Acquiring more hosts with low priced private rooms in Queens can increase revenue.

- **9 out of 10** most reviewed active listings are Private Rooms.
- **5 out of those 10** are in Queens.
- Customers prefer to stay in private rooms priced below \$100.

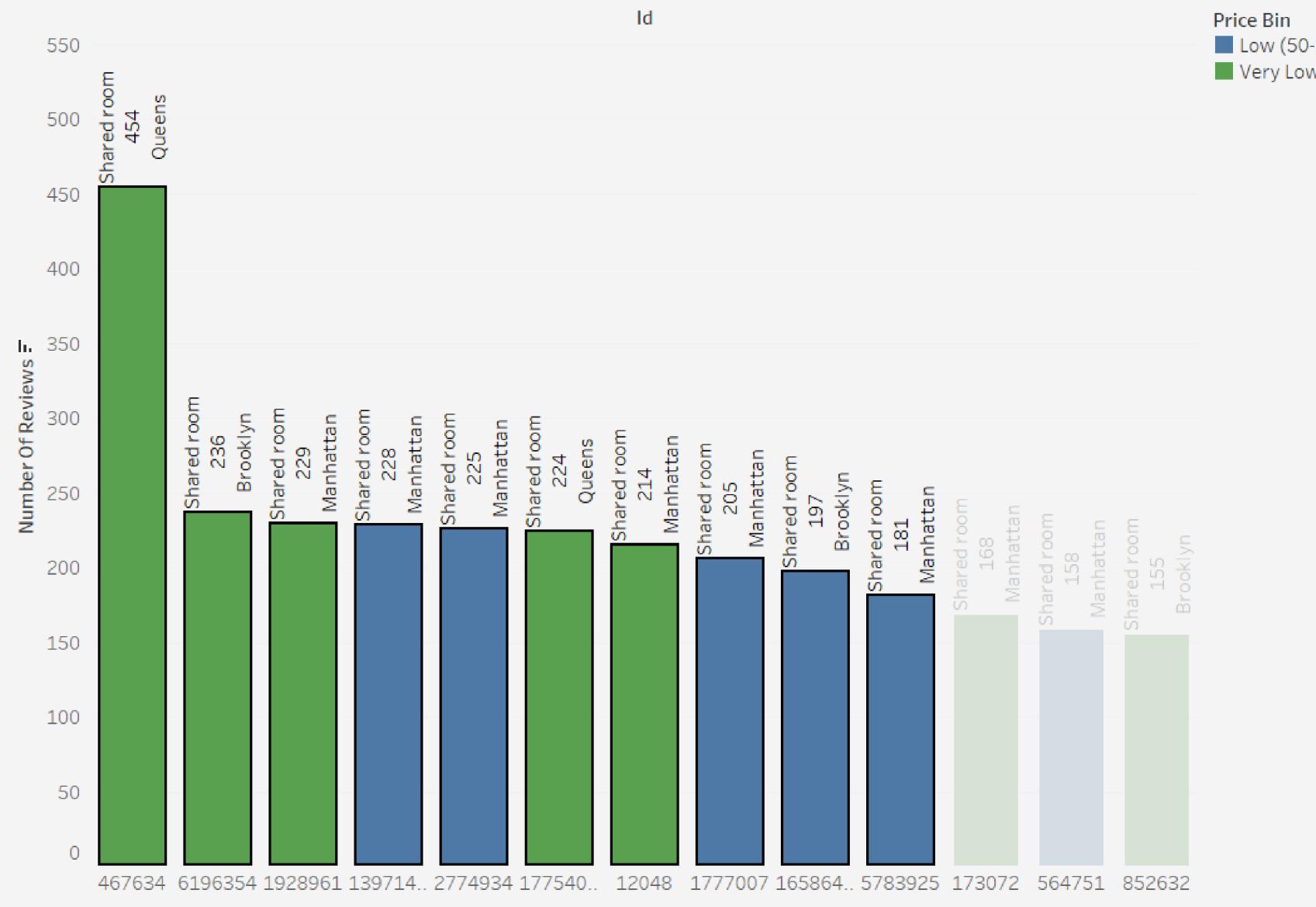


# New Bronx properties can attract more customers

- Customers are willing to pay up to **\$250** for Bronx listings.
- Bronx receives **decent amount of reviews**, considering it accounts for only **2.9%** of total listing counts.
- Bronx is fairly **close to Manhattan**, and most of the bookings are touristic (up to 8 days).



# Providing shared rooms as private rooms will attract younger tourists



- Despite being 2.5% of total room types, shared rooms have received significantly **high number of reviews**.
- The price range preferred is less than \$100.
- Post pandemic guests might prefer to stay alone.

# Appendix

# Data Sources

- The data used in the analysis was provided by Airbnb.
- The data contained 16 columns and 48,896 rows.

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

**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

# Data Methodology

- Data cleaning steps:
  - Removed rows with values such as “Not available”, “No longer available”, etc. in [name].
  - Removed rows with [price] of 0, 9999, 10000.
  - Removed rows with [minimum\_nights] more than 365.
  - Removed rows with [availability\_365] of 0 or more than 365.
  - Removed rows with [number\_of\_reviews] of 0 (considering them inactive).
  - Brought [last\_review] to one format (mm/dd/yyyy).
- Data visualizations were made in Tableau.

# Data Assumptions

- New columns were derived based on the assumptions made:
  - [years\_inactive] assuming [last\_review] the day of the last booking made in a listing
  - [price\_bin] in order to categorize [price]:
    - Very Low (< \$50)
    - Low (\$50-\$100)
    - Medium (\$100-\$250)
    - High (\$250-\$1000)
    - Very High (> \$1000)
  - [minimum\_nights\_bin] in order to categorize customers by the length of their stay:
    - Touristic Booking (1-8 nights) assuming that only tourists will stay in a property for such a short time.
    - Monthly Rent (9-31 nights) assuming that people stay in such properties for living in New York short term.
    - Longer Stay (32-365 nights) assuming that people rent such properties for living in New York.