Analysis of AirBnB Business for the City of New York and Future Steps for Expansion

-by

Srishti Sri

Sushil Kumar

Vishnuvardhan Pinjala

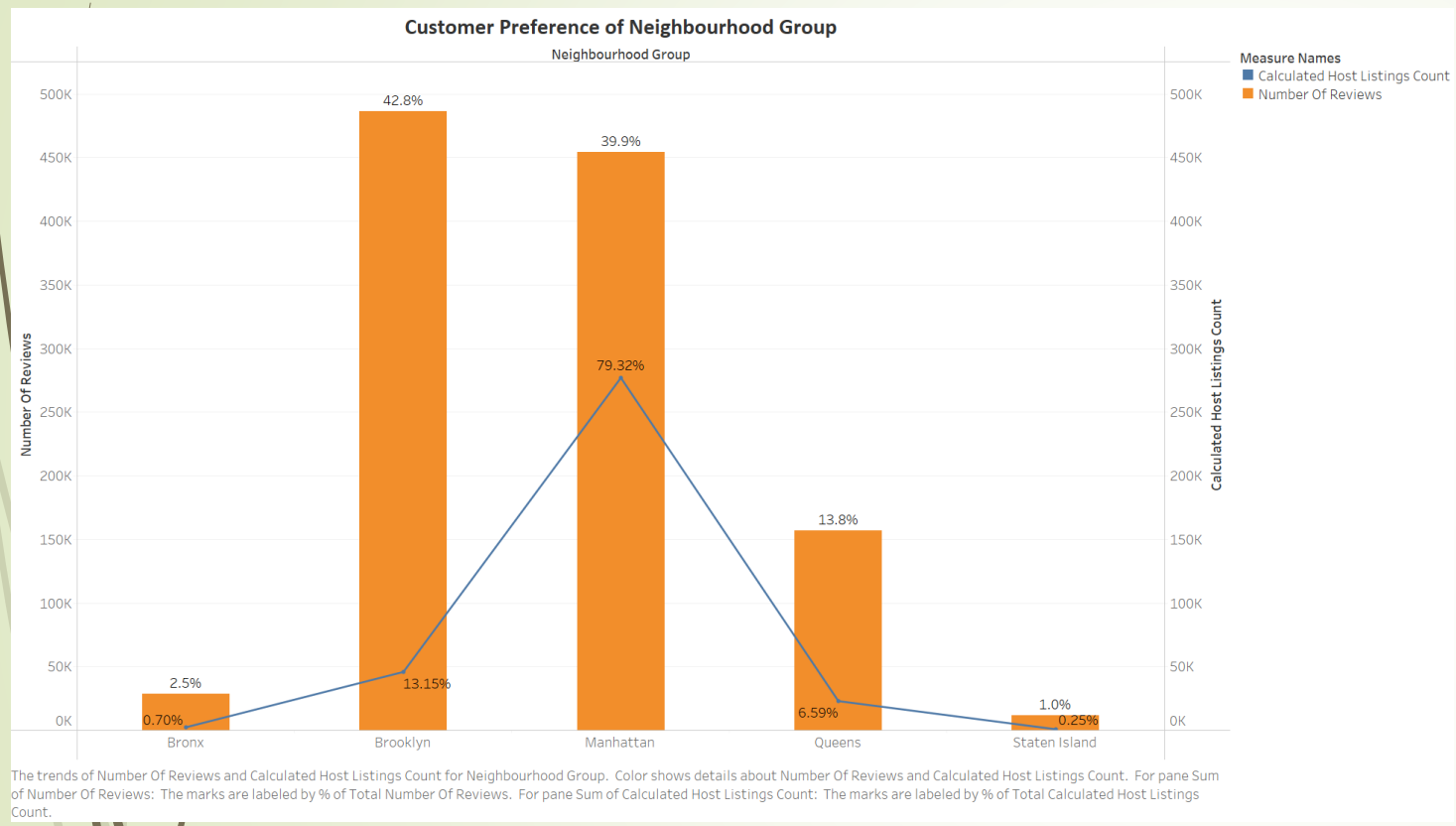
Agenda

- Improve Revenue Generation for AirBnB New York City
- Key to improve Revenue in coming years is to identify trends in the operations to date.
- Classify Customers on Preferences, Identify Profitable Neighbourhoods.
- Steps to be taken to achieve said goal.

Future Steps to Raise Revenue Generation

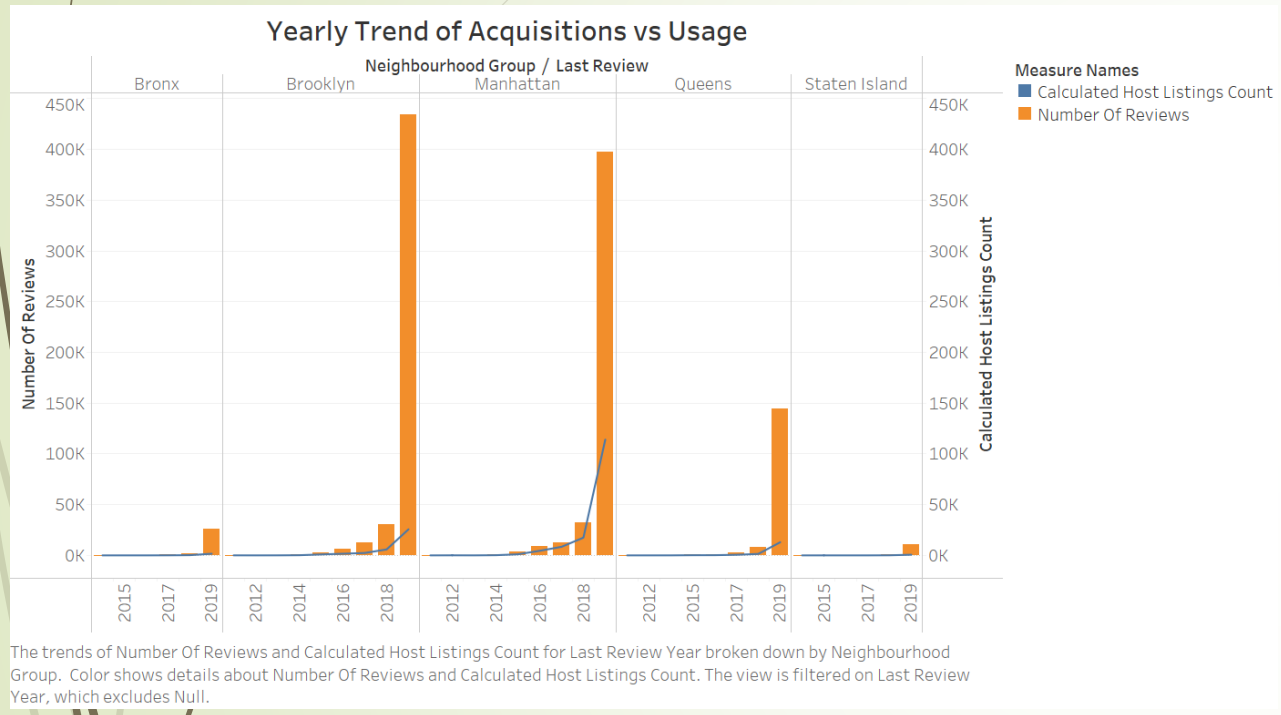
- Acquire more Entire home/apt and Private Rooms in Brooklyn
- Decrease the availability of Entire home/apt and Private Rooms in Brooklyn
- Increase the availability of Entire home /apt and Private Rooms in Manhattan
- Let go of Listings of type Shared Rooms in all Districts.

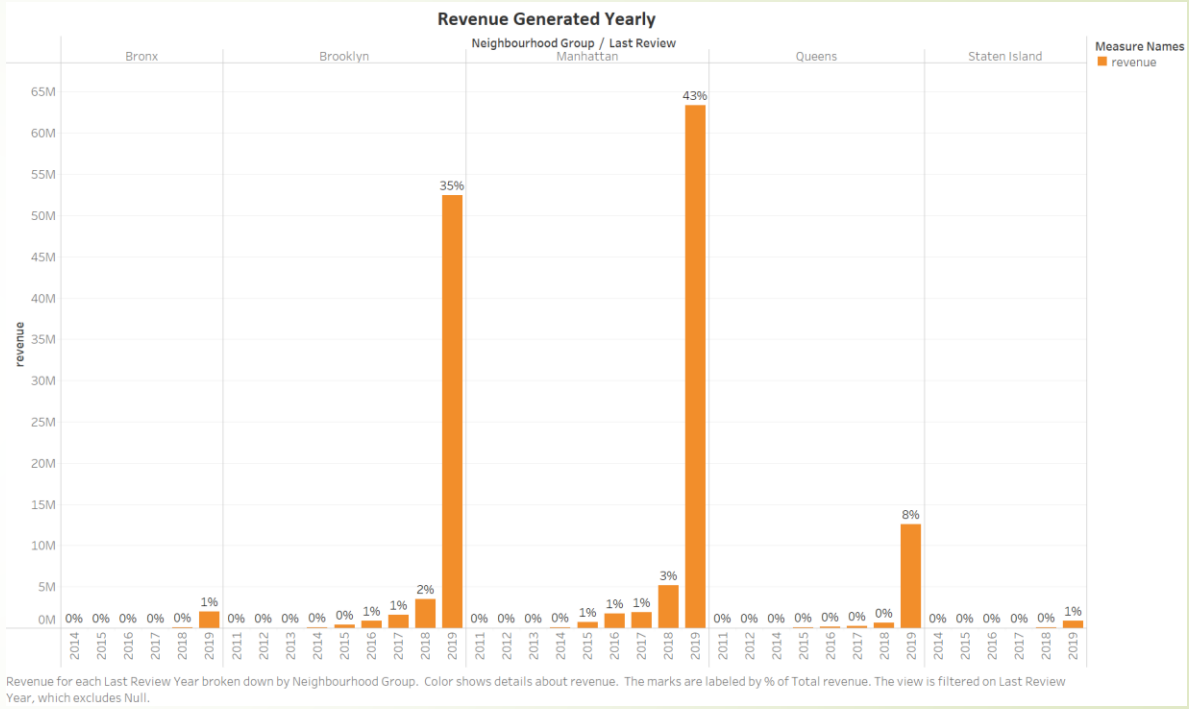
Trends in Operations



- ***Manhattan*** lists nearly ***80%*** of all the listings available with AirBnB
- ***Brooklyn*** has seen ***highest*** number of ***customers*** despite accounting for only 13% of listings.

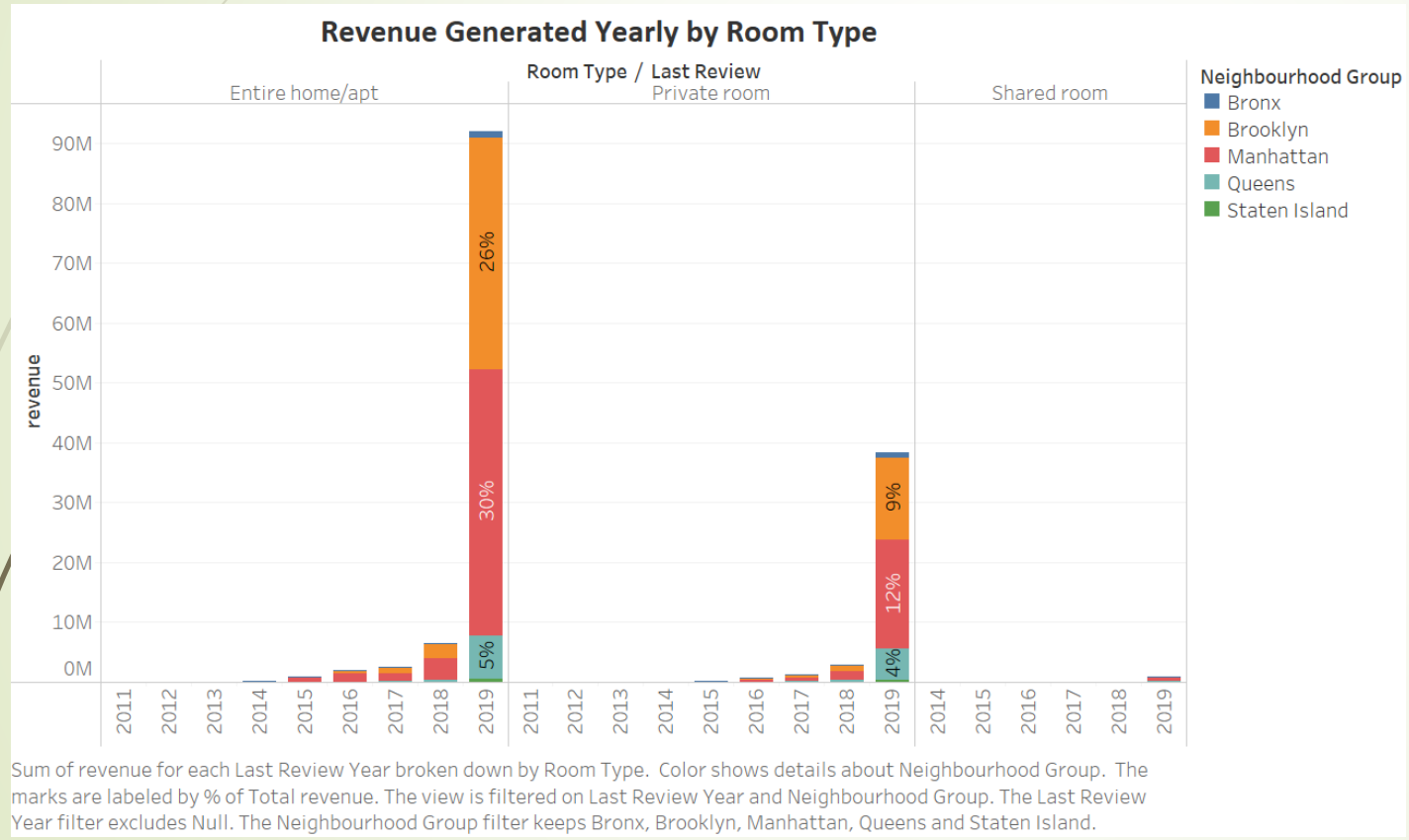
Contd.





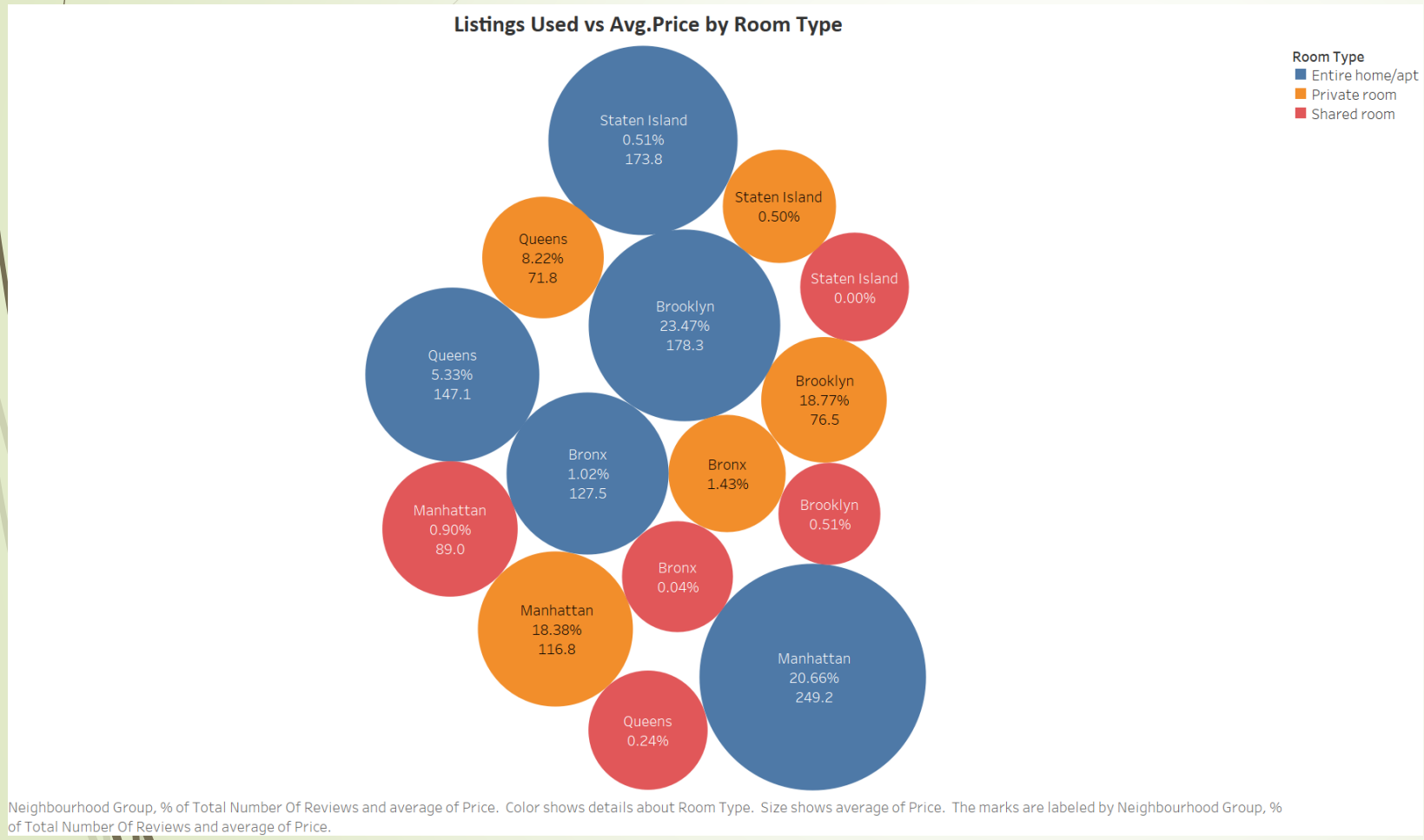
- Spike in operations in year 2019 with acquisitions and bookings with general increasing graph for both.
- Most Revenue generated from Manhattan and Brooklyn with increasing graph for this trend too.

Contd.



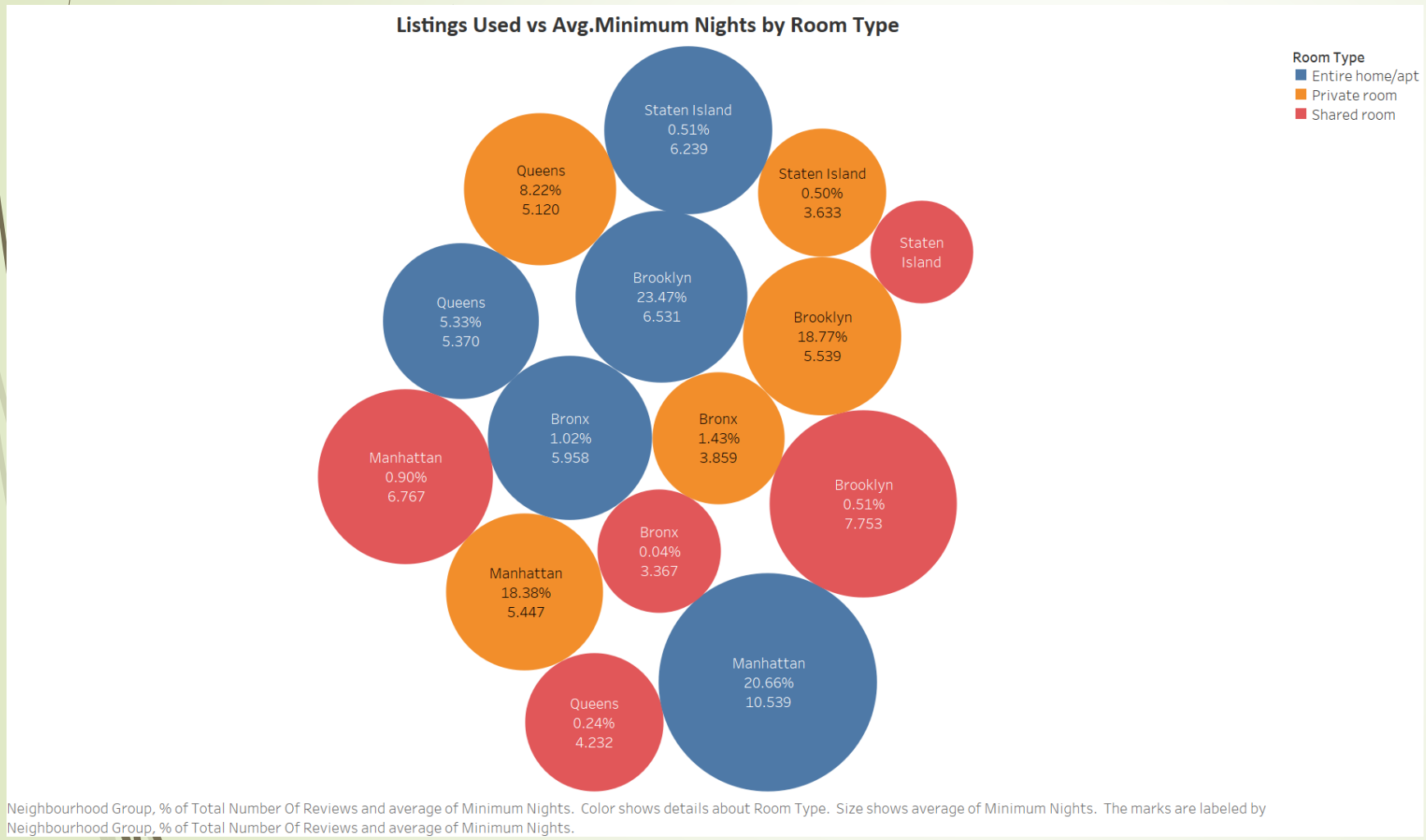
- ***Entire Homes*** lead the pack in Revenue Generation by Room Type.
- While ***Manhattan*** leads Revenue Generation for Neighbourhoods.

Customer Preferences



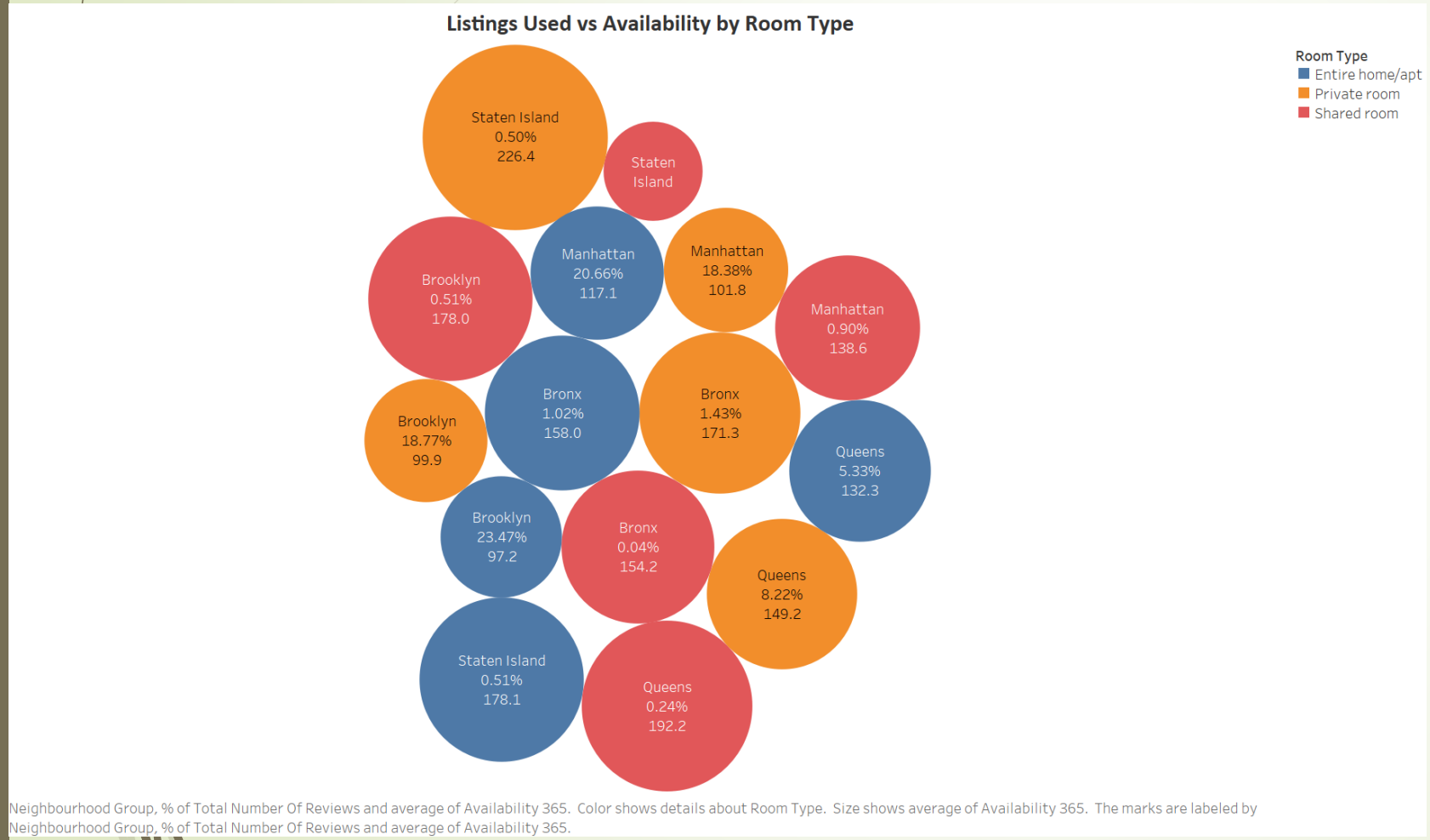
- The Chart here depicts variation in pricing of listing type across Neighbourhoods.
- From here we can gather information that despite high prices customers prefer **Entire homes** followed by **Private Rooms**.
- **Brooklyn** leads as preferred Neighbourhood.

Contd.



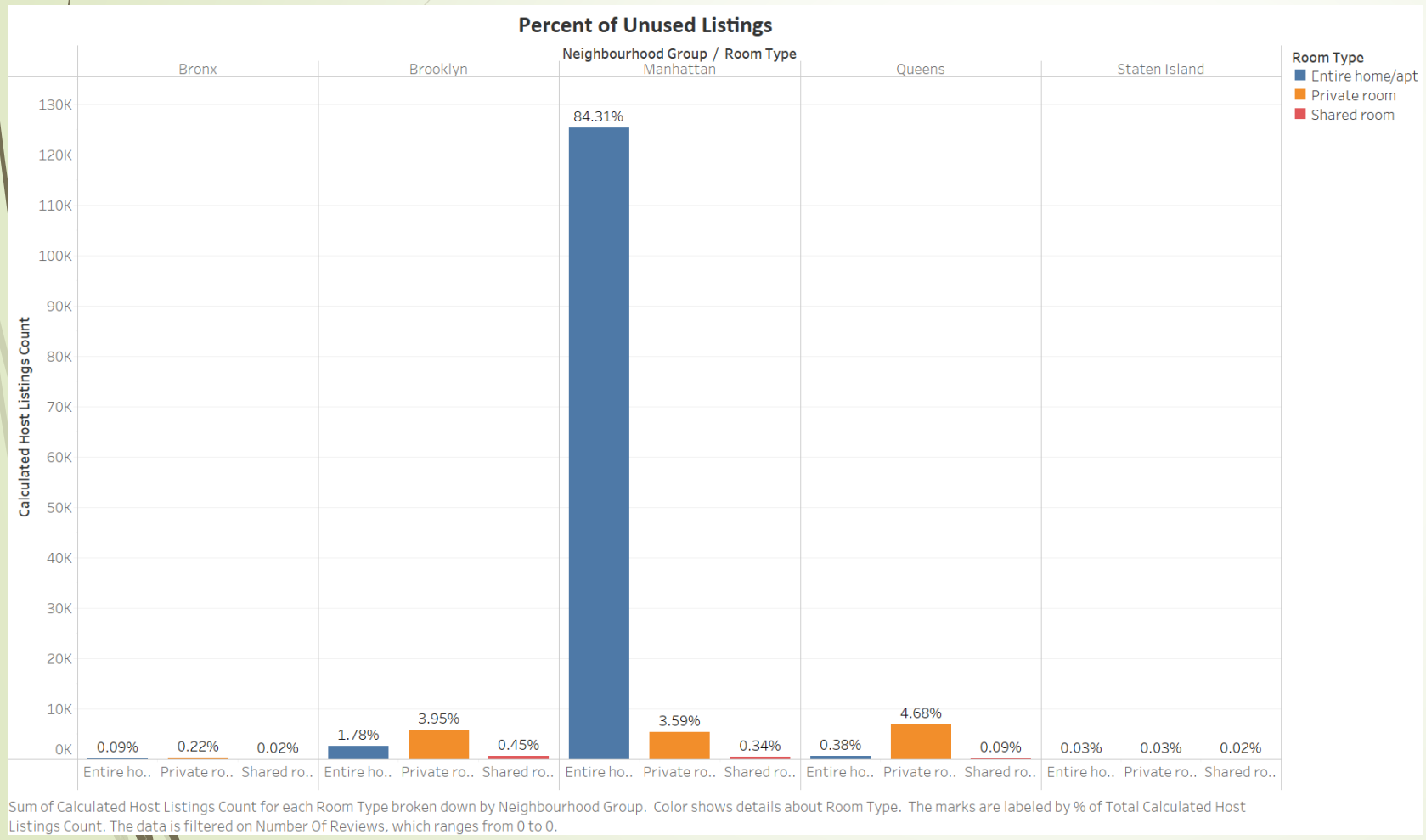
- From This chart it is clearly inferred that Minimum Nights has no bearing towards Customer's influence in bookings.

Adjustment of Booking Availability



- From this chart we can see that **Brooklyn** has **lowest availability** for bookings.
- It can also be understood that non-preferred listings are available for the most time yet have failed to attract customers.
- So further ***decreasing availability in Brooklyn and increasing availability in Manhattan*** will nudge customers towards Manhattan without affecting booking in Brooklyn.

Acquiring Further Listings



Sum of Calculated Host Listings Count for each Room Type broken down by Neighbourhood Group. Color shows details about Room Type. The marks are labeled by % of Total Calculated Host Listings Count. The data is filtered on Number Of Reviews, which ranges from 0 to 0.

- We can see rom this analysis that ***over 120K listings of Entire homes in Manhattan have not generated any revenue***.
- So based on the Customer Preferences seen earlier ***Further Acquisitions of Entire Homes should be targeted in Brooklyn***.

Conclusions

- To improve Revenue for upcoming years focus heavily in Neighbourhoods of Brooklyn and Manhattan.
- In Brooklyn acquire more Entire Homes and decrease the availability for bookings.
- In Manhattan acquiring further listings to be prevented and increase in availability for booking.
- Nos of Listings for shared room to be decreased as they are unappealing to customers.
- Reasons for appeal in Neighbourhoods of Bronx, Staten Island and Queens to be investigated as Data insufficient for Analysis.

Appendix

- DATA CLEANING

- In the Dataset columns (name and host_name) contain blanks. These has been filled with Unnamed and Noname respectively as these datapoints need not be unique.
- Also columns reviews_per_month and last_review contain blanks. Further analysis shows that they coincide with column number_of_reviews having data entry as 0. So they can be left as is.

- DATA WRANGLING

- Column Price is taken as price per booking instead of per night.
- Column Number_of_Reviews is taken as customers and assumed that all customers have posted reviews so Number of Reviews is equated as number of customers.
- New Column Revenue Generated by multiplying data point in Number_of_Reviews with data point in price. Logic being evident from previous points.

- CHARTS

- Bar Charts provide the best visuals for most of the insights gathered from the data.
- Wherever dual-axis charts are used both axis have been synchronized.
- Labels on the Bars in Bar charts display data as percent of total.
- In all Bubble charts the size of the bubble displays the varying characteristics like average price, average availability and average minimum nights.
- In all Bubble charts, color has been used to differentiate between room type, while labels contain neighbourhood name, percent of total visitors for that grouping and varying characteristics.
- In time series data Null data display has been excluded as they inform about unvisited listings.
- For looking into data about unused listings a filter of 0 number of reviews was set to fetch only listings data where number of reviews is set as zero.