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

BY:
Shashwat bhansali

AGENDA

1) Objective

2) Background

3) Key findings

4) Recommendations

5) Appendix:

 a) Data sources
 b) Data methodology
 c) Data model assumptions

OBJECTIVE

1)To provide insights in to the current market situation.

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

**3)Provide early recommendations to our marketing and operations team.**

BACKGROUND

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

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

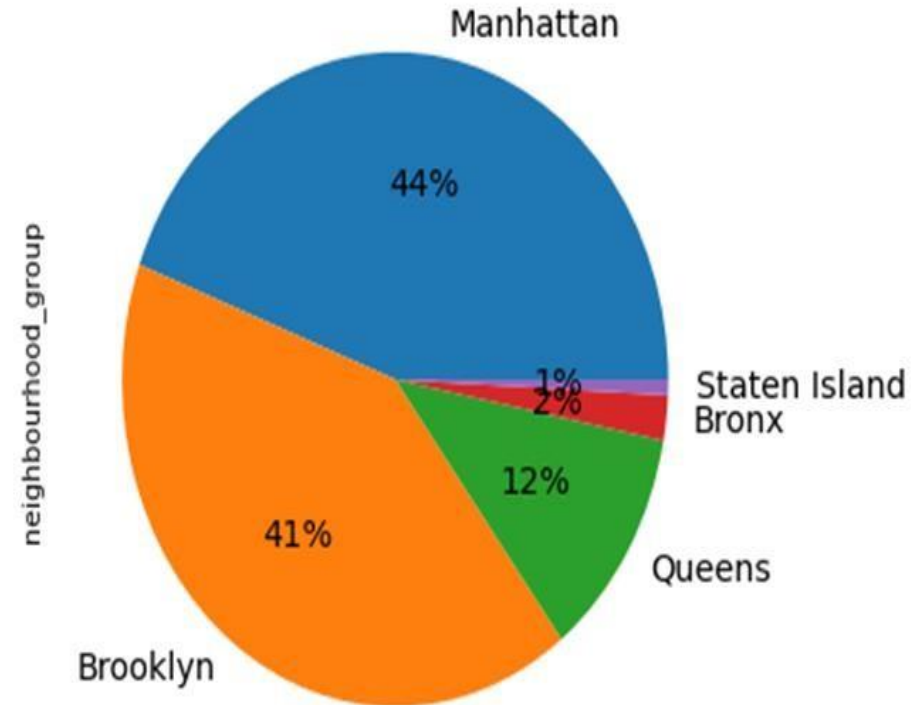3)Airbnb wants to make sure that it is fully prepared for this chance.

The Problems With The Shared Rooms

1) Shared rooms only a count for 2 % of the total types of the rooms.

2) They are less likely to be reviewed
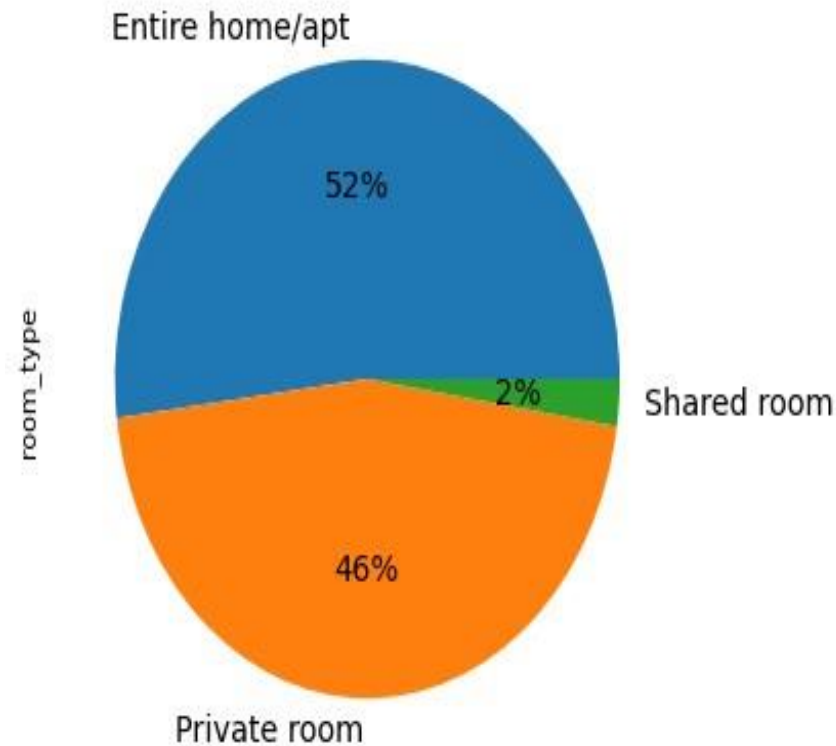
Most Contributing Neighbourhood

1) *85 % of the listing are from Manhattan and Brooklyn neighbourhood group.*

2) *Staten Island has the lowest contribution.*

Conclusion and Recommendations

1) Shared rooms need to inspected upon.

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

3)More than 80% of the listings are from Manhattan and Brooklyn neighbourhood group.

Appendix Data Sources

The columns in the dataset are self- explanatory. You 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 Methodologly

1)Conduct a thorough analysis of New York Airbnb dataset.

2) Cleaned the dataset using python.

3) Derived the necessary features.

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
```

THANK YOU