

Rebuilding Airbnb Business Strategies in Post-Pandemic Era : NYC

Submitted by ,
Sindhu L
DS C63

AGENDA

- Objective
- Background
- Key findings
- Recommendations
- Appendix:
 - ❖ Data sources
 - ❖ Data methodology
 - ❖ Data model assumptions

OBJECTIVE

- To understand the trends on customer preference on properties
- Refining our understanding on post pandemic market situation
- To present the quality of the data for accurate analysis to lead data analysis managers and senior analysts

BACKGROUND

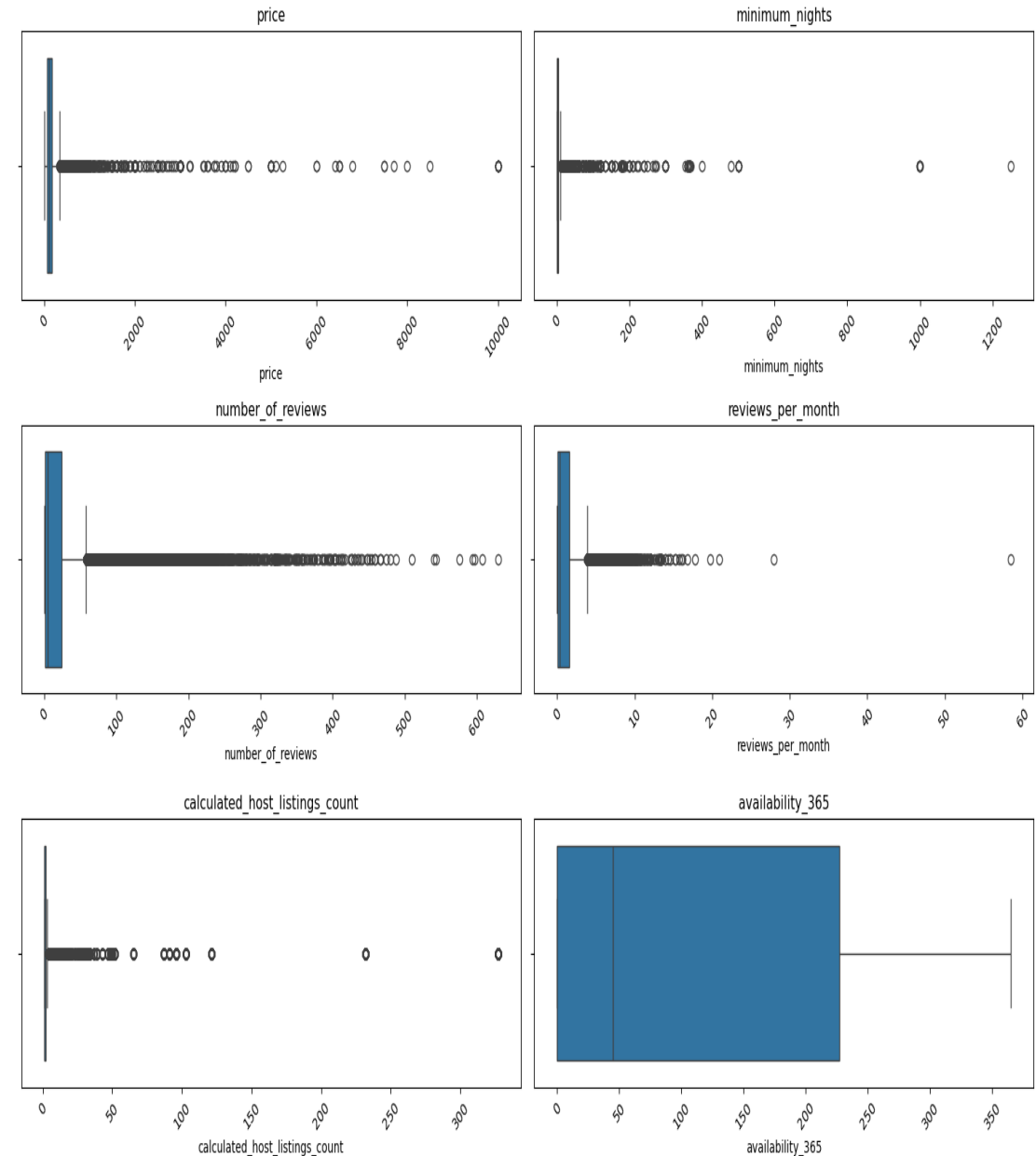
- Airbnb has faced a significant revenue decline in recent months
- With travel restrictions lifting, demand is beginning to rise
- Airbnb aims to prepare strategically for this renewed travel surge

ADDRESSING OUTLIERS AND MISSING VALUES IN DATASET

- Outliers were handled by capping values above the 99th percentile with the 99th percentile value
- Null values in reviews_per_month were filled with 0, assuming these listings have not received any reviews
- Null values in name and host_name are missing completely at random and require further investigation; For this analysis they were left unattended

Percentage of null values in data set

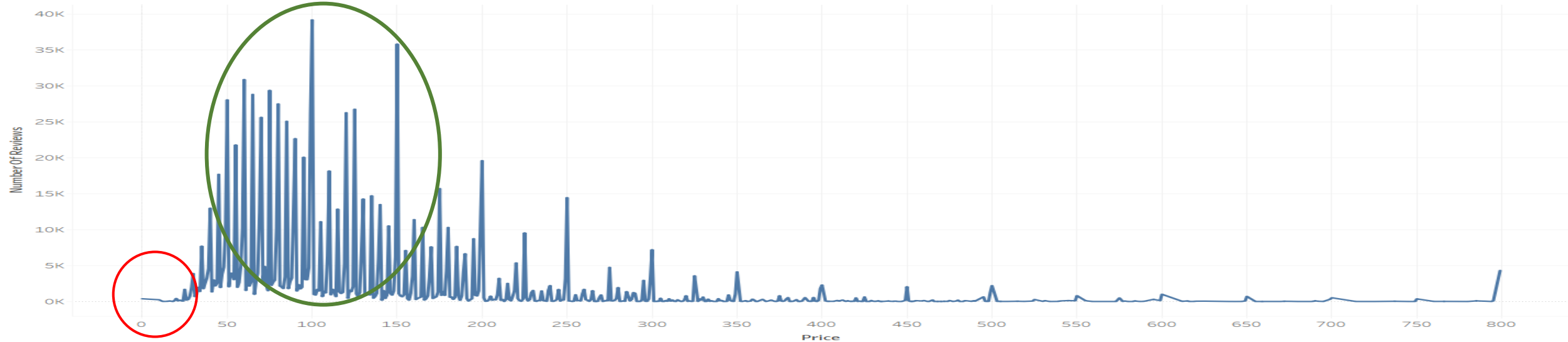
name	0.032723
host_id	0.000000
host_name	0.042949
last_review	20.558339
reviews_per_month	20.558339



INSIGHTS ON PRICE PREFERENCE BY CUSTOMERS

- Customers prefer stays priced between \$50 and \$200
- Properties with zero price have to be examined
- Customer preference declines as price increases

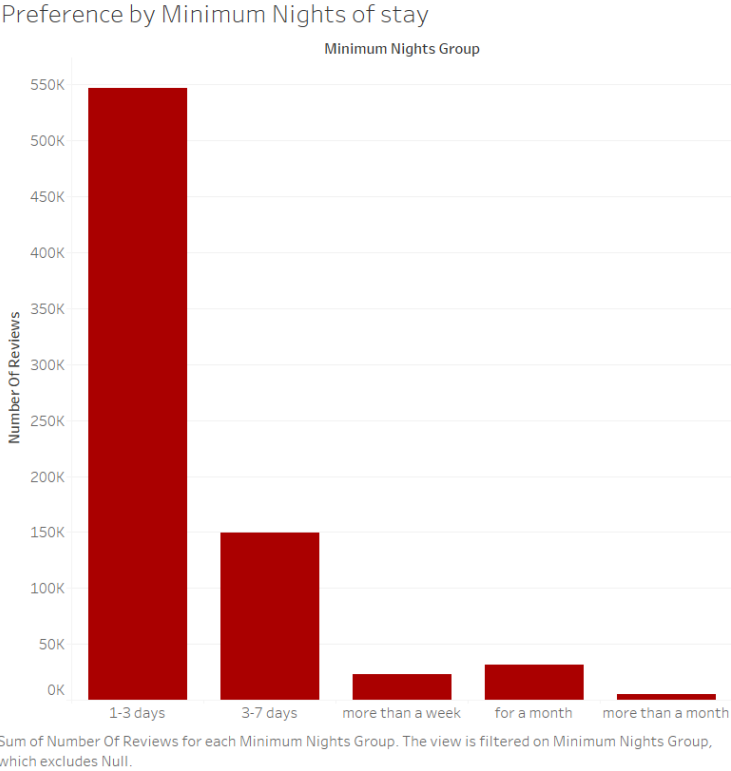
Preference by Price



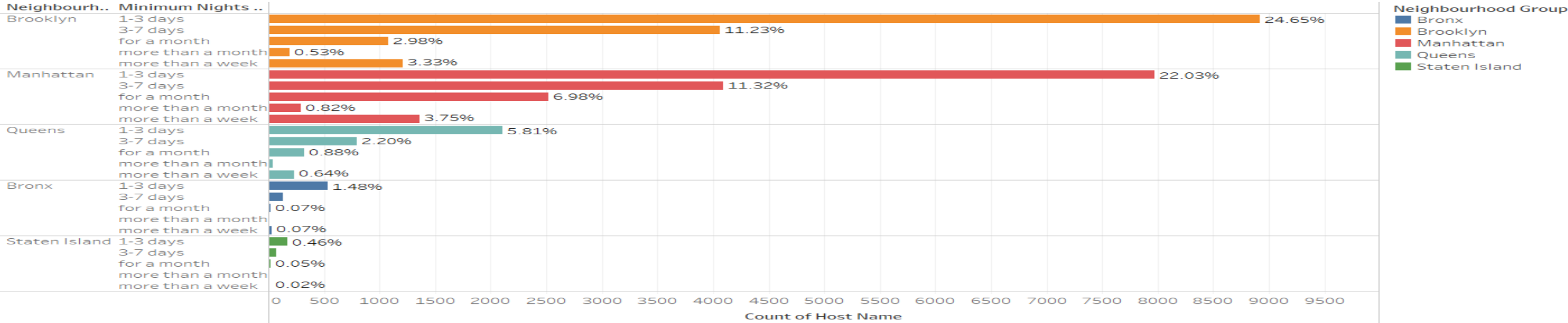
The trend of sum of Number Of Reviews for Price.

KEY TRENDS IN CUSTOMER STAYS AND POPULAR NEIGHBOURHOODS

- 72% Customer prefers 1-3 days for their stays
- Rise in Demand for monthly stays and Manhattan Leads this trend
- Manhattan and Brooklyn dominates the property listings in NYC



Number of listings with Minimum Stay and Neighbourhood



Count of Host Name for each Minimum Nights Group broken down by Neighbourhood Group. Color shows details about Neighbourhood Group. The marks are labeled by % of Total Count of Host Name. The view is filtered on Minimum Nights Group, which excludes Null.

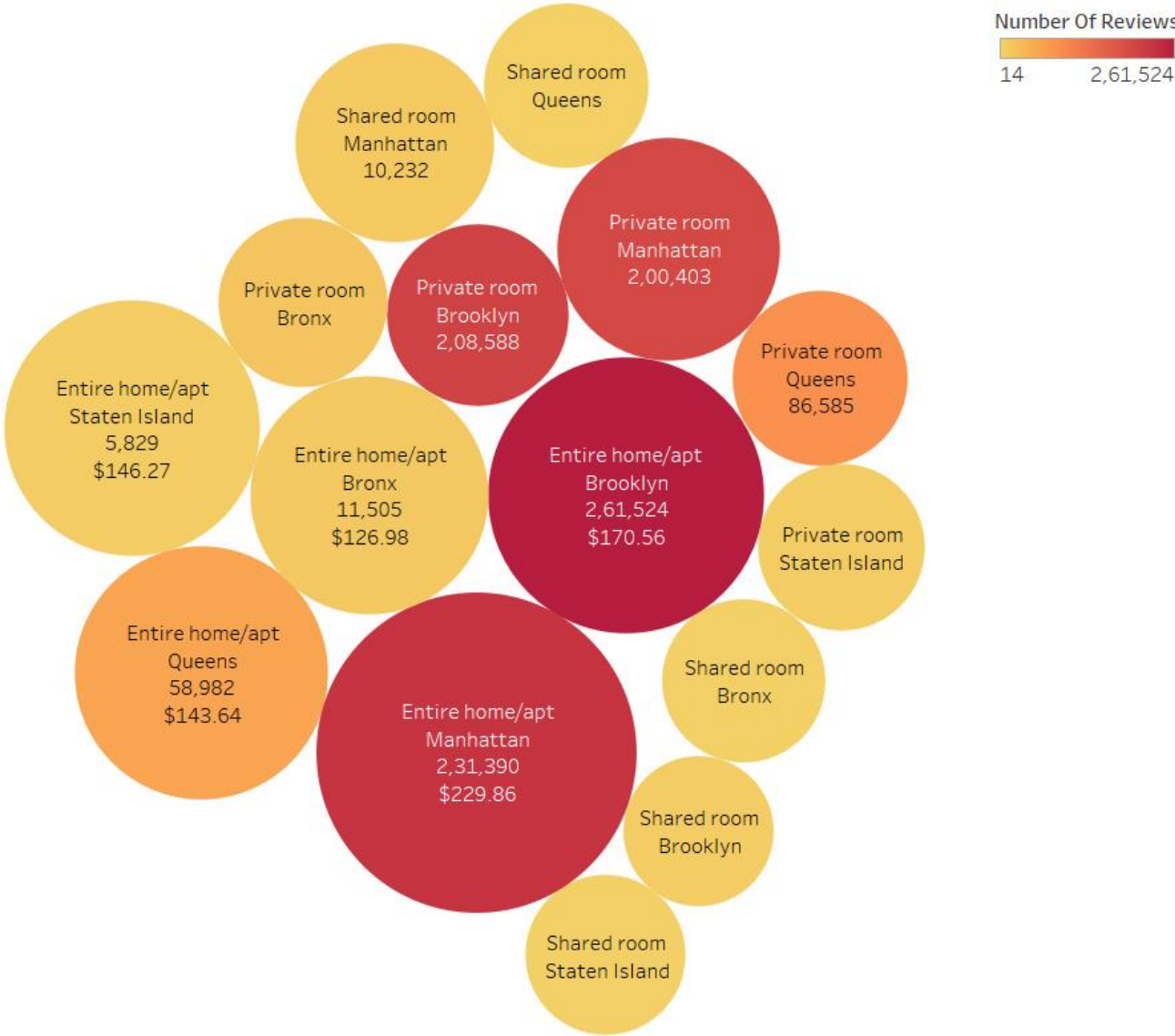
ROOM TYPE AND NEIGHBOURHOOD FAVOURITES OF CUSTOMERS !!

Customer Preferences by Room type, Neighbourhood group

- Customer prefers Entire home/apt for their stays seeking privacy and to feel home-like experience
- Staten Islands attracted the customers
- Improvement required for Shared rooms

Average Number of reviews per listing

Room Type	Neighbourhood Group				
	Bronx	Brookl..	Manha..	Queens	Staten..
Entire home/apt	30.44	27.37	17.53	28.14	33.12
Private room	24.51	20.60	25.13	25.69	29.42
Shared room	7.20	13.97	21.32	12.60	1.56



Room Type, Neighbourhood Group, sum of Number Of Reviews and average of Price. Color shows sum of Number Of Reviews. Size shows average of Price. The marks are labeled by Room Type, Neighbourhood Group, sum of Number Of Reviews and average of Price.

RECOMMENDATIONS

- Focus on property prices, as there are unusual entries are present
- Promote properties with short stays and monthly stays to attract more customers
- Emphasize the affordability of shared rooms in marketing campaigns

APPENDIX – DATA SOURCES

- Here is a snapshot of our data dictionary.
 - **Categorical Variables:** room_type, neighbourhood_group, neighbourhood
 - **Numerical Variables:** Price, minimum_nights, number_of_reviews, reviews_per_month, calculated_host_listings_count, availability_365
 - **Location Variables:** latitude, longitude
 - **Time Variable:** last_review
- Source of Data: Upgrad.com

APPENDIX – DATA METHODOLOGY

We conducted a thorough analysis of Airbnb dataset. The process included:

- Cleaned the data and did thorough univariate ,multivariate analysis
- Handled outliers and null values with appropriate methods
- For effective visualization I used Tableau

For detailed version of Methodology steps –



Methodology_Document.pdf

APPENDIX- DATA MODEL ASSUMPTIONS

- Number_of_reviews was taken as customer preference
- Assumed that null values in the reviews_per_month column indicate that customers have not reviewed those listings; so replaced with 0
- Assuming other null values doesn't impact this analysis