



AIRBNB CASE STUDY

BY: THIPPESH HK
SAPTHARSHI DAS

OBJECTIVE

- ❖ 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.
- ❖ So, analysis has been done on a dataset consisting of various Airbnb listings in New York
- ❖ Airbnb wants to make sure that it is fully prepared for post pandemic change.

Sonder(NYC) and Blueground contributes almost 50% of the total listings in the entire Newyork.

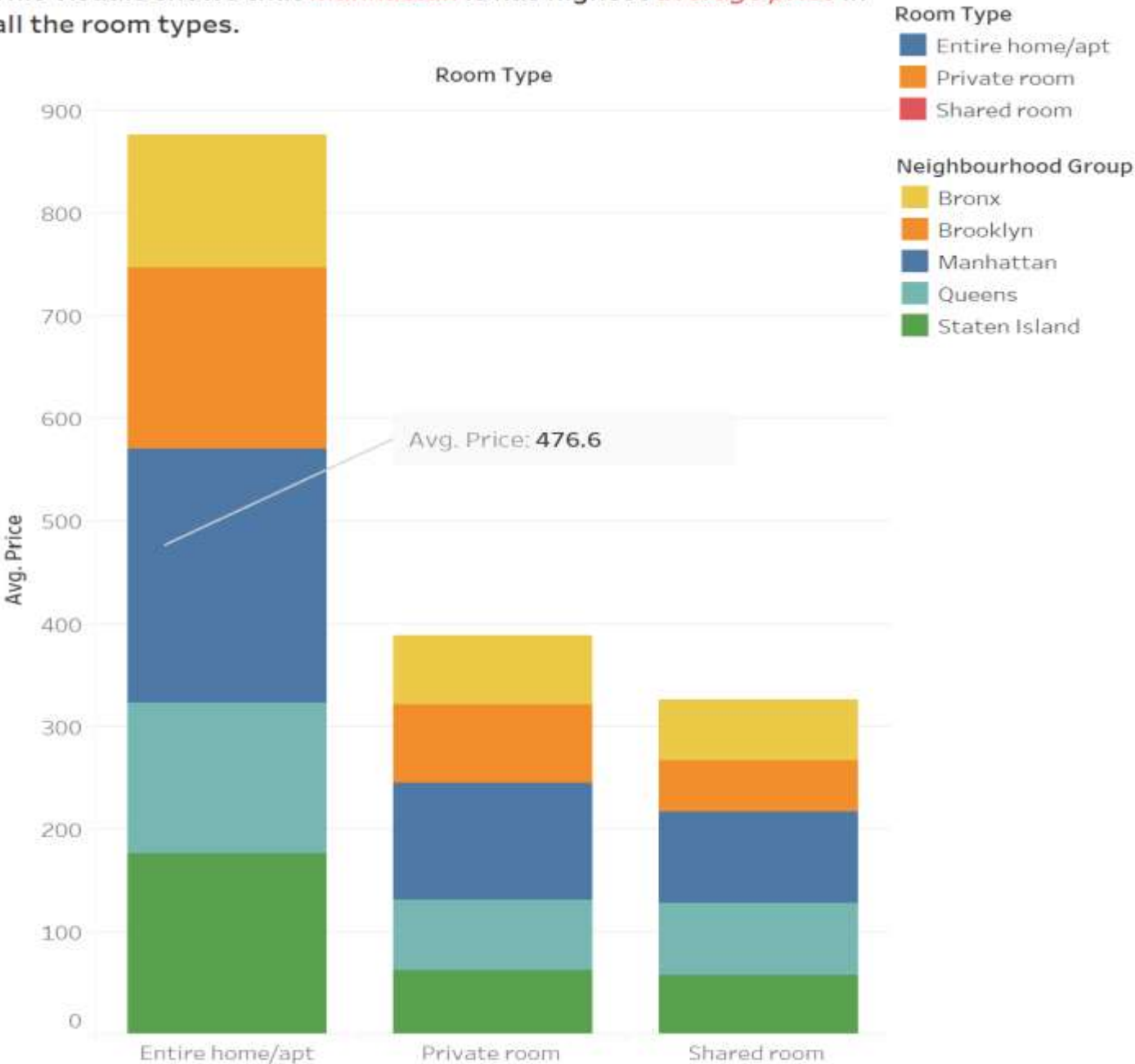
< The Treemap represents the entire listing holders in the Add a caption >

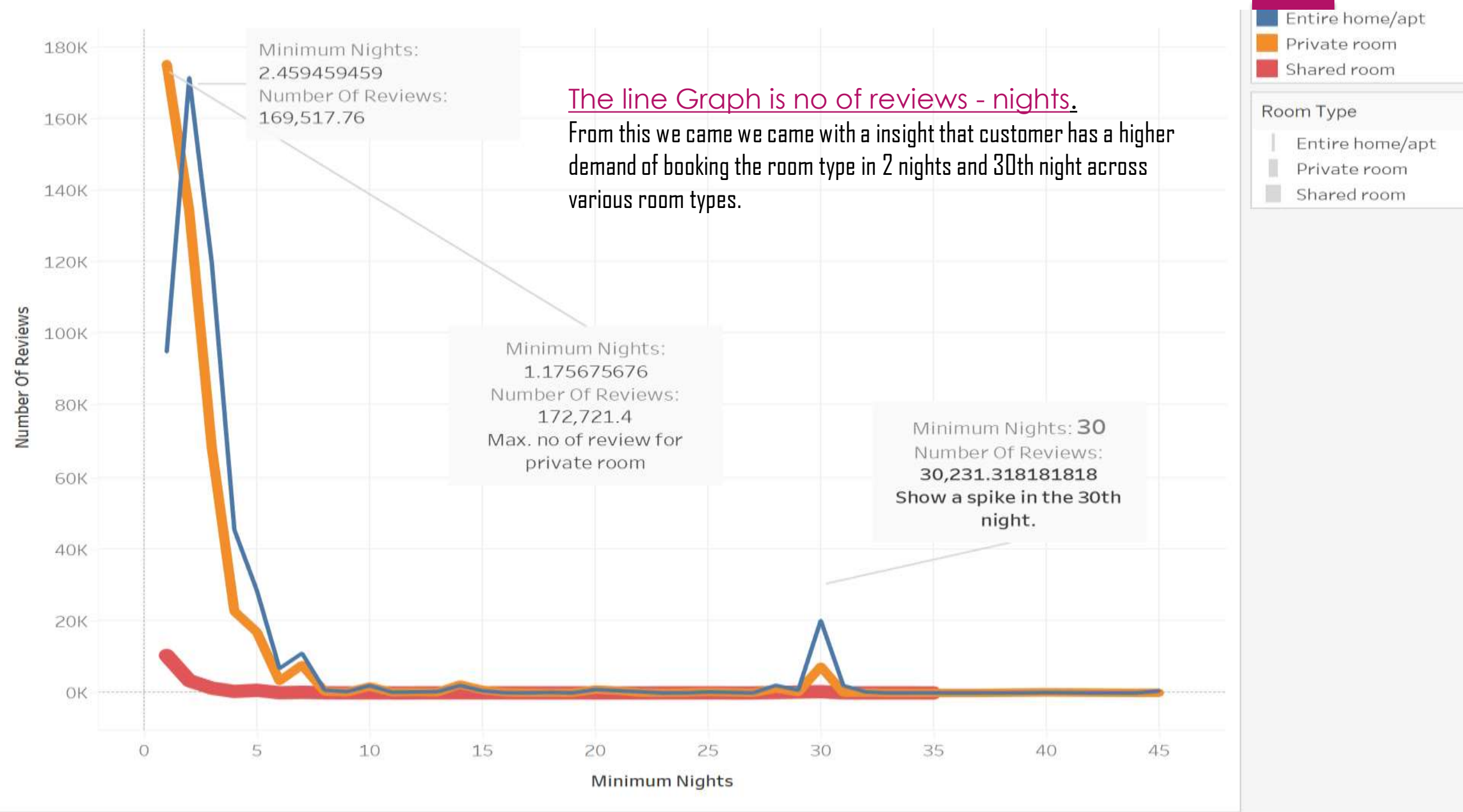


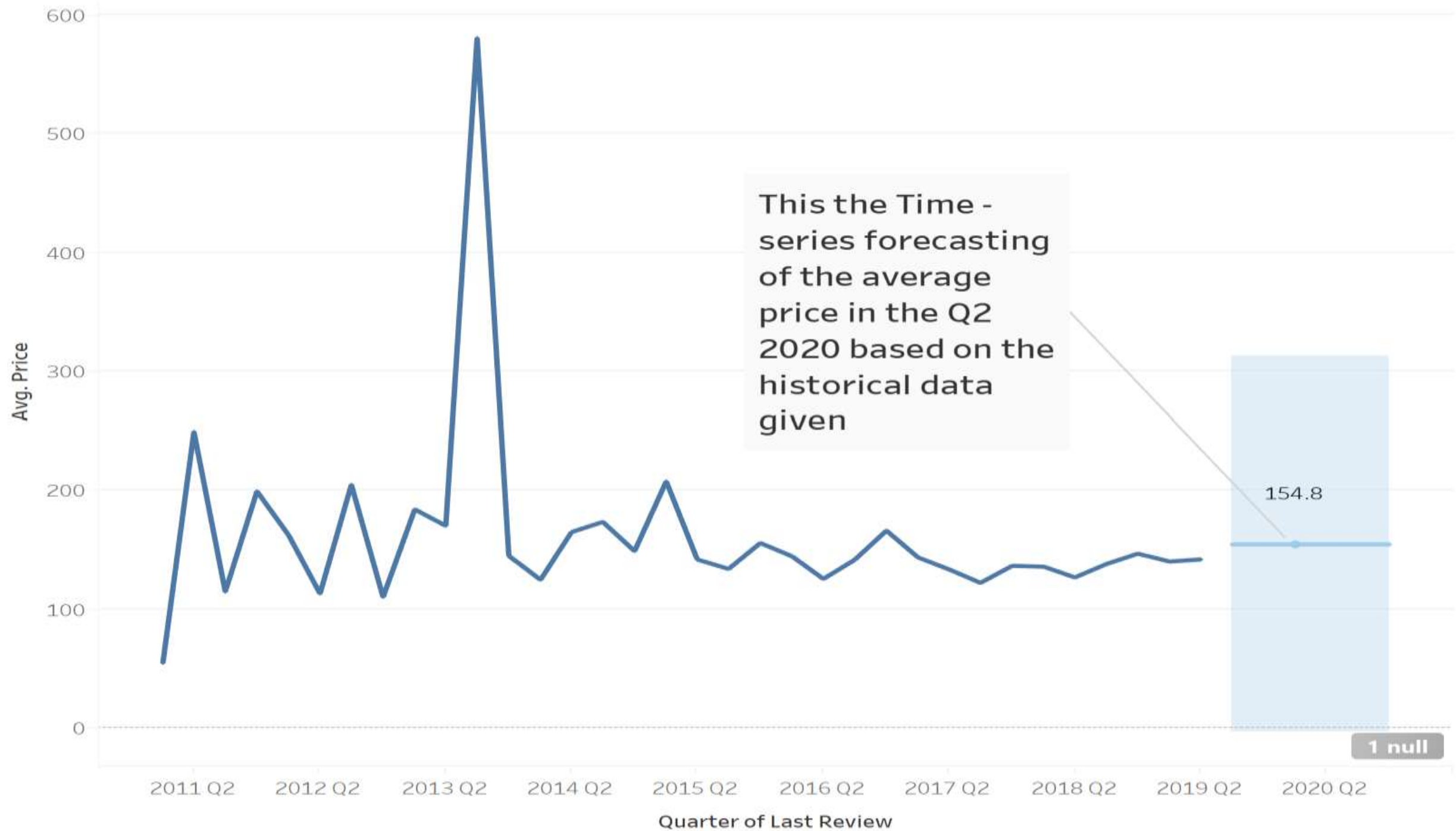
The visual shows that *Manhattan* has the highest average price (Entire home has the max. volume) following *Brooklyn* has the highest no. of reviews with balanced average price.



This visual shows that *Manhattan* has the highest average price in all the room types.









Map shows the top 15 neighbourhoods with most No. of hosts.

Row Labels	Count of host_name
Williamsburg	3830
Bedford-Stuyvesant	3628
Harlem	2585
Bushwick	2410
Upper West Side	1920
Hell's Kitchen	1900
East Village	1801
Upper East Side	1769
Crown Heights	1531
Midtown	1505
Greenpoint	1099
East Harlem	1080





0, 365, 364, 1, 5, 3, 2 are the availabilities which are maximum review and liked and hosted by customers.

Customer_Seg Price_Range

medium

>69\$ - <150\$

low

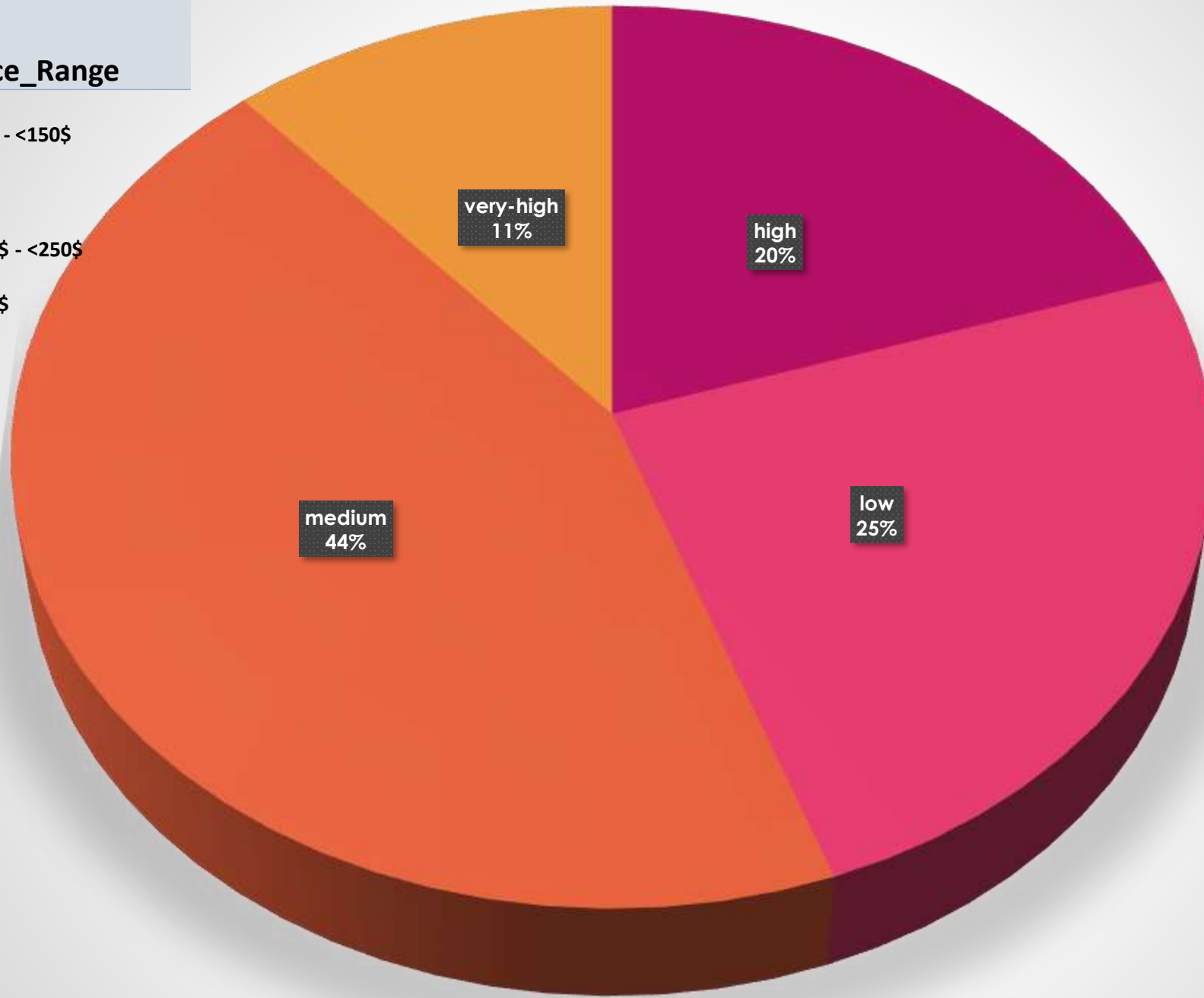
<69\$

high

>150\$ - <250\$

very-high

>250\$



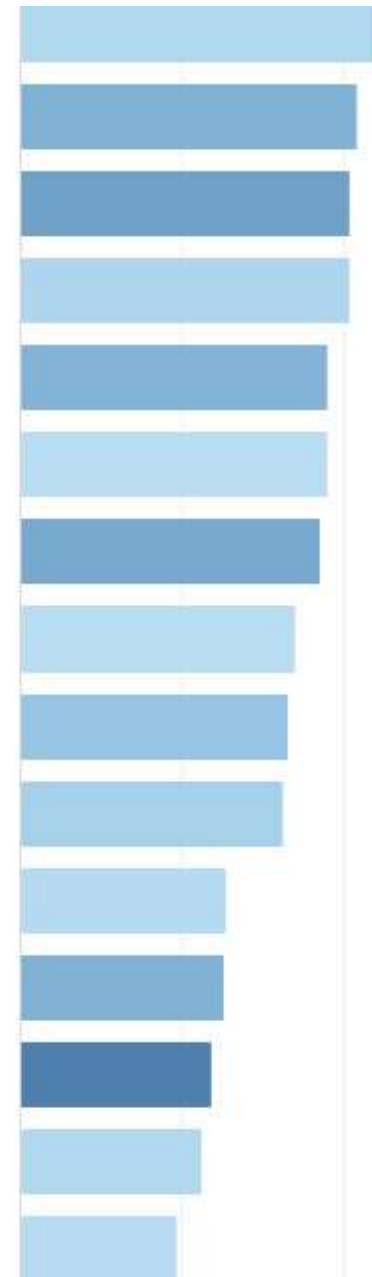
high
low
medium
very-high

Downtown Brooklyn
South Slope
Red Hook
Williamsburg
Windsor Terrace
Greenpoint
Brighton Beach
Cypress Hills
Flatlands
Coney Island
Bay Ridge
Sunset Park
Crown Heights
Bedford-Stuyvesant
Bergen Beach



**Right side are the
Neighbours in Brooklyn and
left side is Manhattan
The deeper shows are the
Places which has very low
Average price with
High no. of reviews
(the price can be improved/
Raised depending on the demand**

Murray Hill
Upper West Side
Hell's Kitchen
Kips Bay
Upper East Side
Midtown Center
East Village
Stuyvesant Town
Lower East Side
Chinatown
Two Bridges
East Harlem
Harlem
Morningside Heights
Roosevelt Island



RECOMMENDATIONS

- ▶ Under the host's segmentations **Sonder(NYC), Blueground , Kara , Ruchi, Corporate housing** are the most important host which contributes 50% of the all listing in Newyork used be the prime customers, Airbnb should provide some extra benefits to them.
- ▶ **Manhattan** and **Brooklyn** are the most popular places with the average price higher compared to other neighbours. **Brooklyn** should be provided with more no of listings as the demand is very high.
- ▶ **Private homes** are the most popular room types with a very high average price , in **Manhattan**, there should be a price cut with a competating prices with all the neighbourhoods.
- ▶ **2 days and 30th day** minimum stays are most popular among all the available. The marketing team should specially advertise and focus more.

RECOMMENDATIONS

- ▶ There is forecasted of **154\$** price average for the upcoming quarters of 2020.
- ▶ **Williamsburg, Bedford-Stuyvesant, Harlem, Bushwick, Upper West Side, Hell's Kitchen, Crown Heights** are the top 5 neighbourhood which has most no of hosts and well demanded.
- ▶ Properties which has a availability of **0, 365, 364, 3, 5, 2** days in the listing has a maximum demand.
- ▶ Almost **44%** of the customer falls within the price range of **70 – 150\$**.
- ▶ **Williamsburg, Bedford, Harlem** these are the certain places with maximum reviews and demand but average price is comparatively low, recommend on increasing as there is a higher demand in these areas.

METHODOLOGY

EDA :-

1. First the data set was imported in excel , the row and columns formatting was done.
2. Then for basic data cleaning , missing values, outliers treatment was done.
3. Name – 16, host_name – 21, last_review and review_month – 10052 values were missing , deleting certain rows where names were missing as its very low compared to total rows in data set
4. The date columns were imputed with mode method.

METHODOLOGY

5. Then outliers fixing was done price, min_nights , no of reviews, availability, calculated hostings has certain values which are extremely high.

6. Applied quartile cut off methods and set a cut off < 98% percentile and deleted the values to make it a normal distribution.

Did descriptive analysis

8. Certain data table which are there in visuals are extracted from Excel pivot.

9 . All the visual were made on tableau desktop.

10 . The python script link in provided for better reference.



THANK YOU

Saptarshi Das
and
Thippesh HK