**AIRBNB, NYC STORYTELLING CASE STUDY**

# DATA DRIVEN

RUSHABH PATEL

C59 BATCH, 2024

# AGENDA

- Objective

- Glance of Problem

- Key findings

- Recommendations Appendix:

# OBJECTIVE

- Offer insights into the present market conditions.

- Improve comprehension of property and host acquisitions, operational processes, and customer preferences.

- Deliver preliminary suggestions to our marketing and operations teams.

# GLANCE OF PROBLEM

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

- Now that 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.
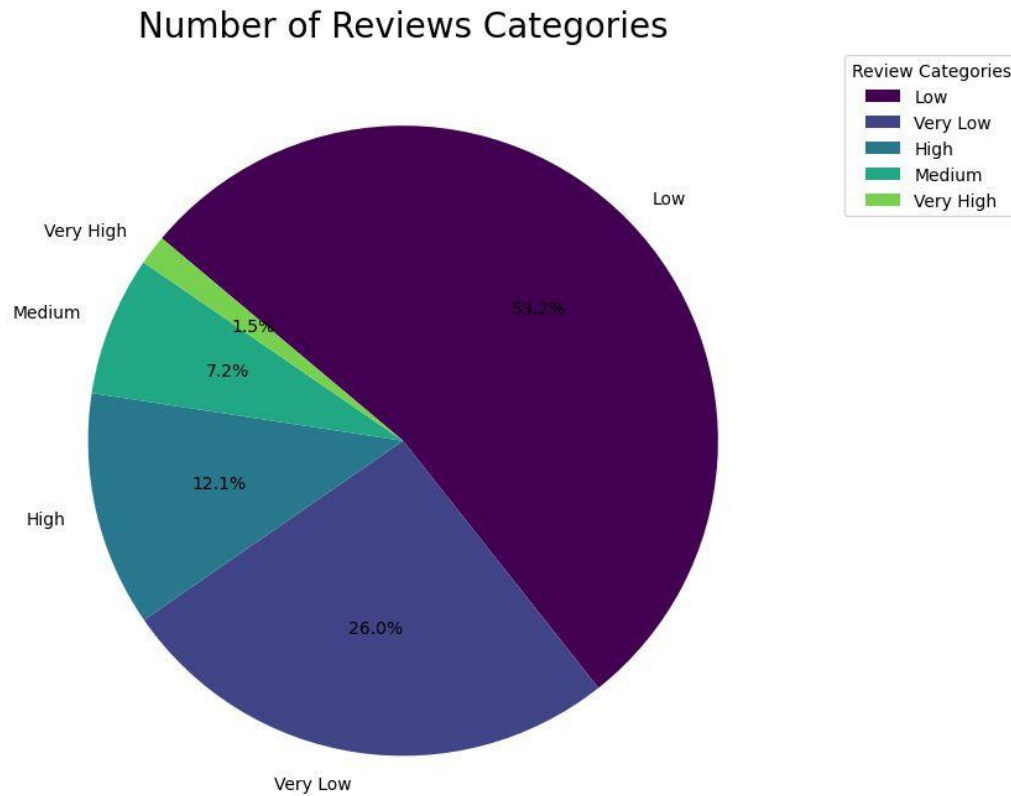
# MOST CONTRIBUTING NEIGHBORHOODS

- 85 % of the listing are Manhattan and Brooklyn neighborhood group.
- Staten Island has the lowest contribution.



Neighbourhood Group Distribution
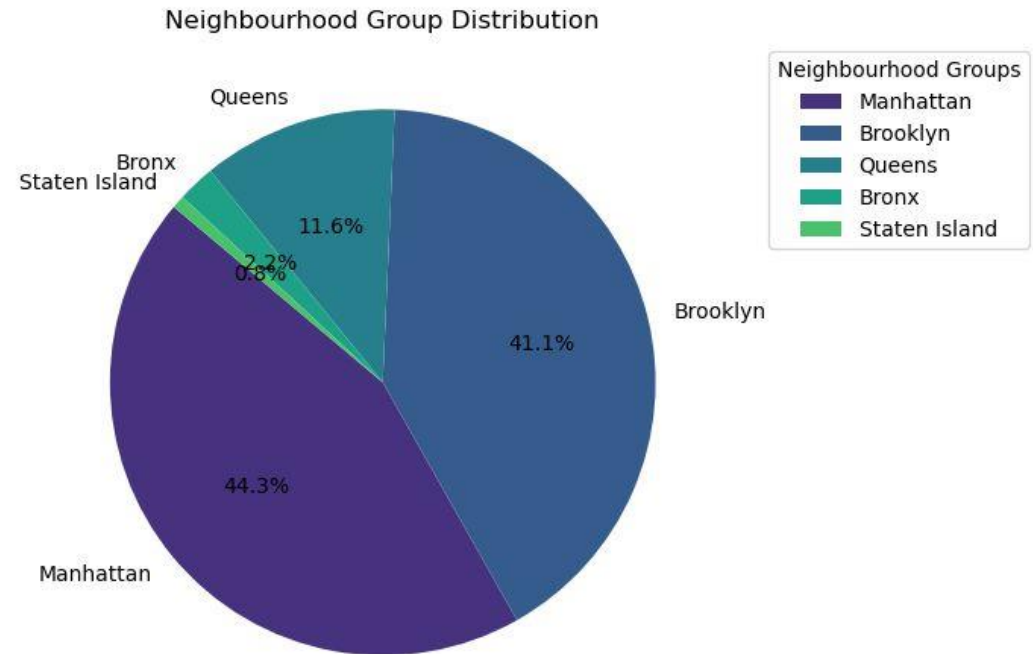
- The top 60 hosts only make up 20% of the total host count.
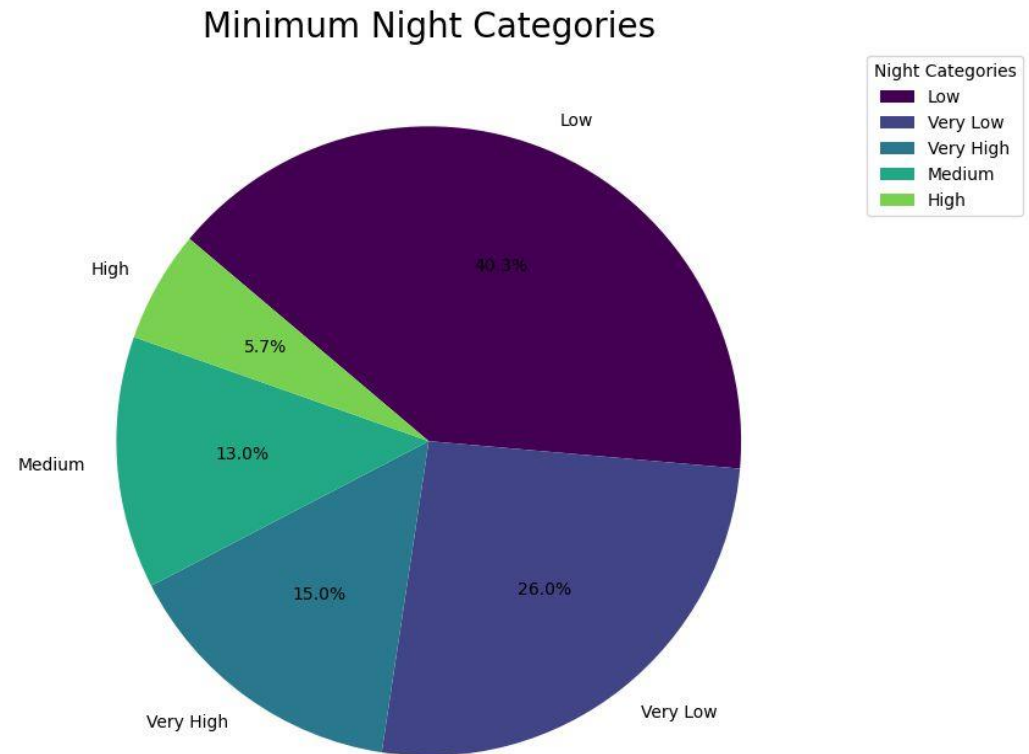
# MINIMUM NIGHT CATEGORIES
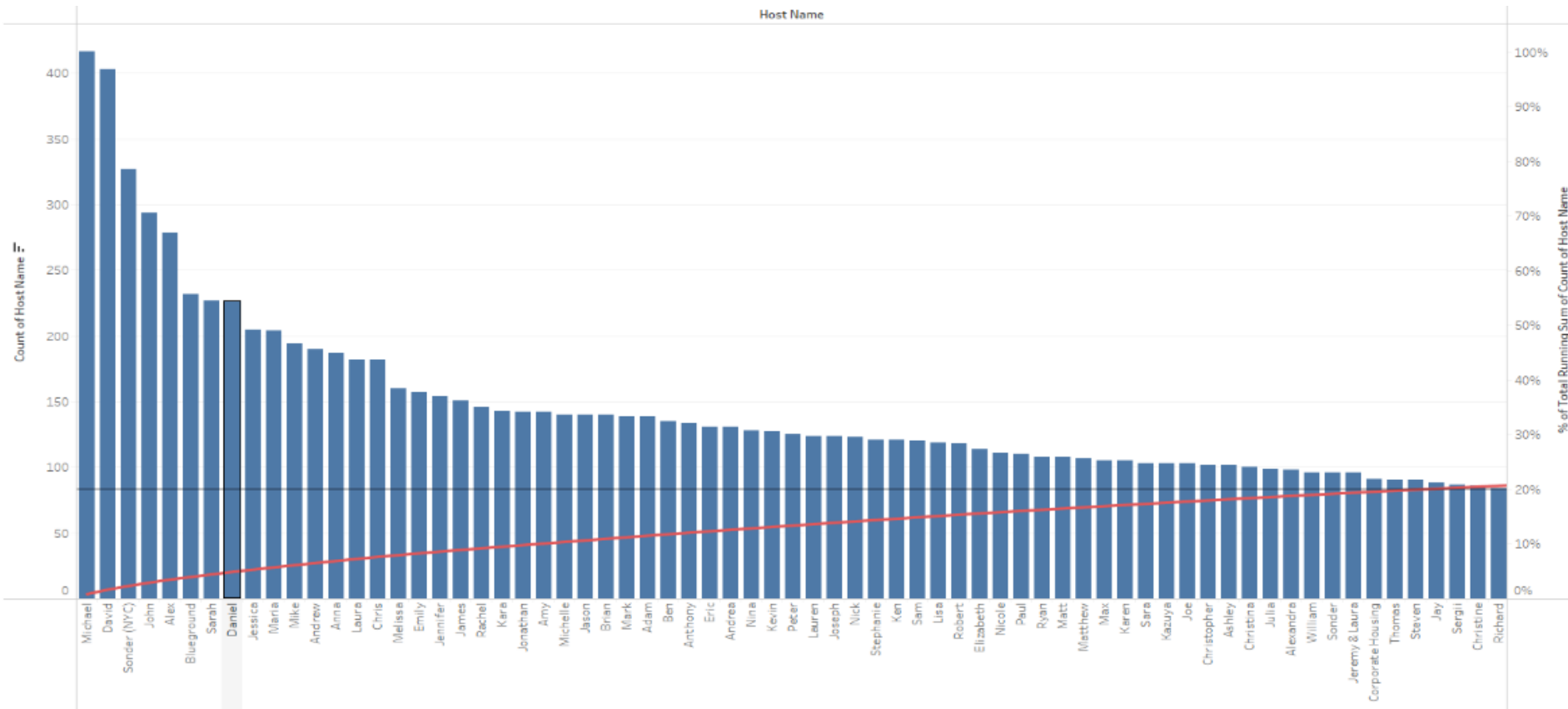


Number of Reviews Categories

• Low category in minimum night feature contributes 40% and Very Low category contributes 26%.

# EFFECT OF MINIMUM NIGHT ON REVIEWS

• Customers are more likely to leave reviews for lower number of minimum nights.



Minimum Night Categories

Night Categories
- Low
- Very Low
- Very High
- Medium
- High

Low 40.3%
High 5.7%
Medium 13.0%
Very High 15.0%
Very Low 26.0%

# CONCLUSION

- Inspections are necessary for **shared rooms**.

- The overall contributions from all hosts are more beneficial than just a few hosts excelling.

- Over 80% of the listings are located in the **Manhattan** and **Brooklyn** neighborhoods.

- Lowering the minimum nights threshold can make properties more appealing to customers.

# APPENDIX - DATA SOURCES

| 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 |

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.

# 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
```

# APPENDIX –DATA METHODOLOGY

- Performed a comprehensive analysis of the New York Airbnb dataset, which included the following steps:

- Data Cleaning: Utilized Python to clean and preprocess the dataset.

- Feature Engineering: Extracted and derived essential features.

- Statistical Analysis: Employed group aggregation, pivot tables, and other statistical techniques.