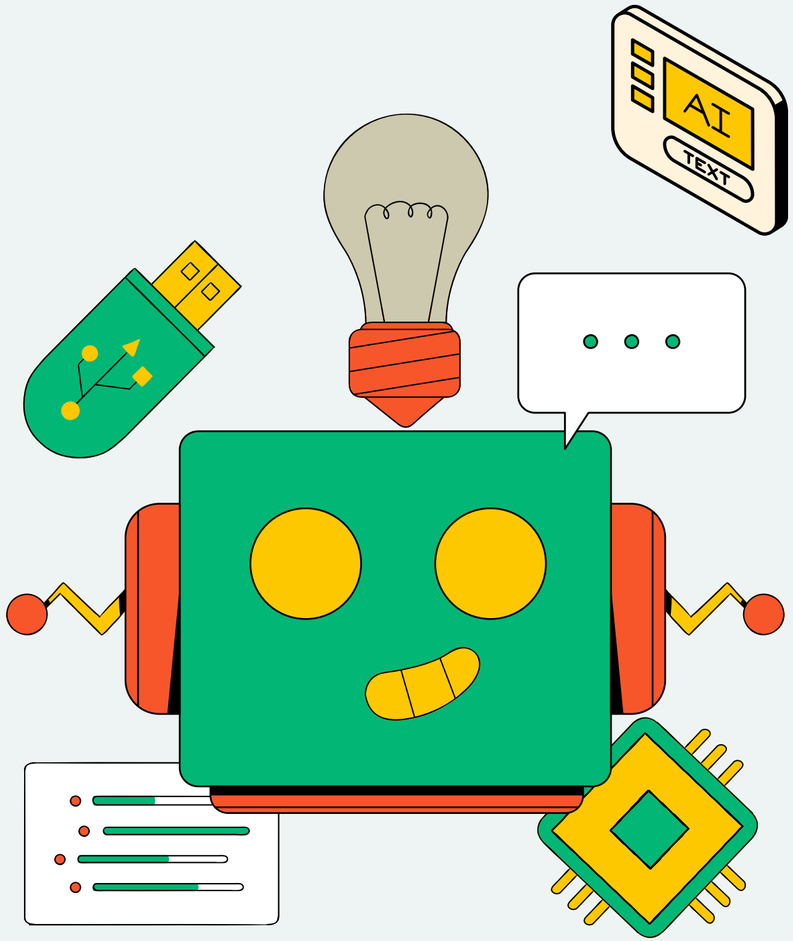
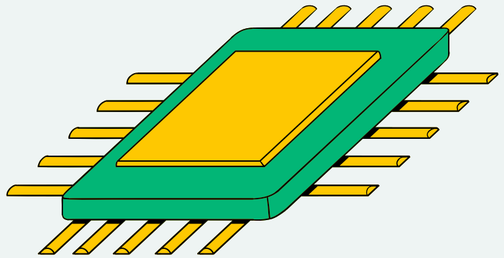
# UNLOCKING INSIGHTS: BUSINESS INTELLIGENCE ANALYSIS OF AIRBNB NYC DATA




## PRESENTATION

**PRESENTED BY: GROUP 3**

DEVISHA BHAYANI
HARSHIT AGGARWAL
HIRA SHAIKH
QI SHI
RAVNEET SINGH
TANIYA KETAN GADKARI



# PRESENTATION OUTLINE

- The description of operations.
- Project objectives.
- Importance of the research to the industry
- Data overview
- Data cleaning and preprocessing.
- Data Visualization
- Power BI Dashboards
- Results and conclusions.

# PROJECT DESCRIPTION

**This project focuses on analyzing Airbnb listings in New York City to uncover key insights that can guide hosts, investors, and stakeholders in optimizing their offerings and strategies.**

**By doing so, the project aims to inform decision-making that can enhance revenue, occupancy rates, and overall success in the competitive NYC Airbnb landscape.**

# PROJECT OBJECTIVES

- **Primary Objective:** To analyze the Airbnb NYC dataset to identify the key factors driving revenue and occupancy rates, providing actionable insights for optimizing pricing strategies, improving host performance, and maximizing profitability.

- **Secondary Objectives:** To assess neighborhood trends and room type performance, offering guidance on strategic investment opportunities and effective market positioning within the competitive NYC Airbnb landscape.

# IMPORTANCE TO THE INDUSTRY

Market Positioning

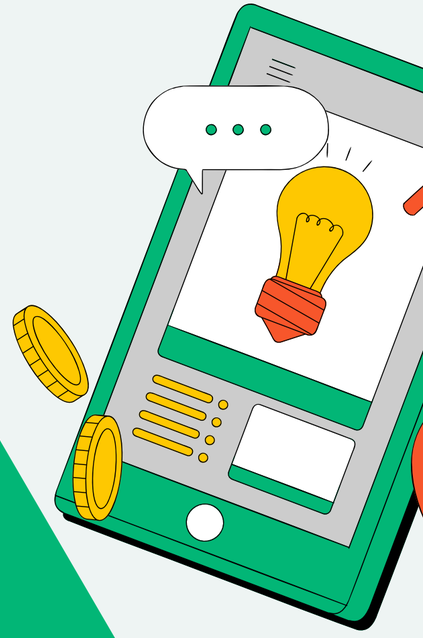
Revenue Optimization

Improved Guest Experience

Data-Driven Decision Making

Investment Strategy

Optimizing Resource Management



# DATA SOURCE

## DATA USED

**Dataset:** AB_NYC_2019.csv from Airbnb, containing listings data for New York City in 2019.

Fields include listing details such as price, availability, neighborhood, room type, and host information.

## DATA CLEANING AND PREPROCESSING

Corrected data types for key columns, e.g., price and availability as integers, reviews as decimals.

Handled missing data to ensure accuracy in analysis and visualization.

Calculated revenue and Occupancy Rate

## VISUALIZED DATA PRESENTATION

Map Visualizations: Plotted listings by geographic location, overlaid with metrics like average price and availability.

Bar & Column Charts: Showcased host performance, room type revenue, and occupancy rates.

# PRICING OPTIMIZATION


**Neighborhood Price Distribution:** High prices in premium areas like Tribeca, Fort Wadsworth, and Woodrow signal strong demand.

**Business Insight:** Hosts should optimize pricing in these high-demand neighborhoods to boost revenue.

**Strategic Focus:** Focus on premium locations for investment and price adjustments to maximize returns.

# HOST PERFORMANCE ANALYSIS:

**Listings Concentration:** Top hosts, including Sonder (NYC), Blueground, and Corporate Housing, focus on Manhattan, Brooklyn, and Queens, with 52 to 327 listings


**Guest Engagement:**

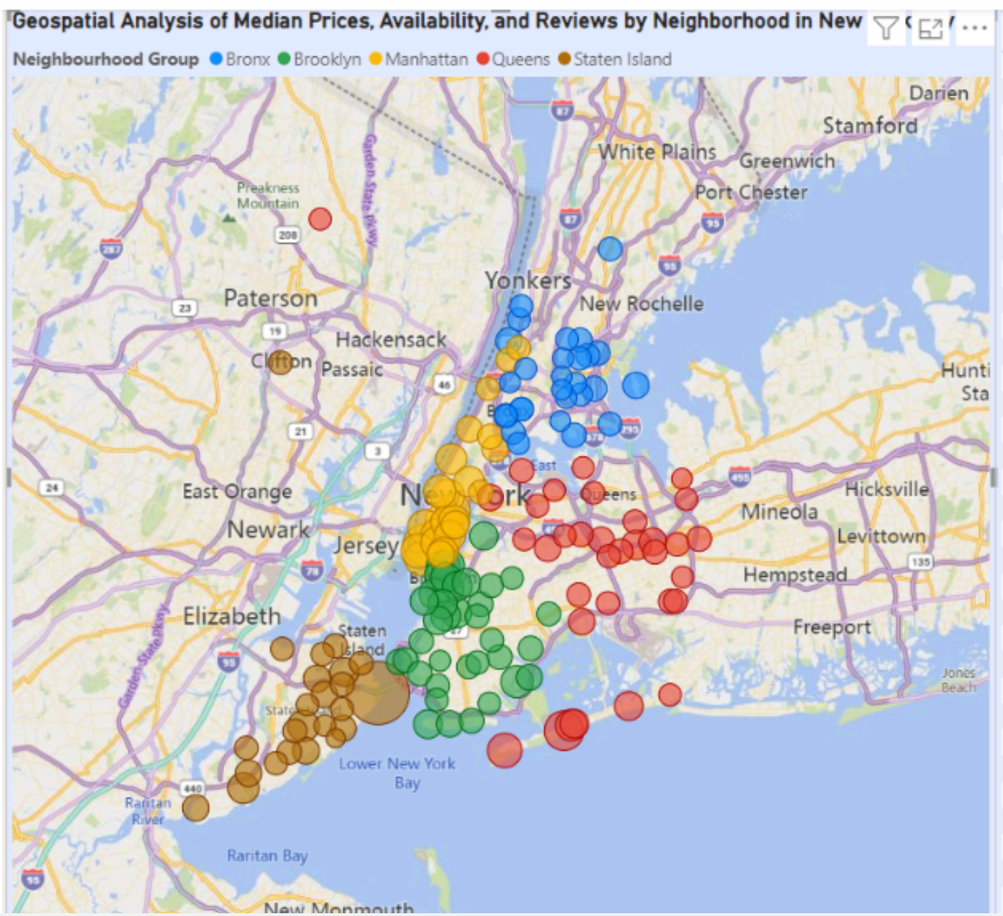
**Sonder (NYC):** 1,281 reviews for 327 listings, indicating strong guest interaction.

**Blueground:** 29 reviews for 232 listings, suggesting lower engagement or different strategies.

**Corporate Housing:** 91 listings with 417 reviews, reflecting a focus on customer service.

**Business Insight:** Key neighborhoods dominate top hosts. Smaller hosts can increase reviews by improving guest experiences.

# NEIGHBORHOOD ANALYSIS



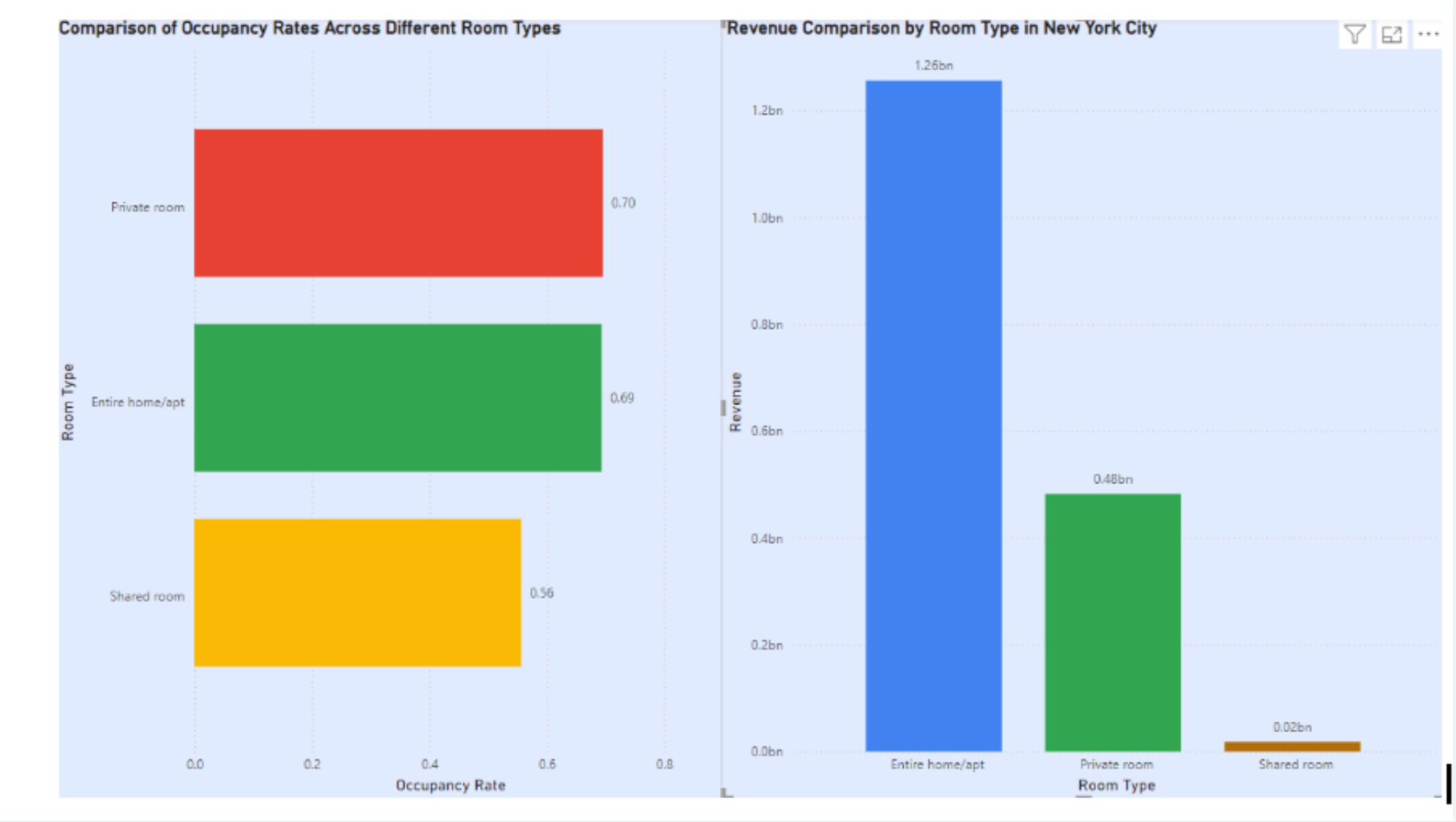

| Neighbourhood | Average Price | Listing Count |
|---|---|---|
| Fort Wadsworth | 800.00 | 1 |
| Woodrow | 700.00 | 1 |
| Tribeca | 490.64 | 177 |
| Sea Gate | 487.86 | 7 |
| Riverdale | 442.09 | 11 |
| Prince's Bay | 409.50 | 4 |
| Battery Park City | 367.56 | 70 |
| Flatiron District | 341.93 | 80 |
| Randall Manor | 336.00 | 19 |
| NoHo | 295.72 | 78 |
| SoHo | 287.10 | 358 |
| Midtown | 282.72 | 1545 |
| Neponsit | 274.67 | 3 |
| West Village | 267.68 | 768 |
| Greenwich Village | 263.41 | 392 |
| Chelsea | 249.74 | 1113 |
| Willowbrook | 249.00 | 1 |
| Theater District | 248.01 | 288 |
| Nolita | 230.14 | 253 |
| Financial District | 225.49 | 744 |
| Gramercy | 222.75 | 338 |
| Little Italy | 222.07 | 121 |
| Murray Hill | 220.96 | 485 |
| Breezy Point | 213.33 | 3 |
| Cobble Hill | 211.93 | 99 |
| Upper West Side | 210.92 | 1971 |
| Brooklyn Heights | 209.06 | 154 |
| Hell's Kitchen | 204.81 | 1957 |
| Kips Bay | 202.41 | 470 |
| DUMBO | 196.31 | 36 |
| Civic Center | 191.94 | 52 |
| Upper East Side | 188.95 | 1798 |
| Vinegar Hill | 187.18 | 34 |

**High Prices:** Tribeca, SoHo, and Flatiron District in Manhattan are premium markets with high returns.

**Listing Volume:** Williamsburg, Bedford–Stuyvesant in Brooklyn, and Harlem in Manhattan have the highest listing counts, indicating high occupancy rates.

**Growth Potential:** Outer areas like Staten Island offer less Intense markets with potential for new hosts.

# OCCUPANCY & REVENUE ANALYSIS



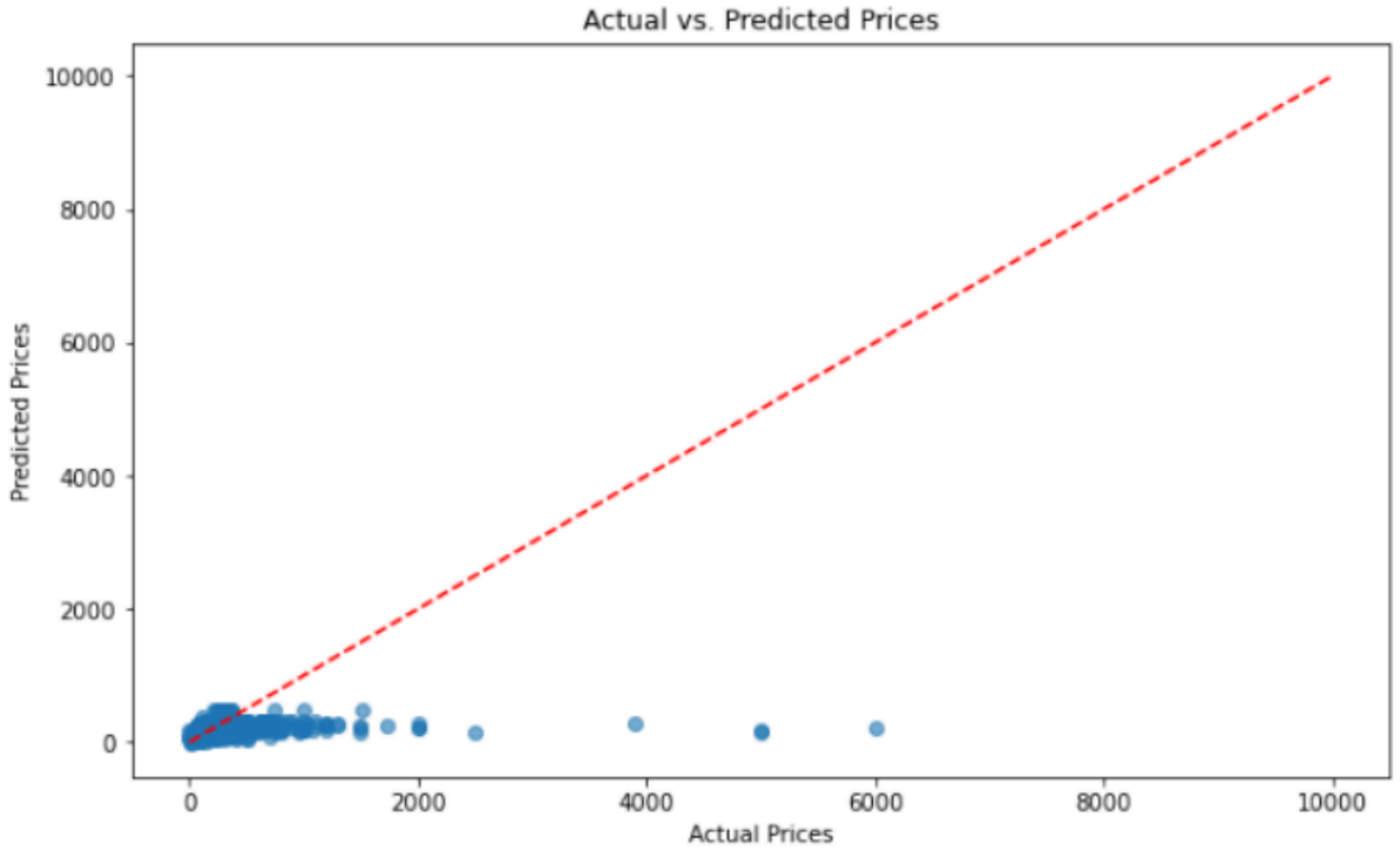

**Occupancy Rates:** Private rooms lead slightly at 0.70, with entire homes/apartments close behind at 0.69. Shared rooms trail at 0.56.
**Revenue:** Entire homes/condos dominate, generating $1.26 billion, far surpassing private rooms ($0.48 billion) and shared rooms ($0.02 billion).

**Insight:** Prioritize entire homes/apartments to maximize revenue, especially in high-demand areas. Private rooms remain valuable for high occupancy, but shared rooms may be best as niche offerings.

# LINEAR REGRESSION MODEL PERFORMANCE ANALYSIS

```
Price Prediction Mean Absolute Error: 58.353413472571035
Price Prediction R² Score: 0.17645067424735705
```

Giving us:




**Mean Absolute Error (MAE): 58.35:** Indicates the average magnitude of errors between predicted and actual prices. Lower values are preferred.

**$R^2$ Score: 0.18:** Suggests that only 17.65% of the variance in the actual prices is explained by the model, indicating a poor fit.

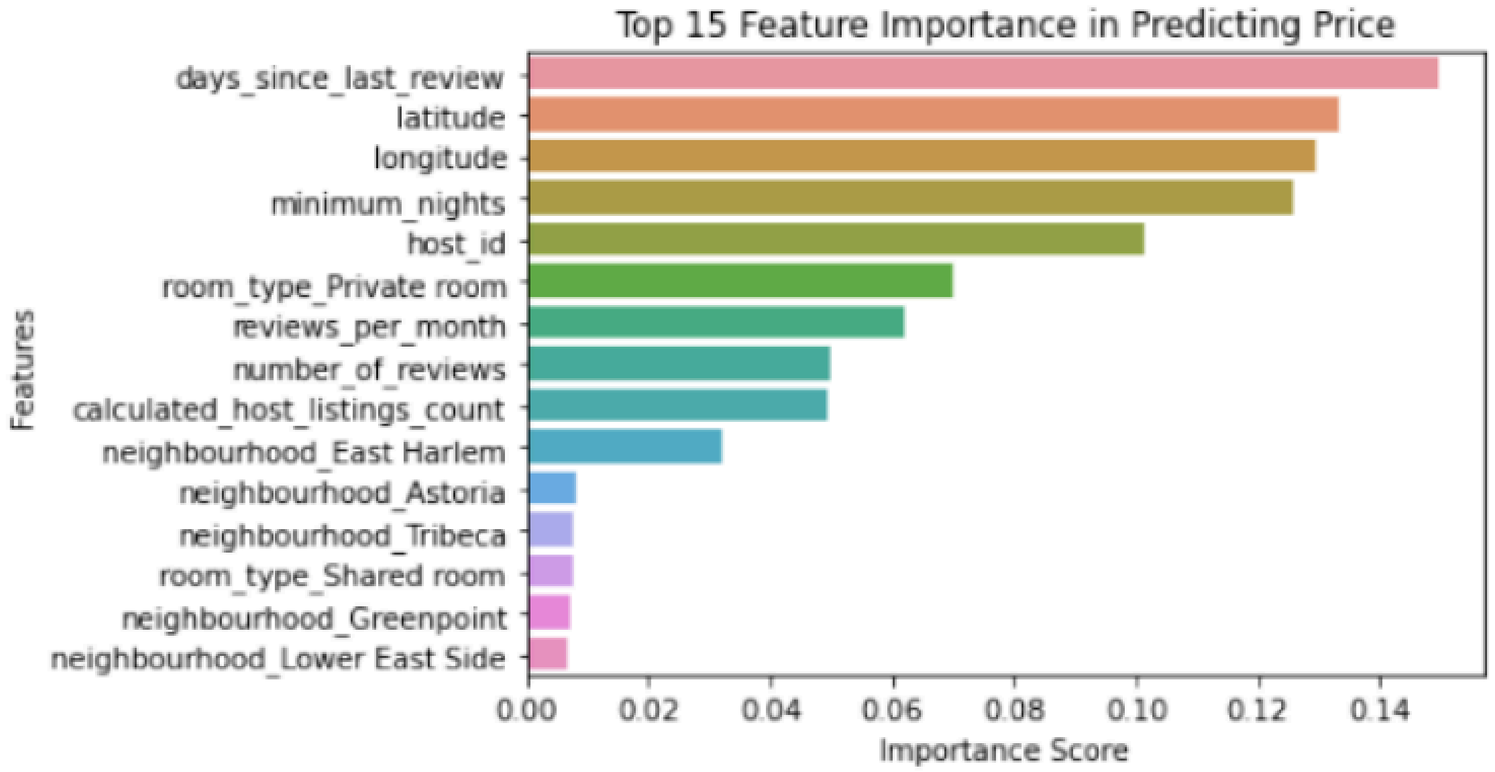
**Insights:**

The model underperforms, especially for higher prices, with many predictions clustered near the origin.

The high error and low $R^2$ indicate that the model or features need refinement for better accuracy.

# TOP 15 FEATURES INFLUENCING PRICE PREDICTION

**Purpose:** Visualizes the most significant features affecting price prediction in the Random Forest model.




**Key Features:**

**days_since_last_review:** Most critical factor influencing price.

**Latitude and Longitude:** Location plays a major role.

**Minimum Nights:** Significantly impacts price prediction.

**Private Room Type and Reviews per Month.**

**Chart Explanation:**

**Y-axis:** Lists the top 15 influential features.

**X-axis:** Shows the importance score, indicating each feature's contribution to the model's prediction.

# RECOMMENDATIONS & CONCLUSION

**Recommendations**

**1. For Hosts:**

Diversification

Customer Engagement

Focus on Entire Homes/Apartments

**2. For Investors:**

Target High-Value Areas

Explore Emerging Markets

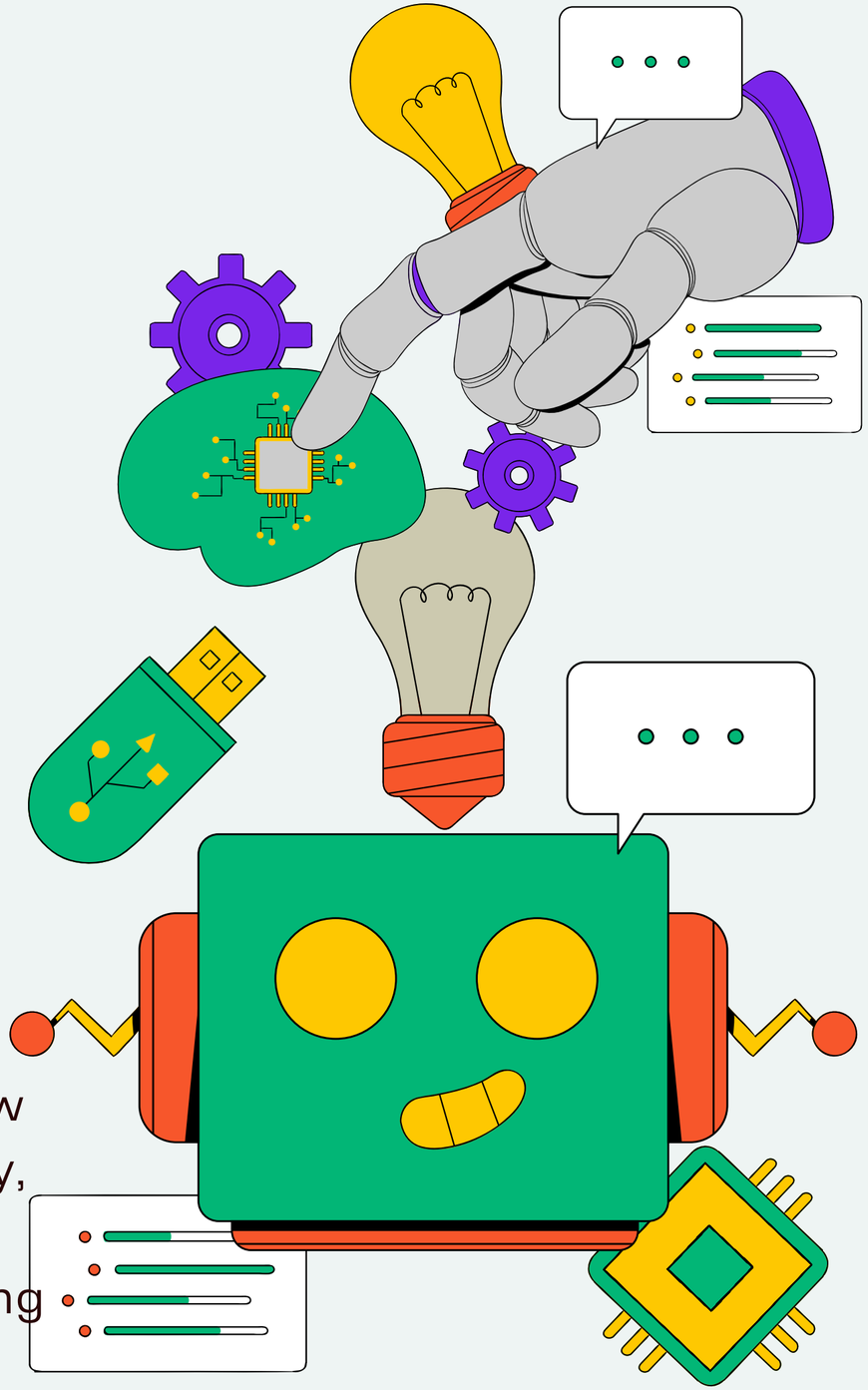
**3. For Policymakers:**

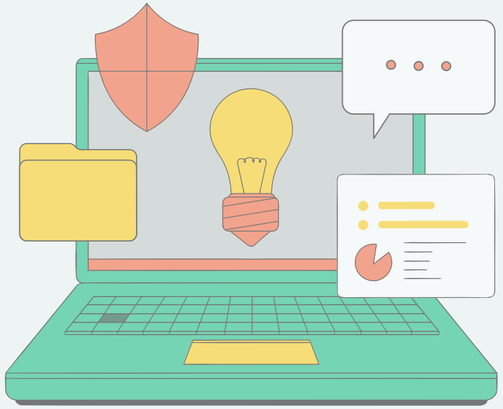
Regulatory Balance

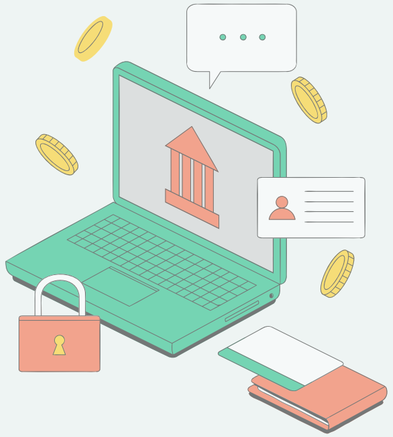
Community Impact

**Conclusion:**

This analysis offers actionable insights for Airbnb stakeholders in New York City. Hosts and investors can optimize strategies for profitability, while policymakers can craft informed regulations. As the sharing economy evolves, data-driven decisions will be essential for navigating its challenges and opportunities.







# THANK YOU!