REVEALING THE SECRETS OF AIRBNB IN NYC: A DATA-DRIVEN NARRATIVE

SUKANYA DEBNATH

SIDHARTH JAYANTH KINI

SATHWIK BIDAYANTH

August 2024

AGENDA

Objective

Background

Key findings

Recommendations

Appendix

- Data sources
- Data methodology
- Data model assumptions

OBJECTIVE

To offer insights into the current market conditions.

To improve our knowledge of property, host acquisitions, operations, customer preferences and handling the website and Airbnb app properly.

To offer necessary suggestions to our operations and marketing team at the earliest.

BACKGROUND

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

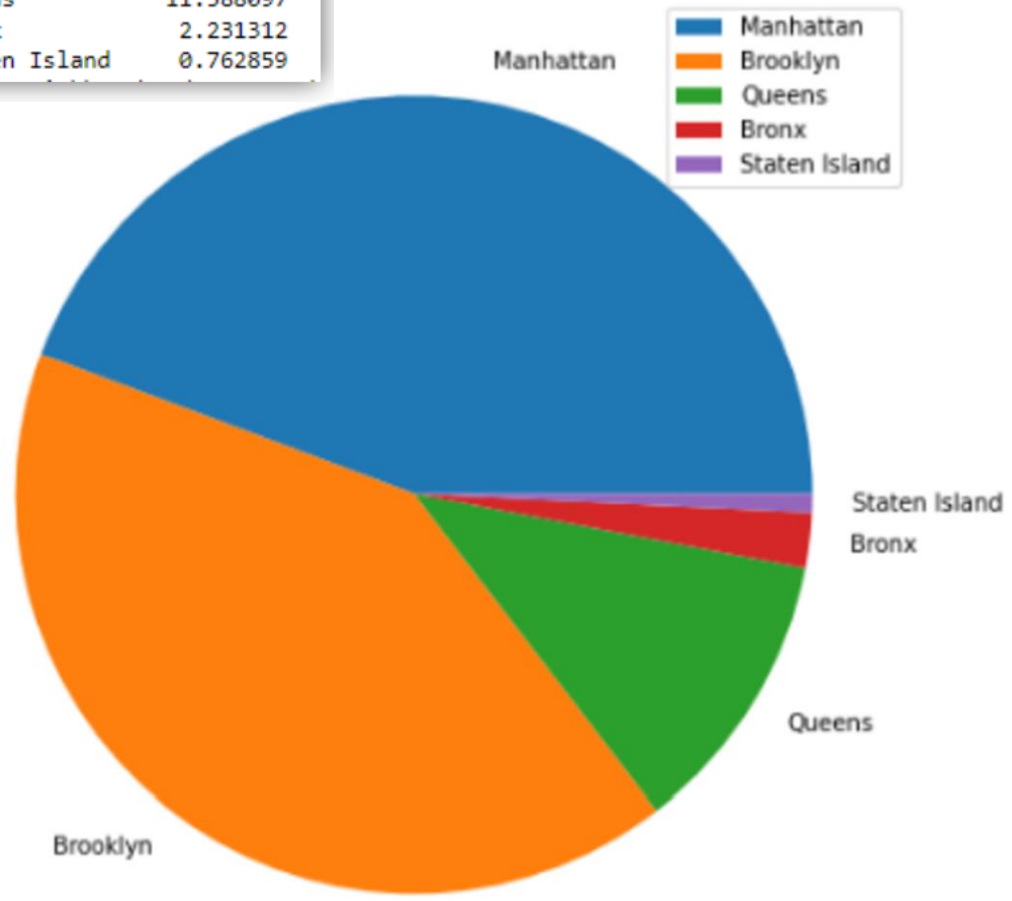
- The restrictions have started lifting and people have started to travel more, Airbnb wants to make sure that it is fully prepared for this change.

- So, the analysis has been conducted on a dataset consisting of numerous Airbnb listings in New York.

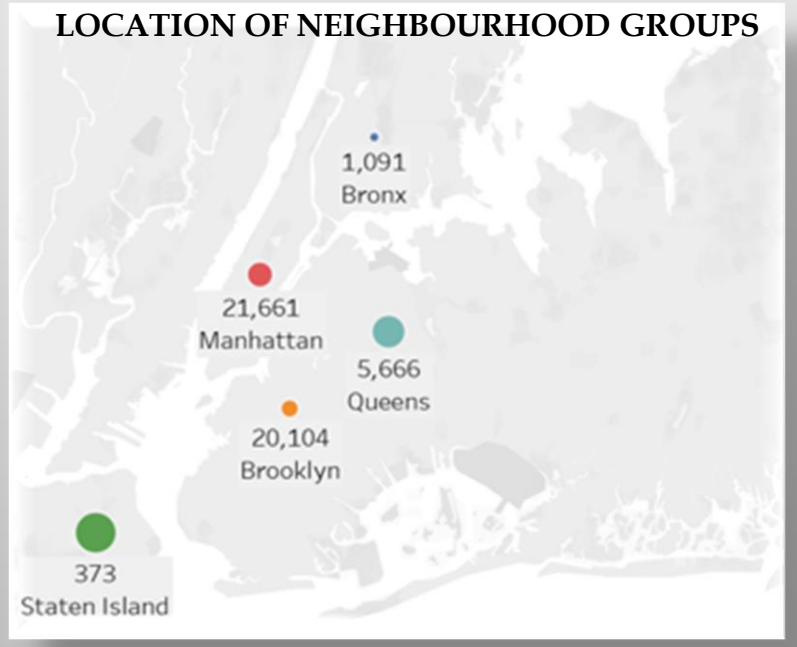
MOST CONTRIBUTING NEIGHBOURHOODS

Neighborhood group percentages

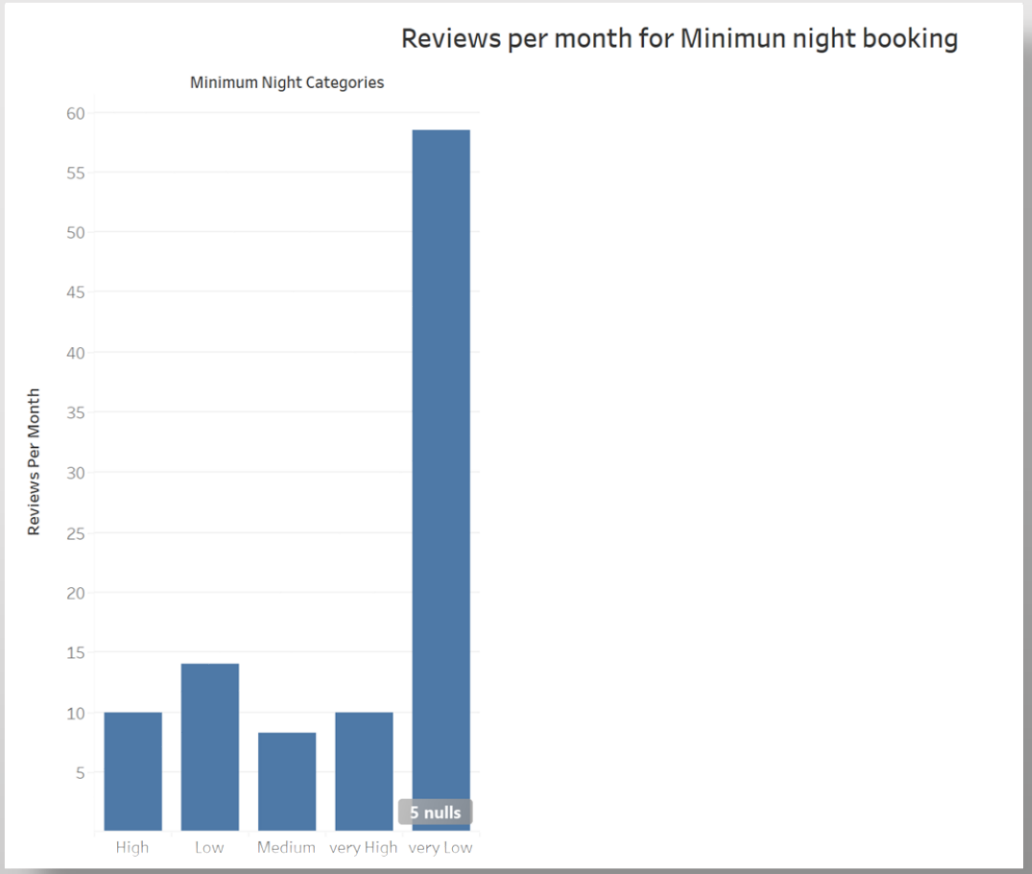
Manhattan	44.301053
Brooklyn	41.116679
Queens	11.588097
Bronx	2.231312
Staten Island	0.762859

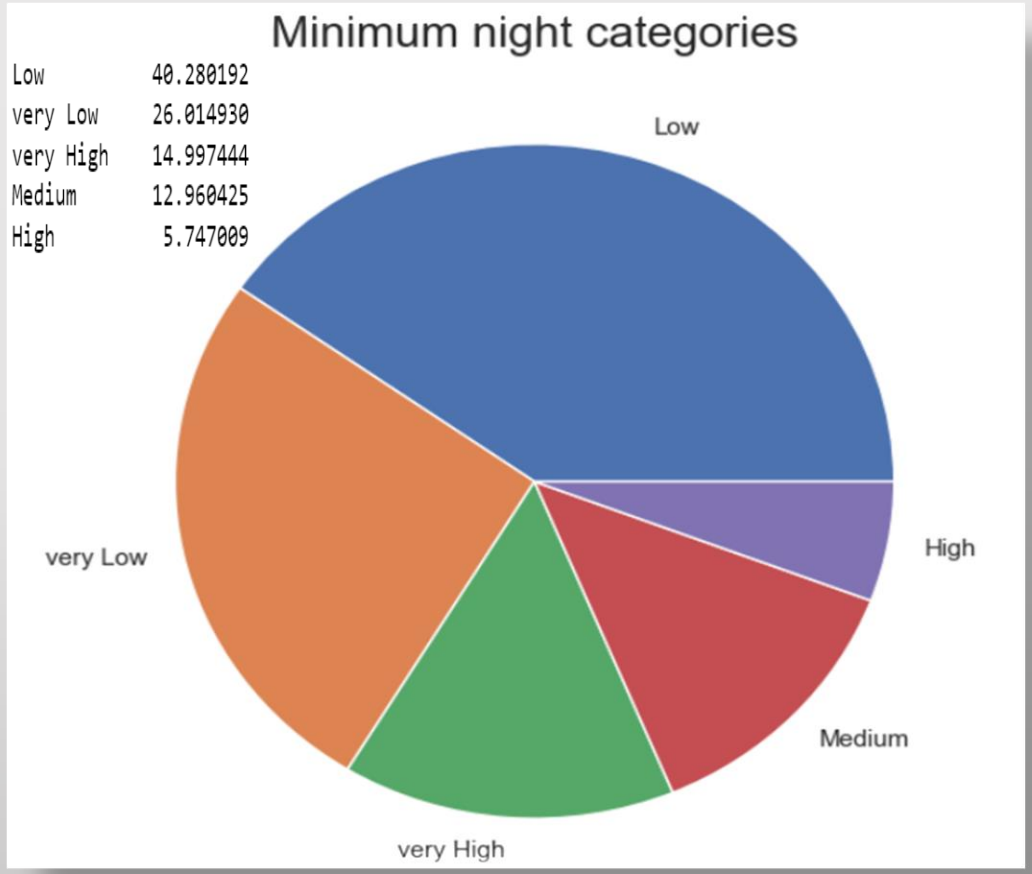


Approximately 85% of the listings belong to Manhattan and Brooklyn neighbourhood group.



CONTRIBUTION OF THE MINIMUM NIGHT FEATURE

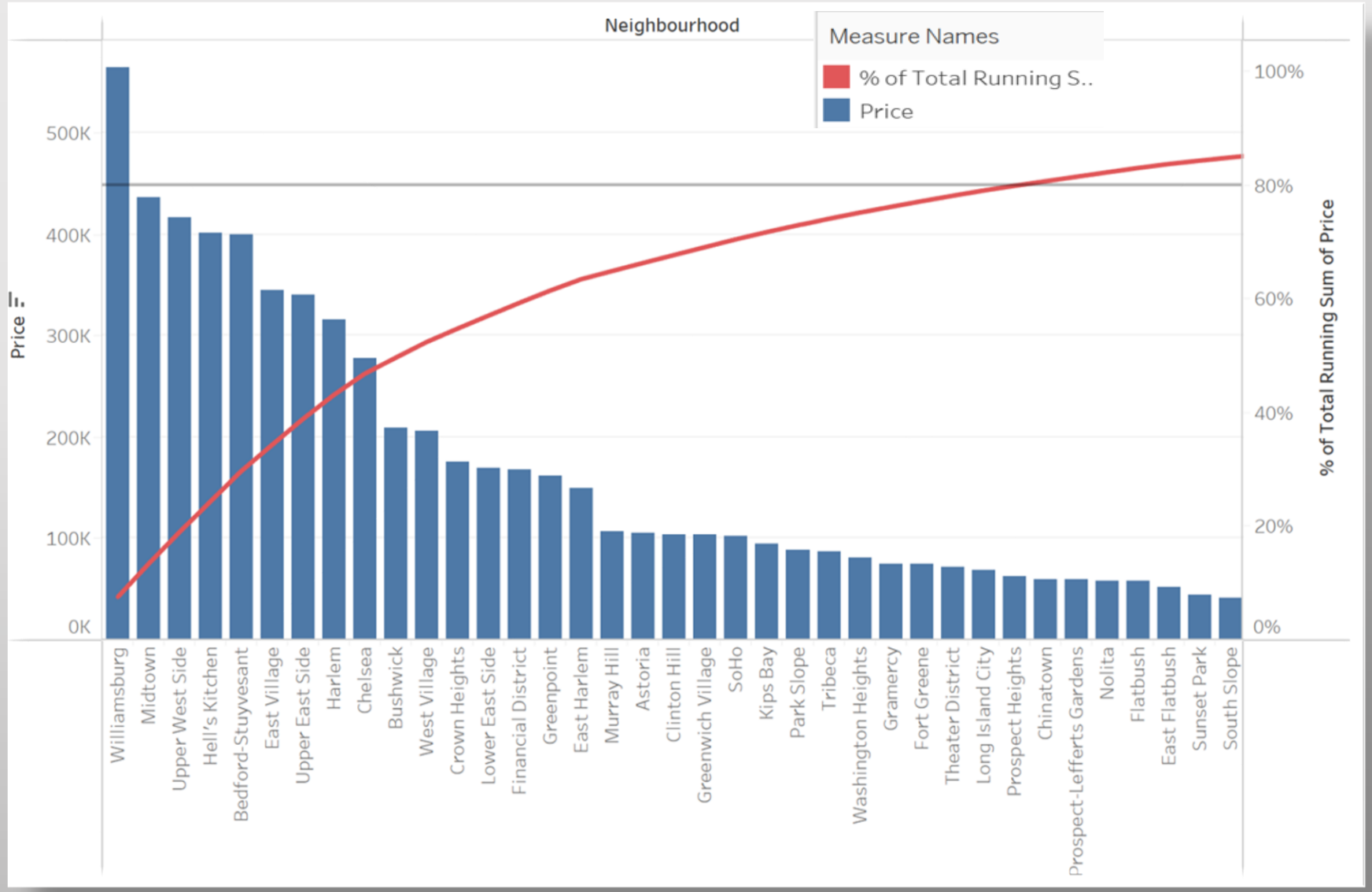




Listings with very low minimum night category tends to receive the highest reviews on a monthly basis.

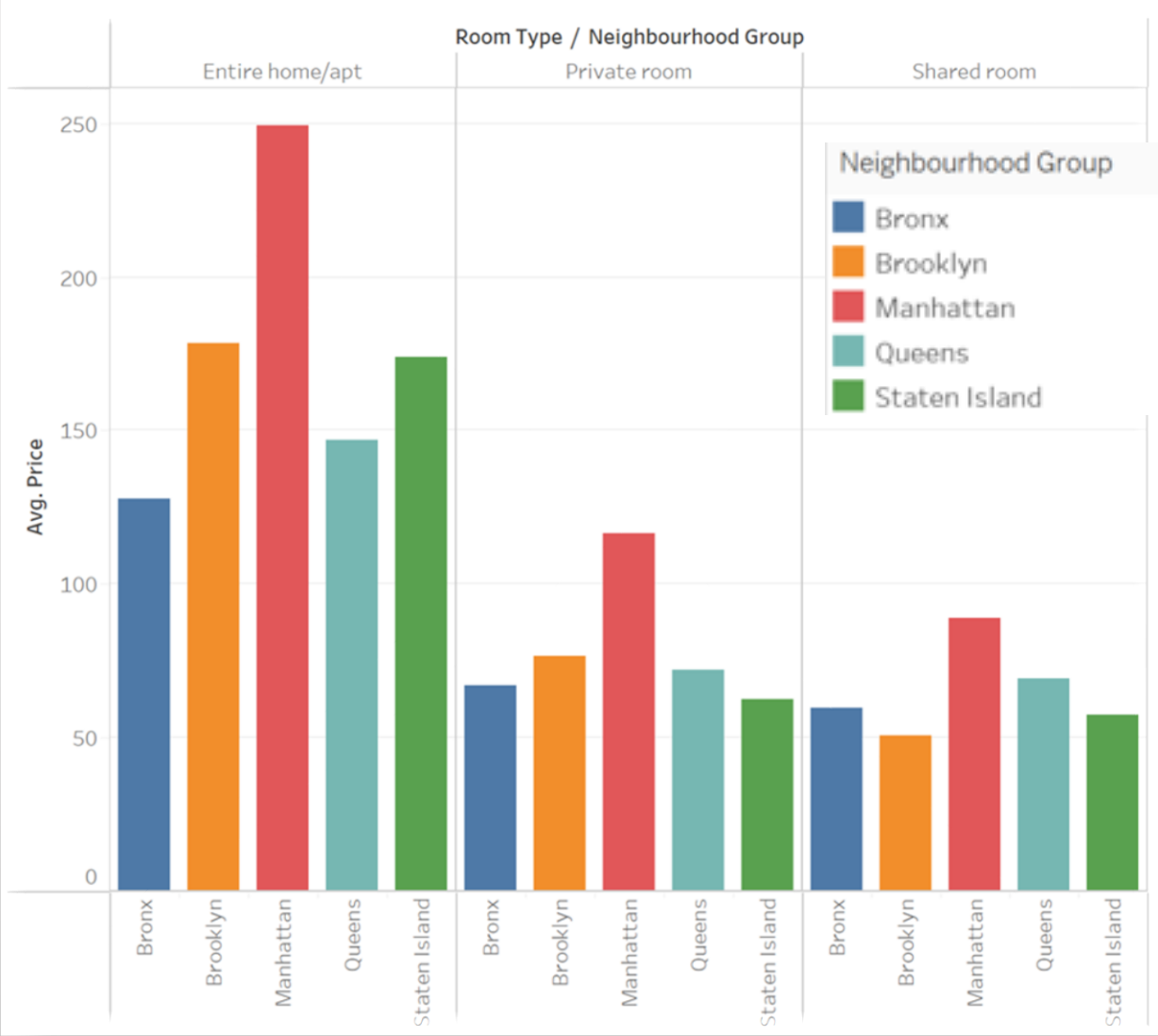
The low minimum night category contributes ~40% of the listings

NEIGHBOURHOOD EVIDENTLY EFFECTING PRICE



Approximately 80% of the total price is contributed by the the top 30 neighbourhoods such as Williamsburg, Midtown, Upper west side, Hells's kitchen etc.

INFLUENCE OF ROOM TYPE & NEIGHBOURHOOD GROUPS ON PRICE



The shared room type in every neighbourhood groups has the least average price

The entire home/apt type in every neighbourhood groups has the highest average price

INFLUENCE OF AVAILABILITY AND PRICE ON REVIEWS

AVG(Reviews Per Month)

0.403 2.130

Price Categories

Availability ..	very High	High	Medium	Low	very Low
very High	1.278	1.278	1.179	1.642	1.563
High	2.062	1.956	2.122	2.098	1.789
Medium	1.939	1.706	1.891	2.130	2.093
Low	1.351	1.450	1.700	2.085	2.092
very Low	0.403	0.428	0.466	0.557	0.498

The listings with very low availability have the lowest reviews per month irrespective of the price categories.

On an average the listings with the high availability have the highest reviews per month across the price categories.

The highest reviews per month can be found for the listing with medium availability and low price.

CONCLUSION & RECOMMENDATIONS

The head of acquisitions & operations needs to pay a good amount attention on the **shared room type**. They need to check whether there is any issues with the service.

Ensuring **strategic optimization in the price of the properties among the cities** will ensure improvement in overall user experience ,higher engagements and conversion rates.

The **minimum night booking** threshold should be on the **lower** side in order to attract the customers as well as ensuring the customer satisfaction.

It is necessary to **optimise the order of property listing** in Manhattan and Brooklyn in order to get every property the optimal amount of traction in Queens, Bronx and Staten Island.

The **availability** of the listings **needs to be increased** in order to make it more customer oriented and to provide the customer **stable booking options** throughout the year. This increase will also contribute towards **revenue generation.**

APPENDIX - DATA SOURCES

The columns in the dataset are self-explanatory. You can refer to the diagram given below to get a better idea of what each column signifies.

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

Performed a comprehensive analysis of the New York Airbnbs dataset.

Cleaned the dataset using Python.

Extracted the essential features.

Applied group aggregation, pivot tables and other statistical techniques.

Developed charts and visualizations using Tableau.

APPENDIX - DATA ASSUMPTIONS

```
Categorical Variables:
    - room_type
    - neighbourhood_group
    - neighbourhood

Continous Variables(Numerical):
    - Price
    - minimum_nights
    - number_of_reviews
    - reviews_per_month
    - calculated_host_listings_count
    - availability_365
- Continous Variables could be binned in to groups too

Location Varibles:
    - latitude
    - longitude

Time Varibale:
    - last_review
```