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

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

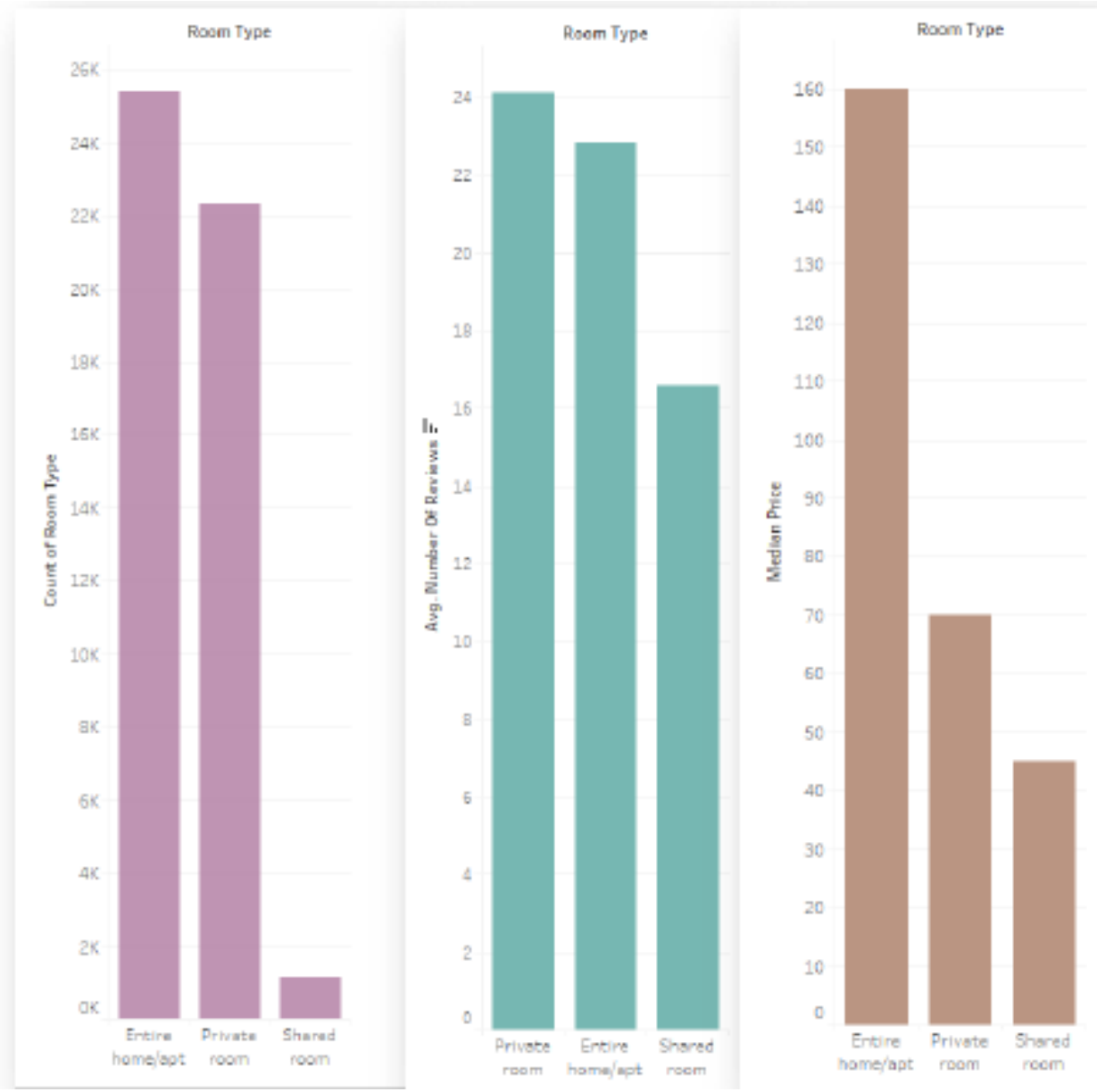
For the past few months, Airbnb has seen a major decline in revenue.

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

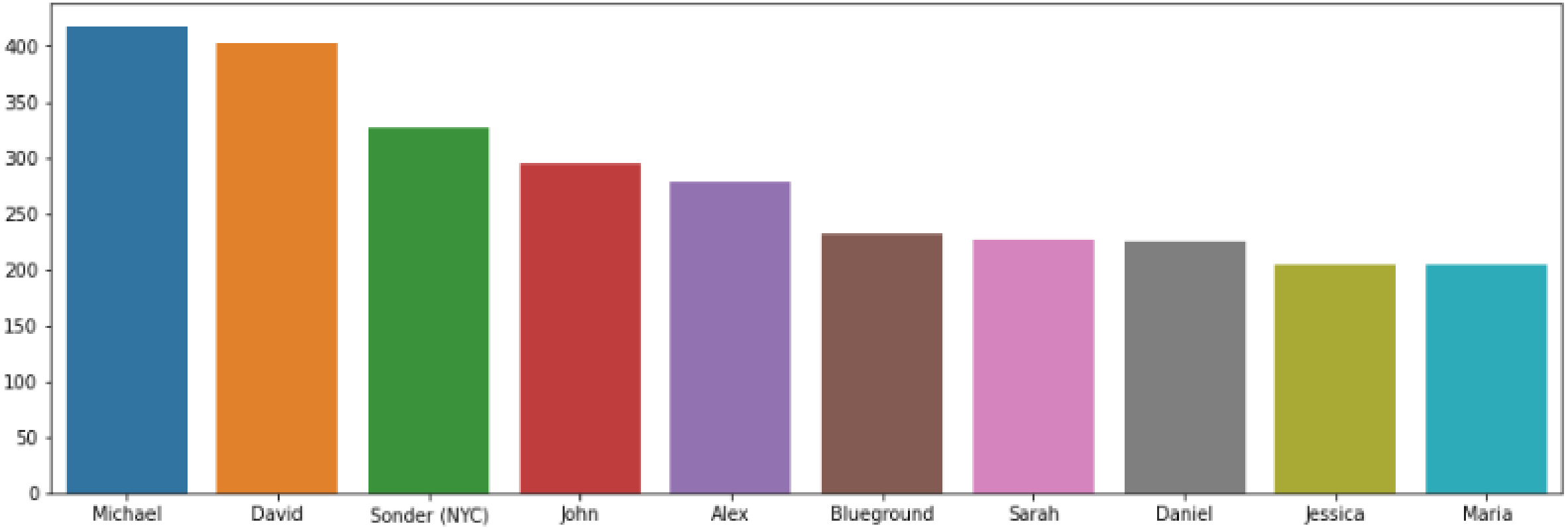
Airbnb wants to make sure that it is fully prepared for this change.

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.

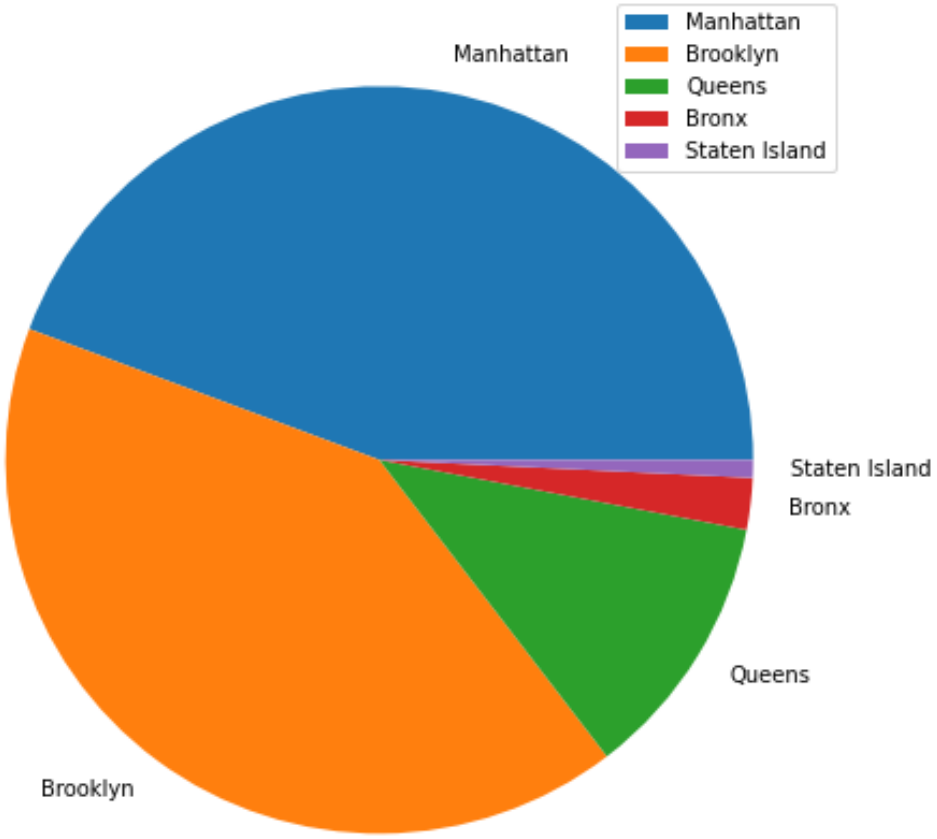


Top 10 HOST MATTER



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

MOST CONTRIBUTING NEIGHBOURHOODS



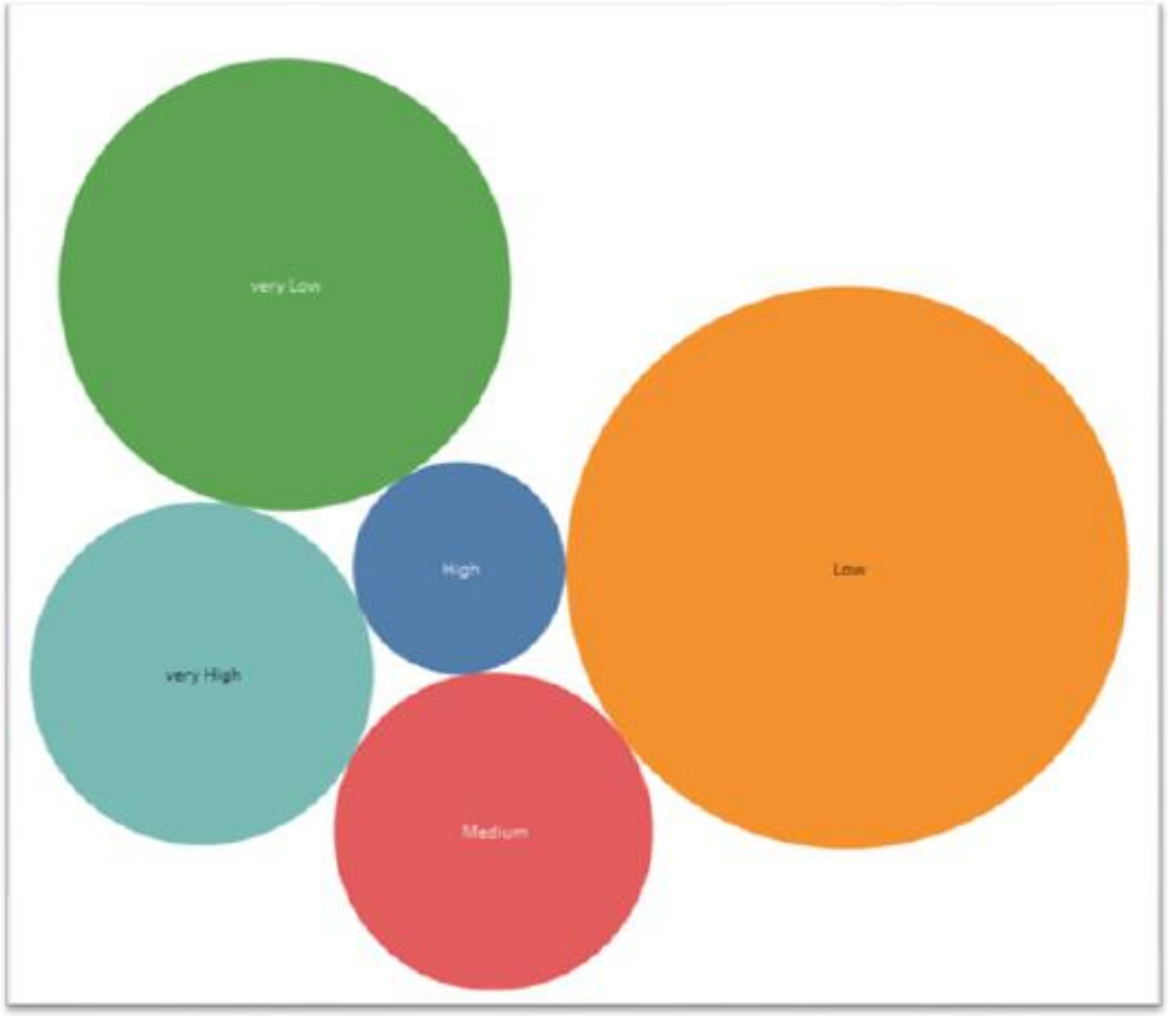
6.4 neighbourhood_group

```
inp0.neighbourhood_group.value_counts()
```

```
Manhattan        21661
Brooklyn         20104
Queens            5666
Bronx             1091
Staten Island      373
Name: neighbourhood_group, dtype: int64
```

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

MINIMUM NIGHT CATEGORIES

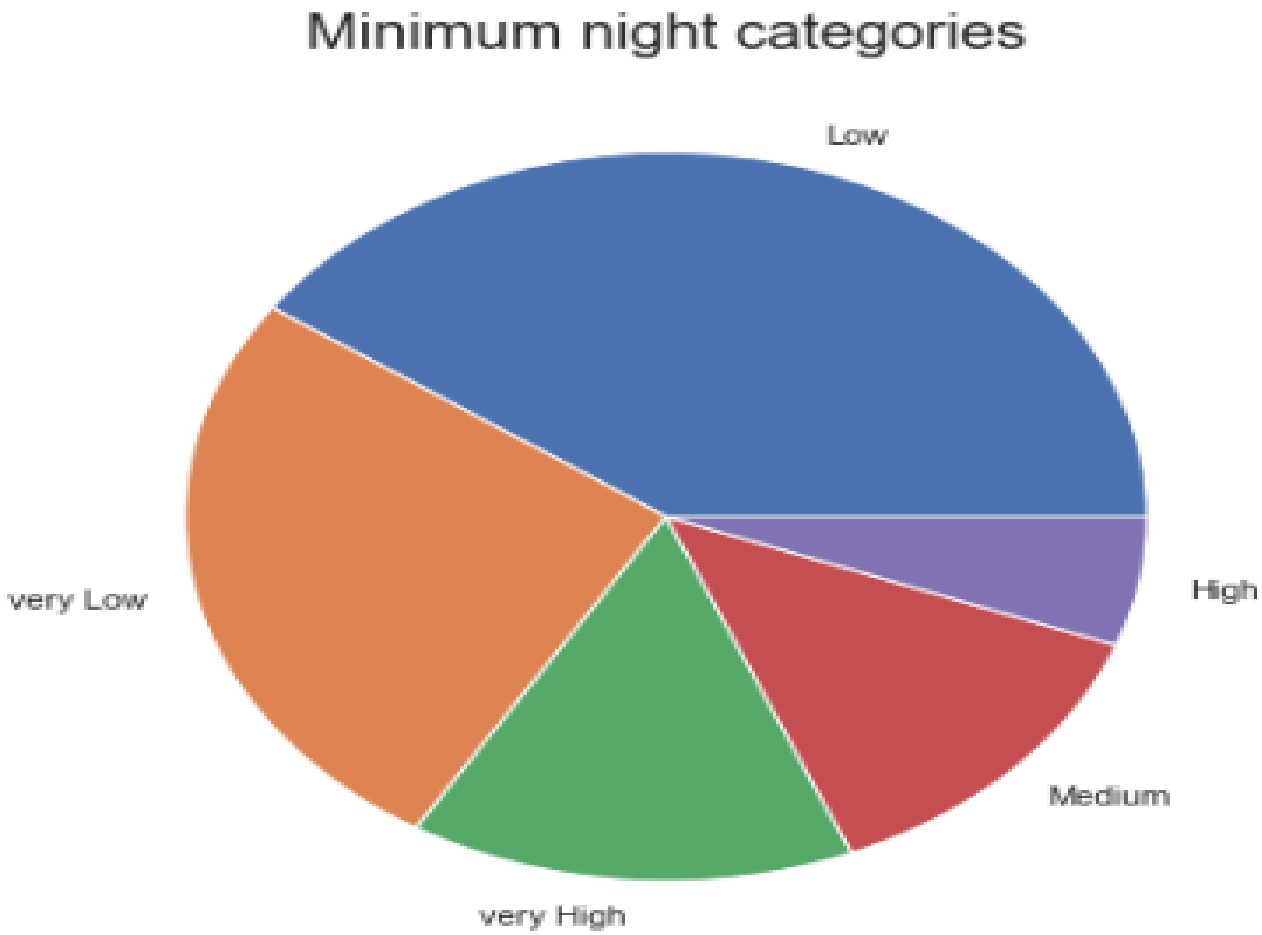


Minimum night category percentages

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 neighbourhood 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 usingTableau.

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
```