

# AIRBNB STORYTELLING CASE STUDY- PPT 1

By  
Vaibhav Vijay  
Aaradhya



# OBJECTIVE

- Airbnb is an online platform that allows people to rent out their unused accommodations.
- During the COVID-19 pandemic, Airbnb experienced a significant decline in revenue.
- With travel resuming, Airbnb is focused on rebuilding its business and is prepared to serve customers once again.

# OVERVIEW

- In recent months, Airbnb has experienced a significant drop in revenue.
- With travel restrictions easing and people resuming travel, Airbnb aims to be fully prepared for this shift.
- To support this effort, an analysis has been conducted on a dataset of various Airbnb listings in New York.



# DATA CLEANING AND ANALYSIS

- **Tools Used:** Jupyter Notebook (Preprocessing), Tableau (Visualization)
- **Dataset:** AB\_NYC\_2019.csv (48,895 rows, 16 columns)
- **Steps Taken:**
  - Removed irrelevant columns (ID, Name, Last Review)
  - Checked for duplicate entries (None found)
  - Handled missing values (Dropped 'name' due to minimal impact)
  - Identified and analyzed outliers



# DATA ANALYSIS & VISUALIZATION

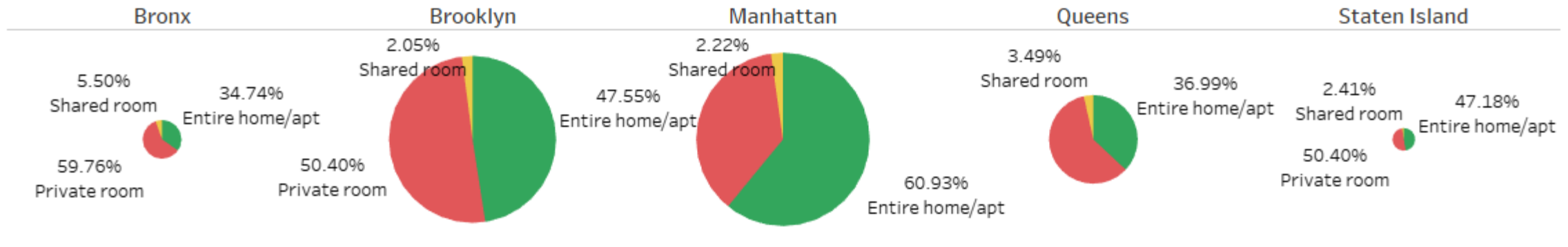
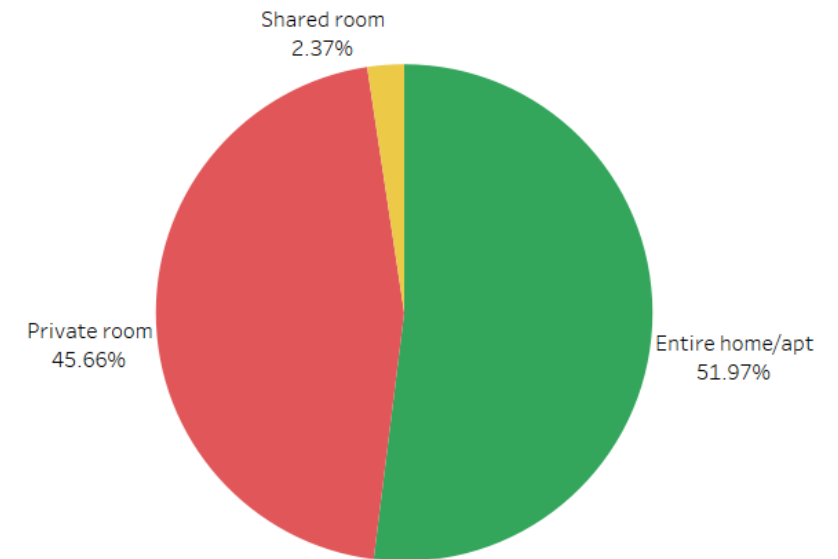
## Key Insights and Visualizations

- 1. Top 10 Hosts Analysis** – Treemap showing the most active hosts.
- 2. Room Type Preferences by Neighborhood Group** – Pie chart showing distribution.
- 3. Price Variance by Neighborhood** – Box plot showing price fluctuations.
- 4. Average Price by Neighborhood** – Bubble chart representing price trends.
- 5. Customer Booking Trends by Minimum Nights** – Binned histogram showing booking durations.
- 6. Most Popular Neighborhoods** – Bar chart of top 20 reviewed locations.
- 7. Neighborhood vs Availability** – Dual-axis chart comparing availability and pricing.



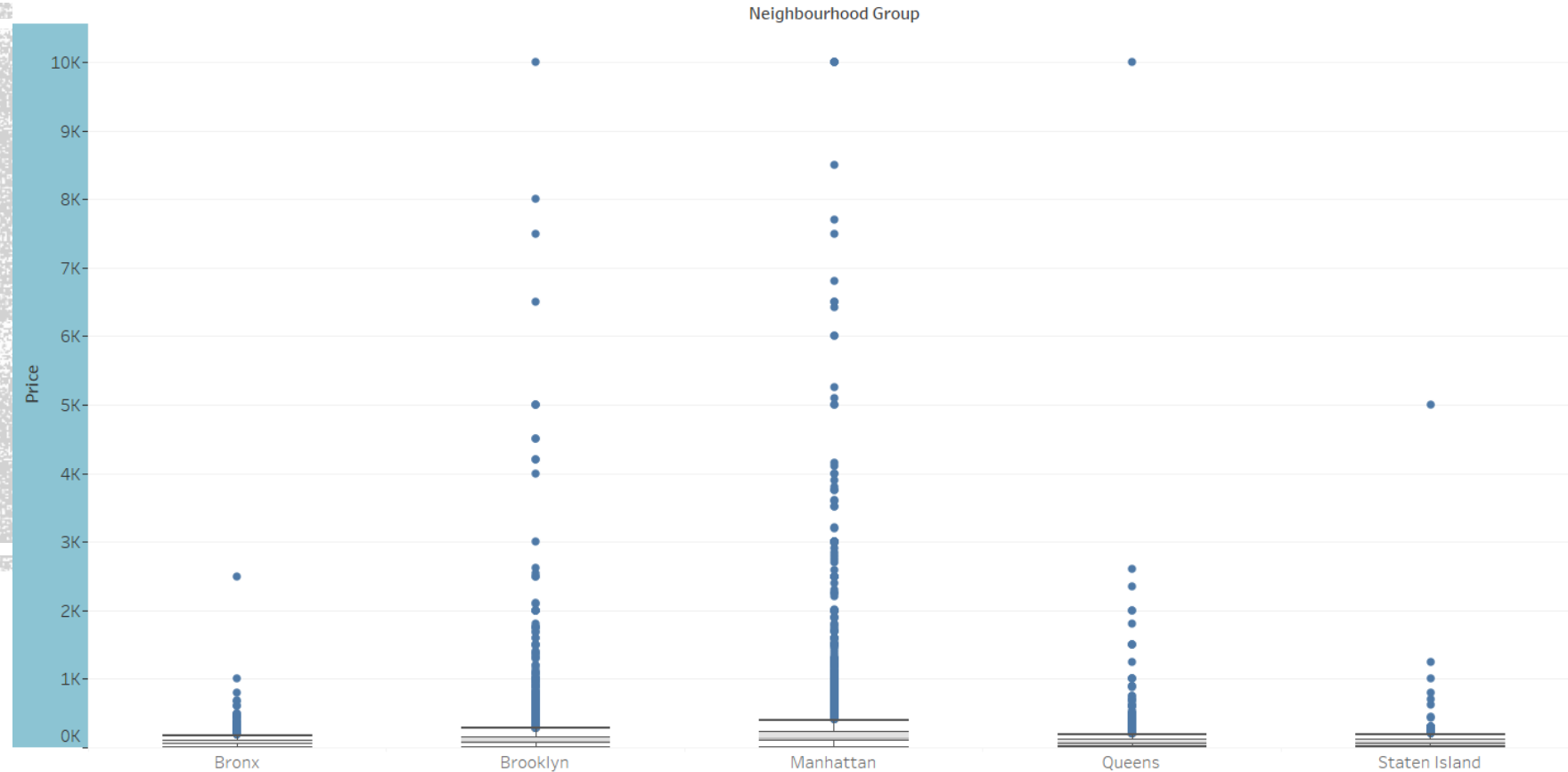
# Room Type Preferences by Neighborhood Group

- Manhattan & Brooklyn are the most popular areas for bookings.
- **Manhattan:** Highest number of entire home/apartment listings (60.93%).
- **Bronx:** Dominates private room listings (59.76%).
- **Shared rooms** are available in very limited numbers across neighborhoods.



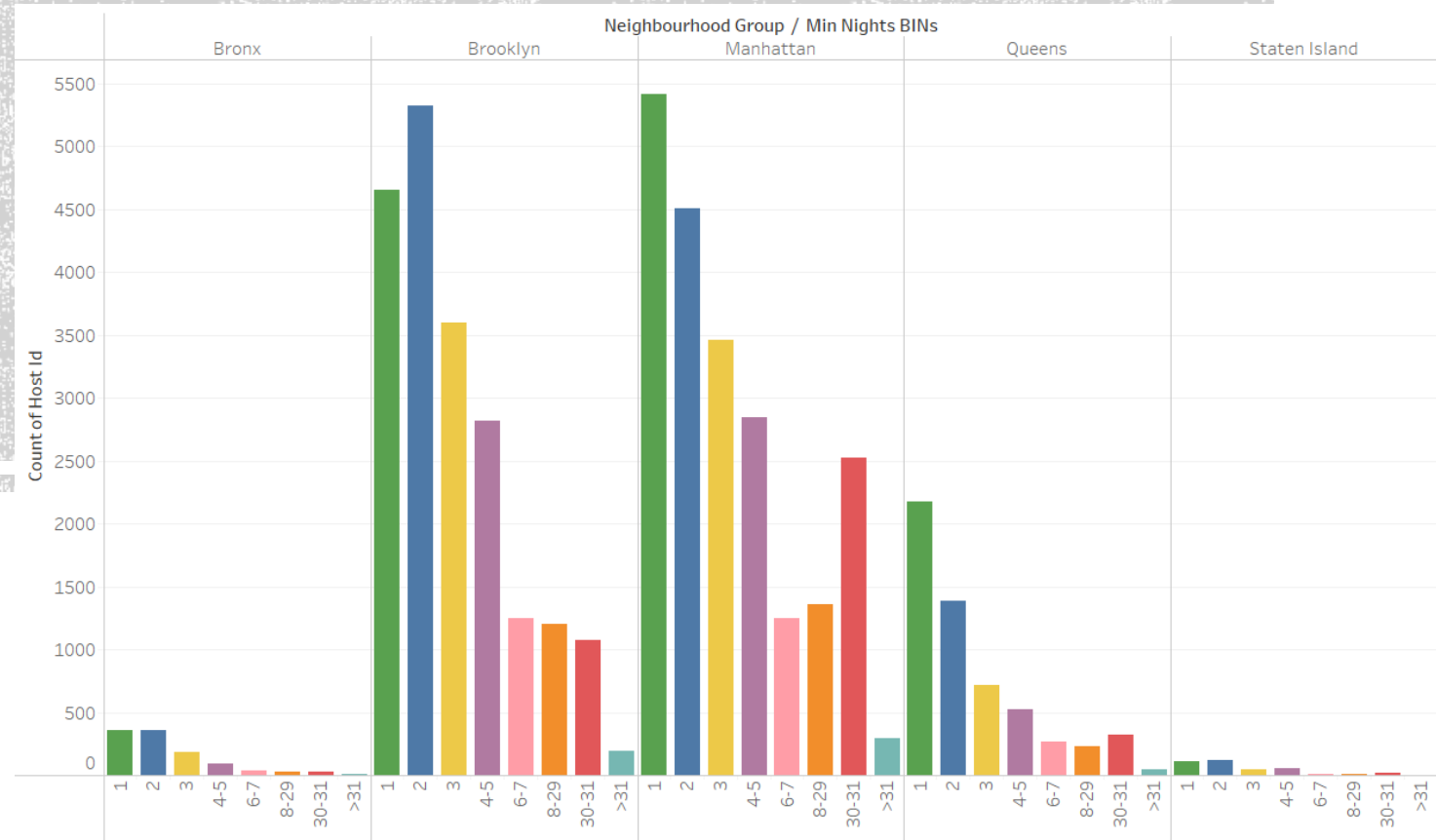
# Neighborhood Price Analysis

- Price outliers are primarily found in Brooklyn and Manhattan.
- Manhattan exhibits the widest price range for listings.
- The Bronx has the most affordable listings among all neighborhoods.
- The median price across all neighborhood groups falls between \$80 and \$300.
- Prices show a strong positive skew, with the median positioned close to the lower quartile, along with several outliers, as indicated in the boxplot.



# Minimum Nights wrt Customer Booking

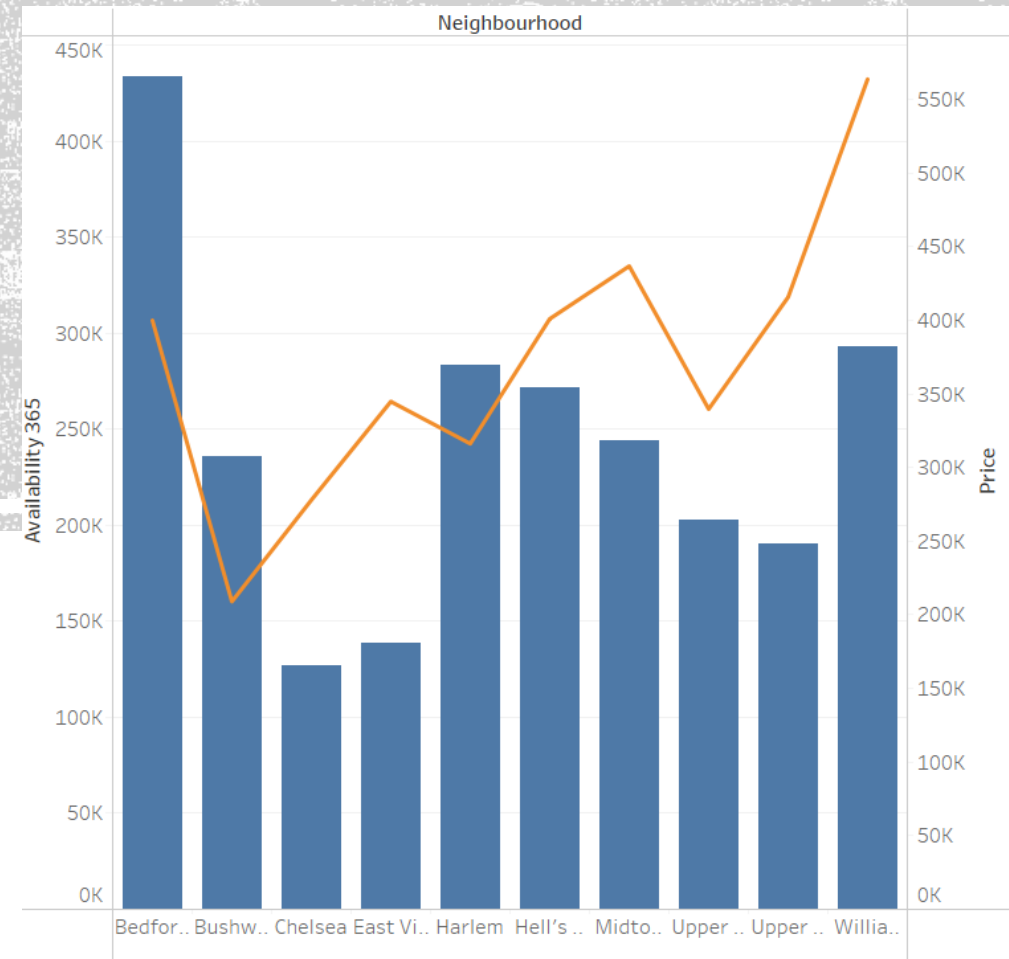
- Most bookings are for stays between **1-5 nights**.
- A noticeable spike occurs at **30 days**, indicating monthly rental trends.
- **Manhattan & Queens** have a higher number of 30-day bookings, possibly due to:
  - Tourists opting for extended stays.
  - Mid-level employees choosing budget-friendly accommodations for work trips.





# Neighbourhood Prices vs Availability

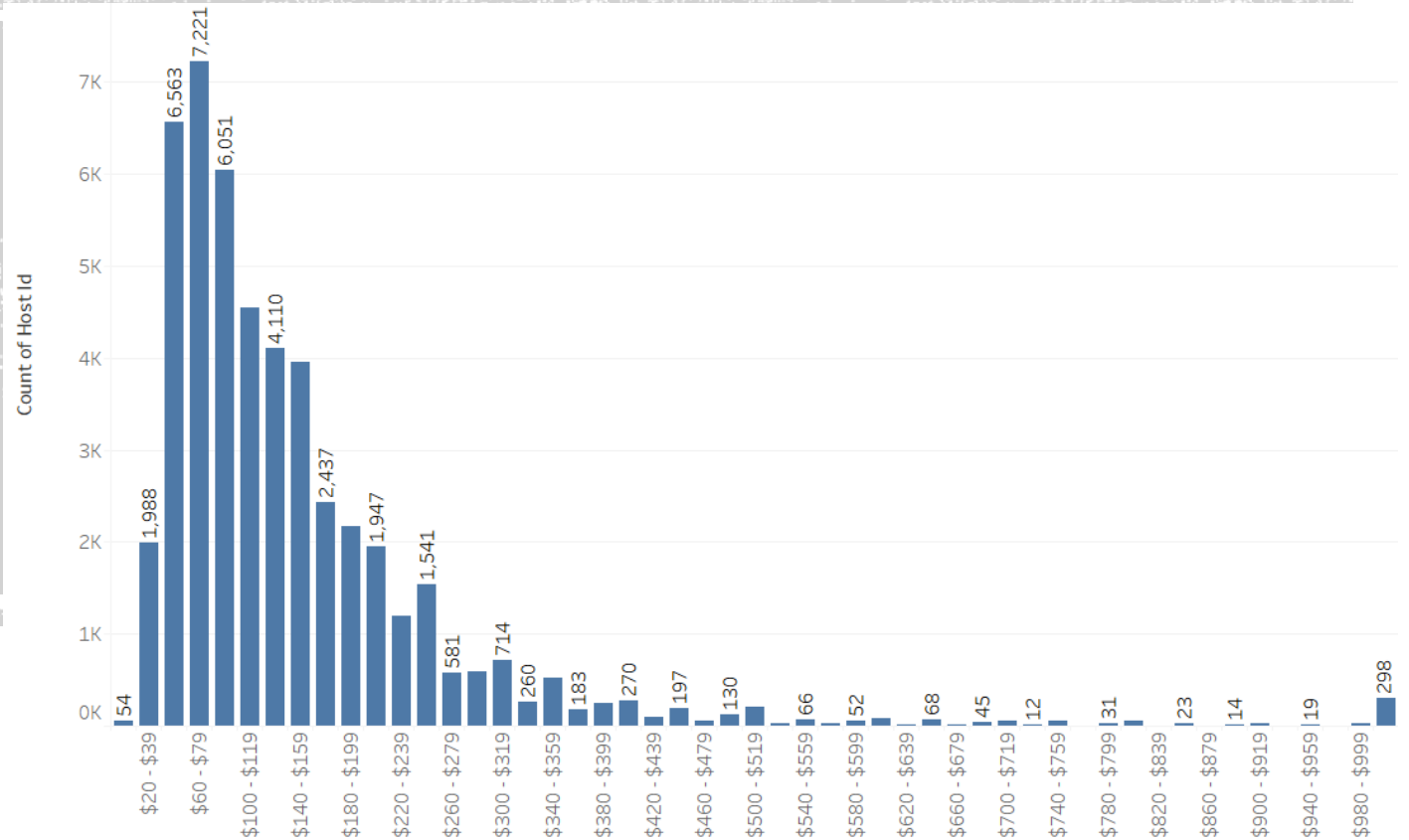
- **Bedford-Stuyvesant** has the highest availability with relatively lower prices.
- **Harlem** follows a similar trend.
- **Chelsea** has limited availability but is relatively expensive.
- **Williamsburg** has high prices with moderate availability.





# Preferred Price Ranges

- Analyzed pricing preferences based on booking volume.
- Most **preferred price range** is between **\$60 and \$200**.



# Price variation as per Room Type/Neighbourhood

- **Entire home/apartment:** Most expensive in **Manhattan** (~\$250).

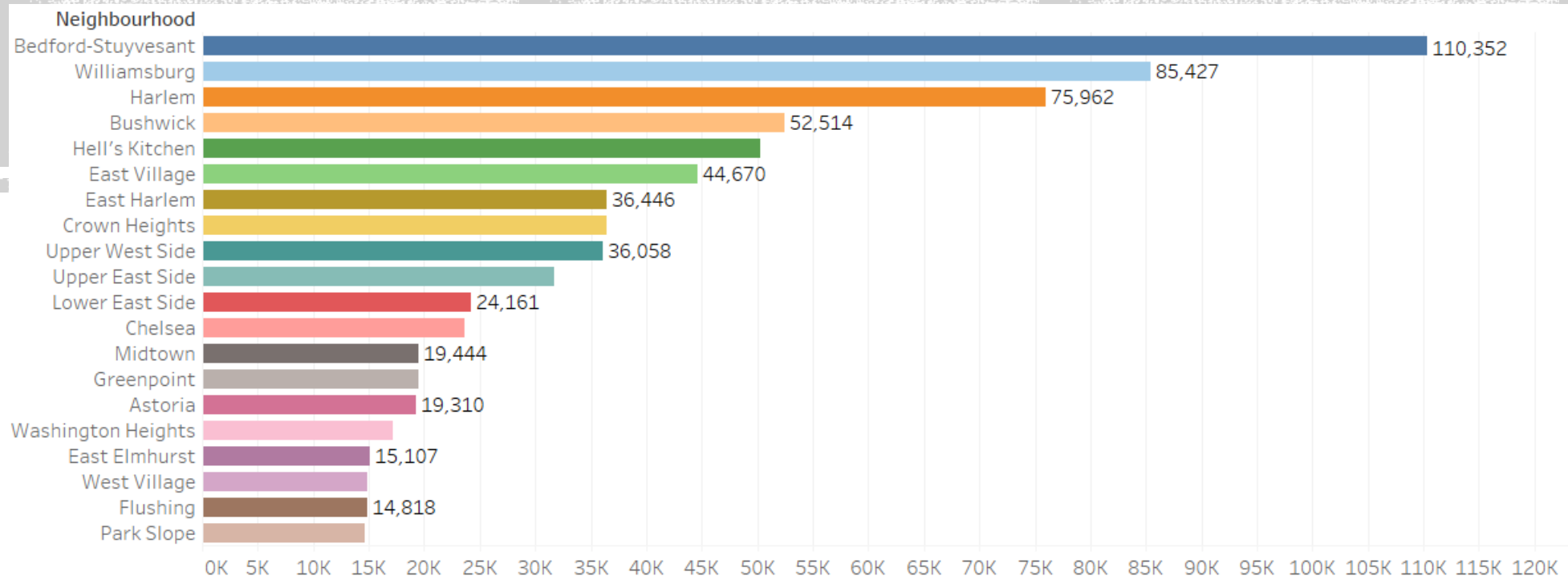
- **Private rooms:** Highest average prices in **Manhattan & Brooklyn**.

- **Shared rooms:** Most affordable in **Brooklyn** (~\$50.5).

		Neighbourhood Group				
Room Type		Bronx	Brooklyn	Manhattan	Queens	Staten Island
Entire home/apt		127.5	178.3	249.2	147.1	173.8
Private room		66.8	76.5	116.8	71.8	62.3
Shared room		59.8	50.5	89.0	69.0	57.4

# Popular Neighborhoods

- **Bedford-Stuyvesant (Brooklyn):** Most popular, with **110,352 reviews**.
- **Williamsburg** follows closely in Brooklyn.
- **Harlem (Manhattan):** Highest number of reviews, followed by **Hell's Kitchen**.
- **Higher number of reviews suggests greater customer satisfaction** in these areas.



# APPENDIX- DATA SOURCES

- Data File- [AB NYC 2019.csv](#)
- Prepared Data File- [airbnb final.csv](#)

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



# APPENDIX- DATA METHODOLOGY

- Data methodology File- [Airbnb Case Study Methodology.pdf](#)
- **Step 1:** Data Cleaning & Preparation (Jupyter Notebook).
- **Step 2:** Data Analysis & Visualization (Tableau).
- **Step 3:** Additional Insights & Refinements.
- **Step 4:** Conclusions & Strategic Recommendations.



# APPENDIX- DATA ASSUMPTIONS

## 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

