



# UNVEILING THE SECRETS OF AIRBNB IN NYC: A DATA-DRIVEN STORY

MOHAMMAD ASIF

MARCH 2024

# AGENDA

Objective

Background

Key findings

Recommendations

Appendix:

- Data sources
- Data methodology
- Data model assumptions

# OBJECTIVE



To Provide insight into the current market situation



Enhance our understanding of property and host acquisitions, operations, and customer preferences.



Provide early recommendations to our marketing and operations teams

## BACKGROUND

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

A light gray downward-pointing arrow indicating a flow from the first text box to the second.

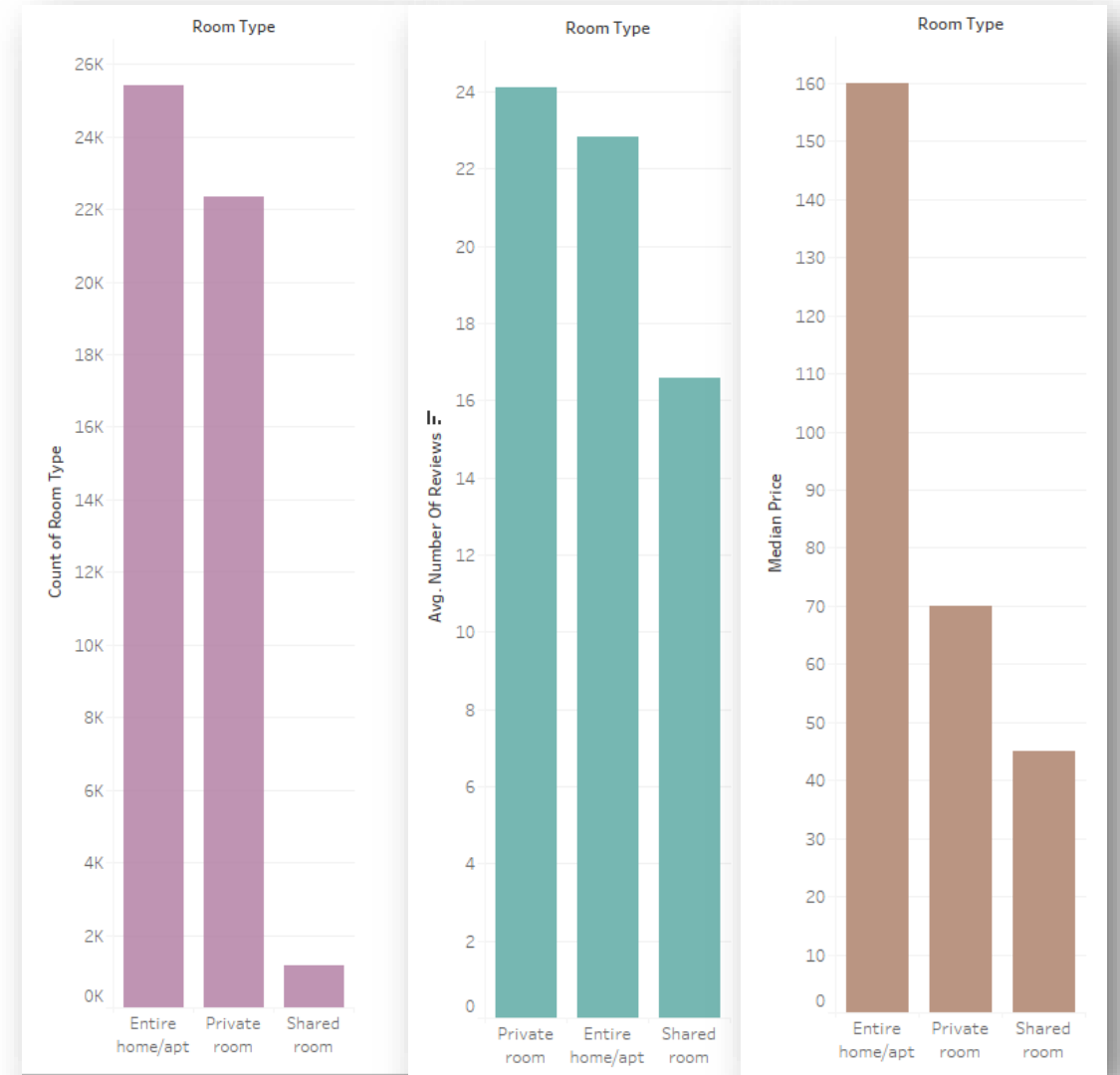
Now that the restrictions have started lifting and people have started to travel more.

A light gray downward-pointing arrow indicating a flow from the second text box to the third.

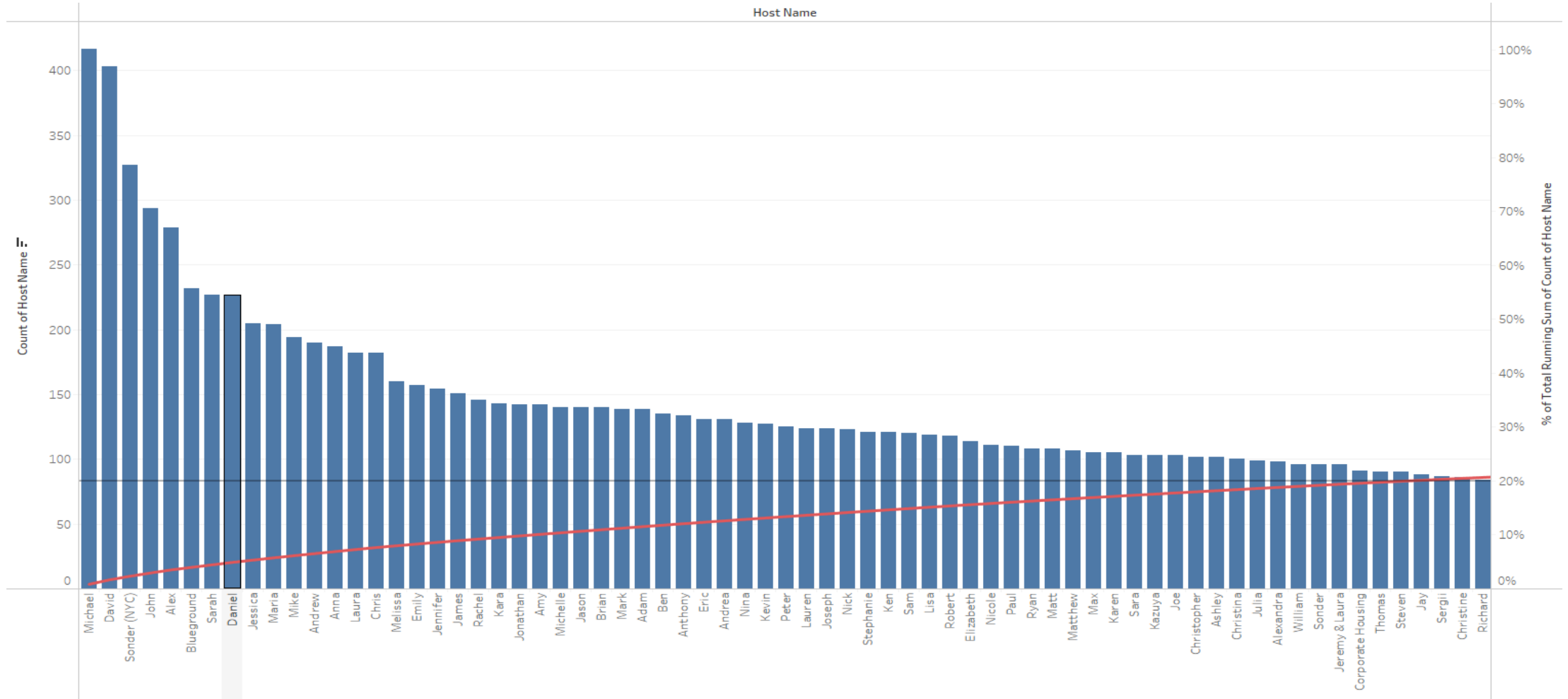
Airbnb wants to make sure that it is fully prepared for this change.

# THE PROBLEMS WITH SHARED ROOMS

- Shared rooms only account for 2 % of the total types of rooms.
- They are less likely to be reviewed.
- Median rates for shared rooms are significantly lower.

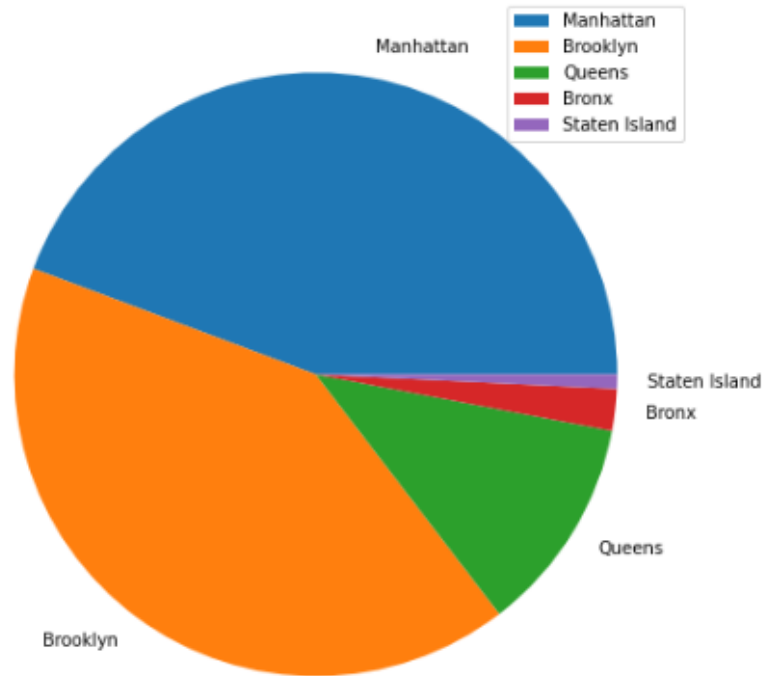


# EVERY HOST MATTER



- The top 60 hosts only make up 20% of the total host count!

# MOST CONTRIBUTING NEIGHBORHOODS

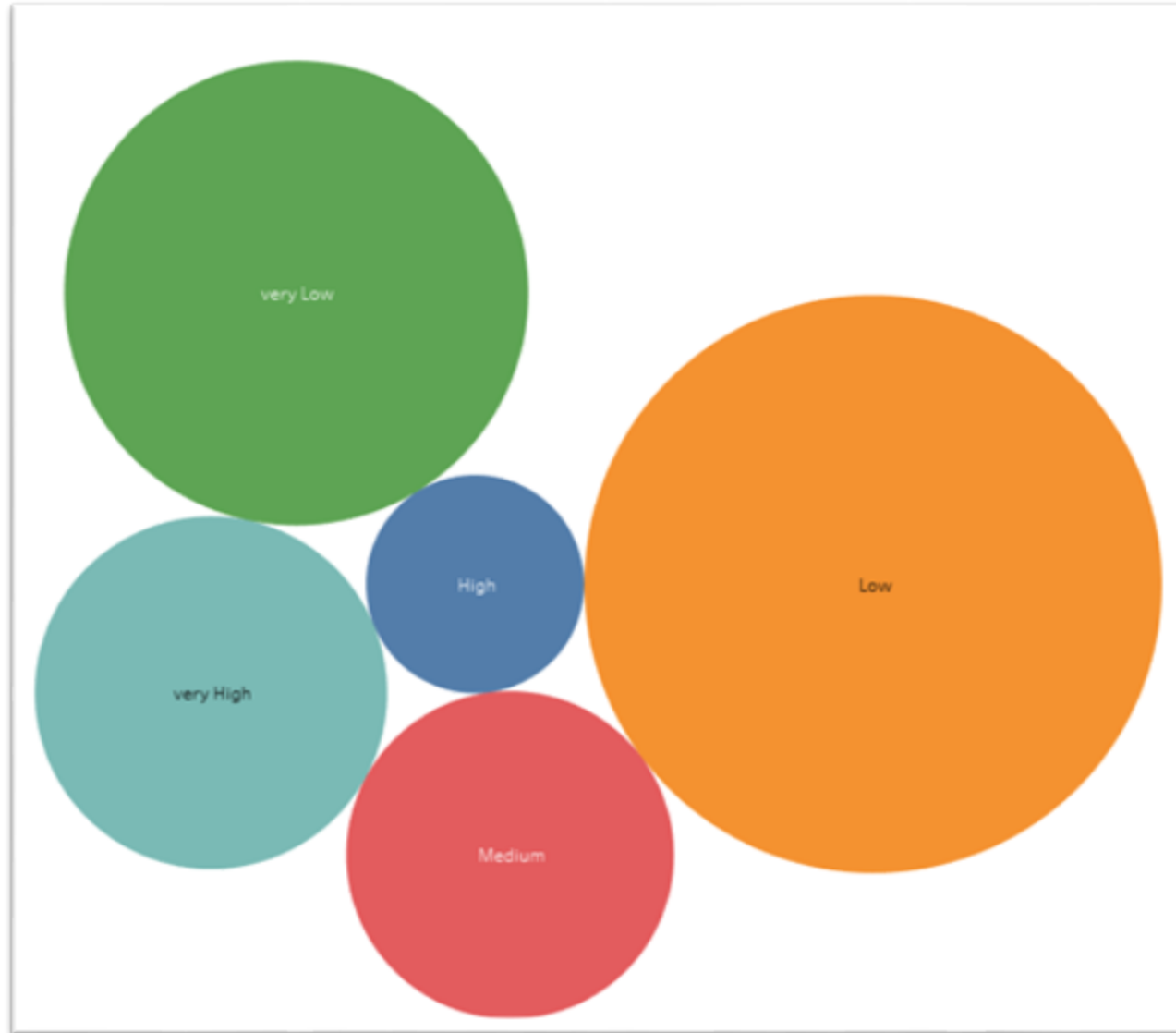


Neighborhood group percentages

Manhattan	44.301053
Brooklyn	41.116679
Queens	11.588097
Bronx	2.231312
Staten Island	0.762859

- 81 % of the listing are **Manhattan** and **Brooklyn** neighborhood group
- **Staten** Island has the lowest contribution.

# MINIMUM NIGHT CATEGORIES



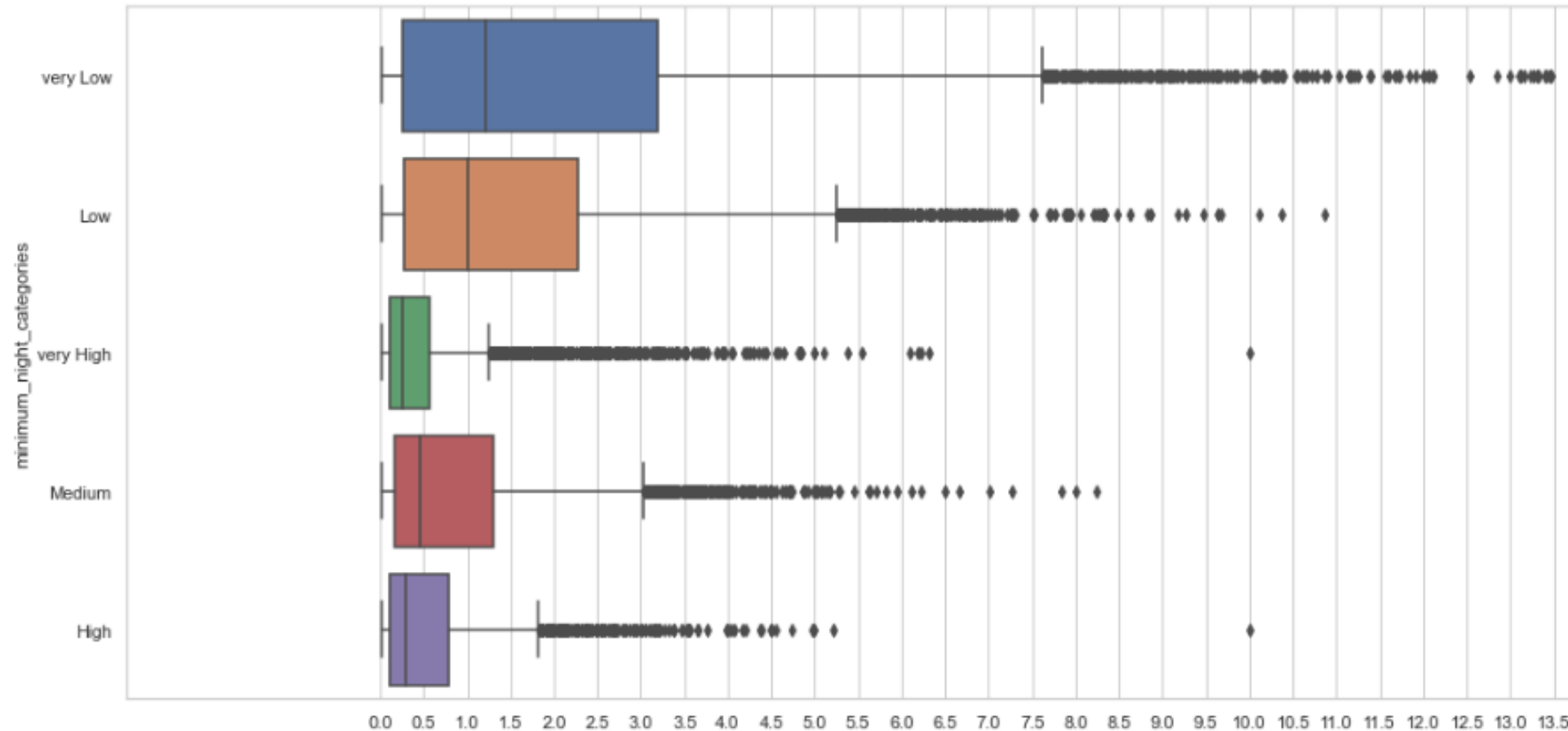
Minimum night category percentages

Low	40.280192
very Low	26.014930
very High	14.997444
Medium	12.960425
High	5.747009

- Low category in minimum night feature contributes 40 %



# EFFECT OF MINIMUM NIGHT ON REVIEWS



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

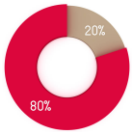
# CONCLUSION & RECOMMENDATIONS



Shared rooms need to be **inspected** upon.



The **cumulative contribution** of all hosts is better than a few hosts doing well.



More than 80 % of the listing are **Manhattan** and **Brooklyn** neighborhood group



Minimum nights threshold should be on the **lower side** to make properties more customer-oriented

# 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

- Conducted a thorough analysis of NewYork Airbnbs Dataset.
- Cleaned the data set using python.
- Derived the necessary features.
- Used group aggregation, pivot table and other statistical methods.
- Created 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 Variables:

- latitude
- longitude

## Time Varibale:

- last\_review