

Capstone Project

Airbnb Booking Analysis

Sangamesh Chandankera

Points for Discussion

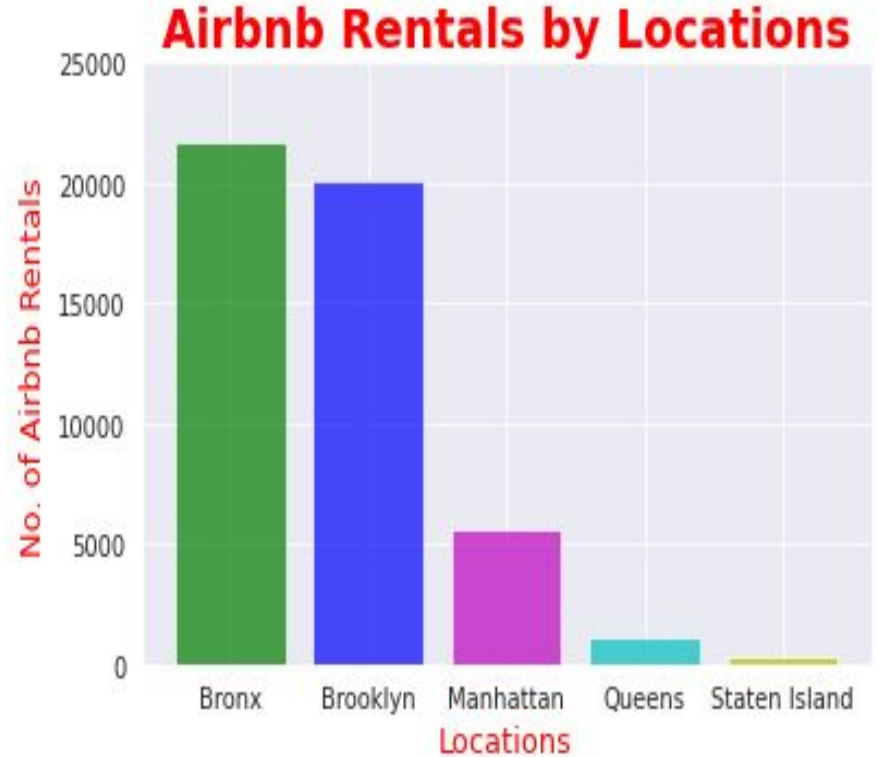
- Data Summary
- Location-wise Airbnb rentals
- Type of rental rooms
- Price variations among different locations
- Pricing based on room category
- Availability of rooms over the year
- Top hosts id's
- Best and top of the class neighbourhoods
- Spread over map using Latitude & Longitude

Data Summary

- **id & host_id** : 'id' holds a unique number for each row whereas 'host_id' holds the unique number code for each host.
- **neighbourhood_group** : It contains the location information of the rentals.
- **neighbourhood** : This contains the sub-locations of the neighbourhood_group.
- **host_name** : It contains the names of the many hosts across the different neighbourhoods.
- **room_type** : It contains the category rooms among the rentals.
- **price** : It contains all the price information of different rooms of different places.
- **number_of_reviews** : It contains the number of reviews.
- **availability_365** : It contains the information about the availability of the rooms throughout the year.
- **latitude** : It contains the latitudinal coordinates.
- **longitude** : It contains the longitudinal coordinates.

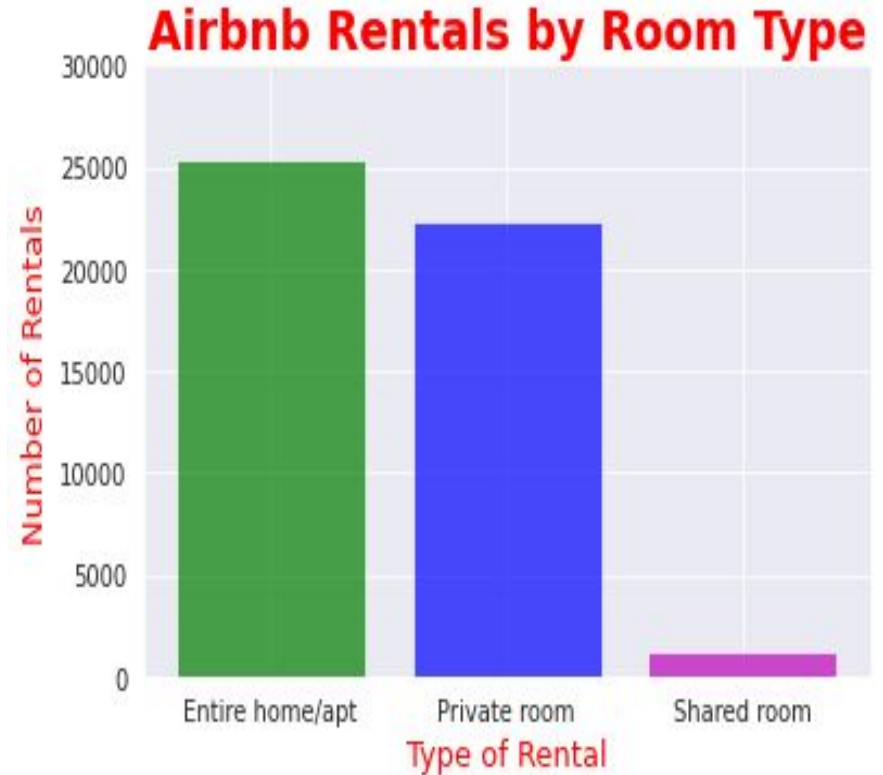
Airbnb location-wise listed rentals

- The bar-graph on the right represents the number of rentals from the given different locations listed on Airbnb.
- The x-axis represents the locations and the y-axis features the number of listed rentals.
- The majority of listed Airbnb rentals are in the neighbourhood of 'Bronx' & 'Brooklyn' touching the mark of 20,000.
- The lowest number of Airbnb rentals are in the neighbourhood of 'Queens' & 'Staten Island' ranging less than less than 1,000.



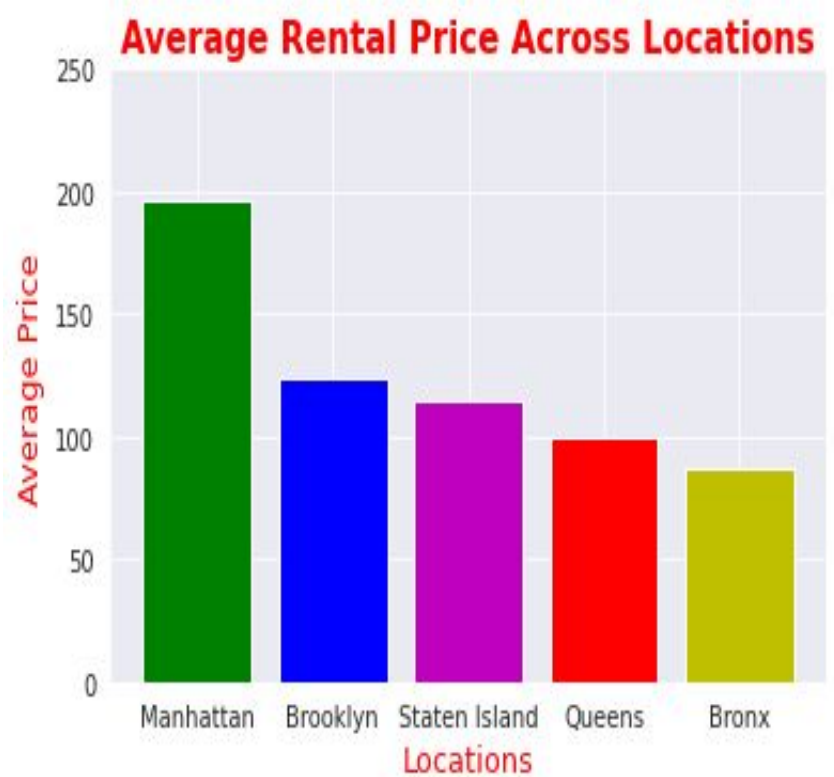
Rentals Rooms Categories/Types

- This particular bar-graph on the right gives us an idea of the number of different types of rental rooms that are listed on Airbnb.
- The x-axis represents the type of the rental rooms whereas; the y-axis features the total number of rooms/rentals.
- It is clear from the bar-graph that the most in demand and numbers are the 'Private rooms' & 'Entire home/apartment' that are in the range of numbers 20,000 - 26,000 each.
- Number of shared rooms are in the range less than 1,000. The reason associated to this maybe security and no privacy. Hence; there are very few demands for shared rooms.



Average price of different types of rooms across different locations

- This particular bar-graph provides us with the information of different average price ranges per night for different locations provided in the data.
- Conclusively 'Manhattan' has the highest average price compared to other locations for per night. Price ranging from 150-200 dollars.
- 'Bronx' location provides with the cheapest of all rental room listed in the Airbnb data, with an average price range of 50-100 dollars.
- The rest three locations have around the same average price, putting them all in the same range of 100-150 dollars.
- Also; we have assigned a fixed colour for each locations and they are what shown in the graph.

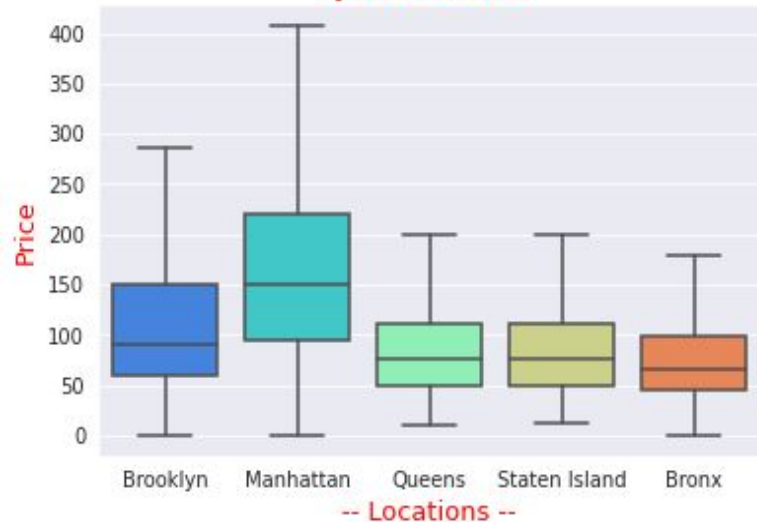


- There were problems in calculations of the average prices over the different provided location due to some outliers of top of the class rooms/rentals whose prices were too expensive.
- Therefore; we will be showing below two boxplots : one with outliers(on the left) and the other without outliers(on the right).

**Price Distribution Across neighbourhood_group
price < 500**



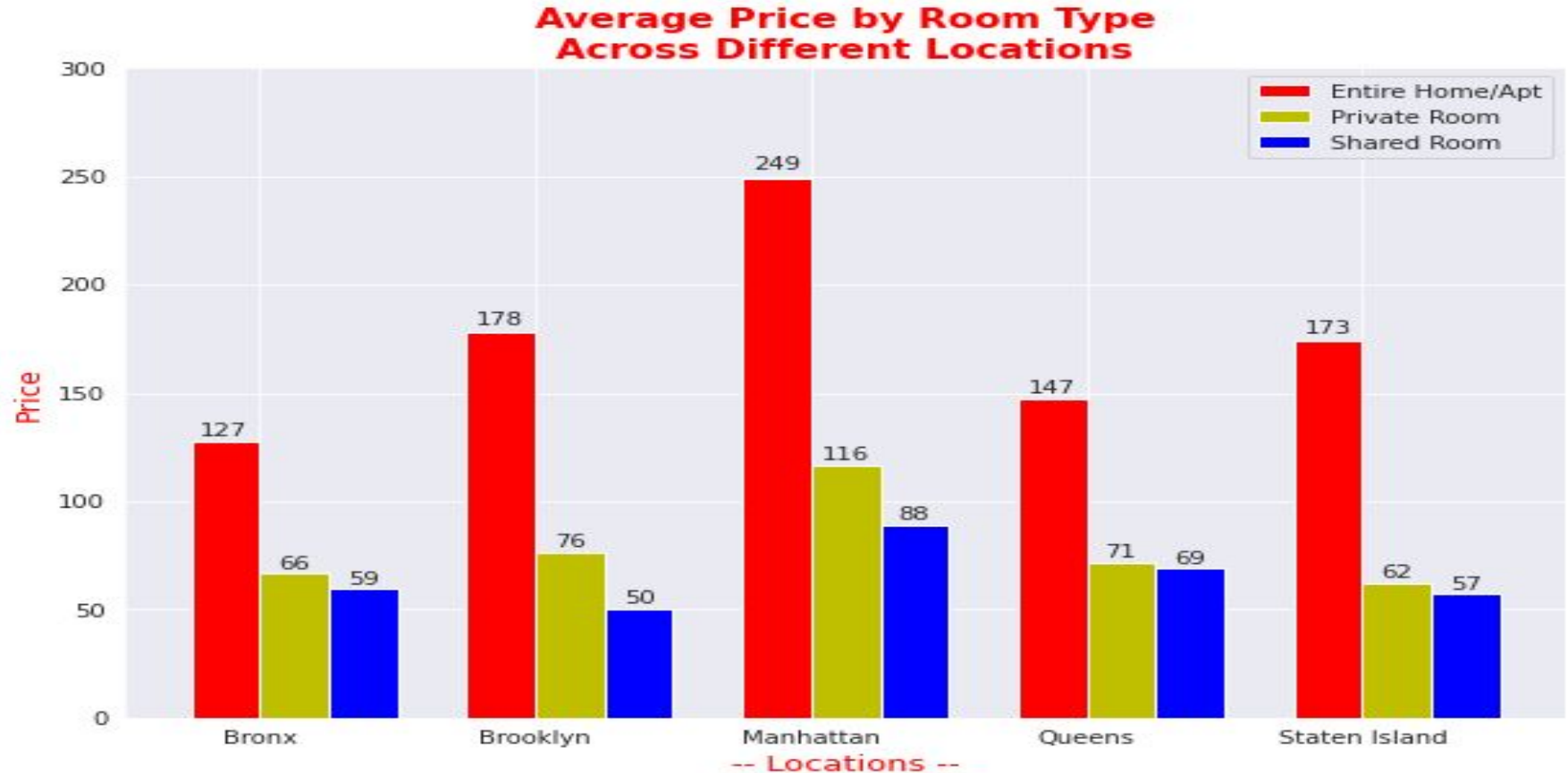
**Price Distribution Across Different Locations
price < 500**



- Next page is the table of the average prices giving a better idea.

Location	Room type	price
Bronx	Entire room/apt	127.50
	Private room	66.78
	Shared room	59.80
Brooklyn	Entire room/apt	178.32
	Private room	76.50
	Shared room	50.52
Manhattan	Entire room/apt	249.23
	Private room	167.77
	Shared room	88.97
Queens	Entire room/apt	147.05
	Private room	71.76
	Shared room	69.02
Staten Island	Entire room/apt	173.84
	Private room	62.29
	Shared room	57.44

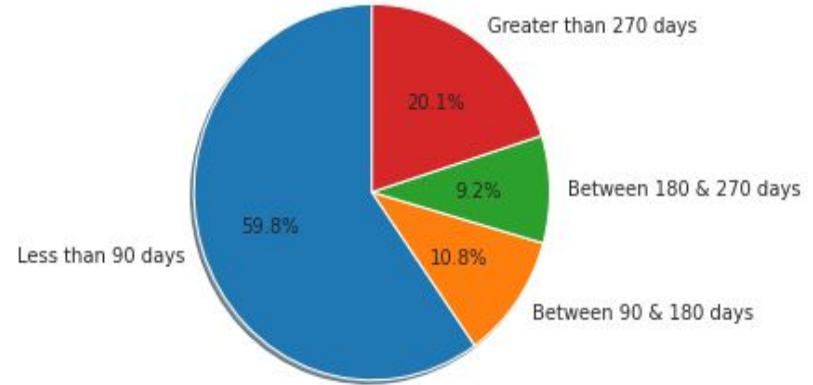
- We can get a better insights of the average prices of different rooms over different locations using a bar-graph.



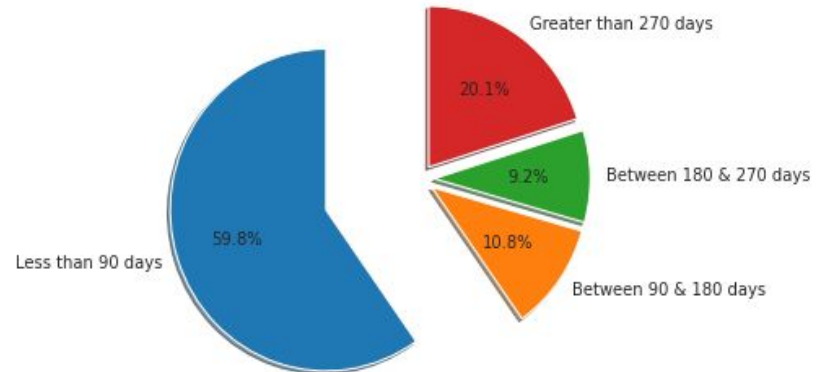
Room Availability throughout the Year

- As going through the data we found a very important column of data i.e., the `availability_365`. It contained the information of how many days a rental is available throughout the year & it cannot go unnoticed. So we calculated the percentage of all the available rentals.
- On the right we have Pie-Charts showing us information about the availability of rentals.
- Major portion of the rentals (around 60%) are only available only for less than 90 days in a year.
- Whereas; only 20% of the rentals function close to a year (more than 270 days).
- And rest of the 20% rentals are operational from a range of 90 - 180 days.

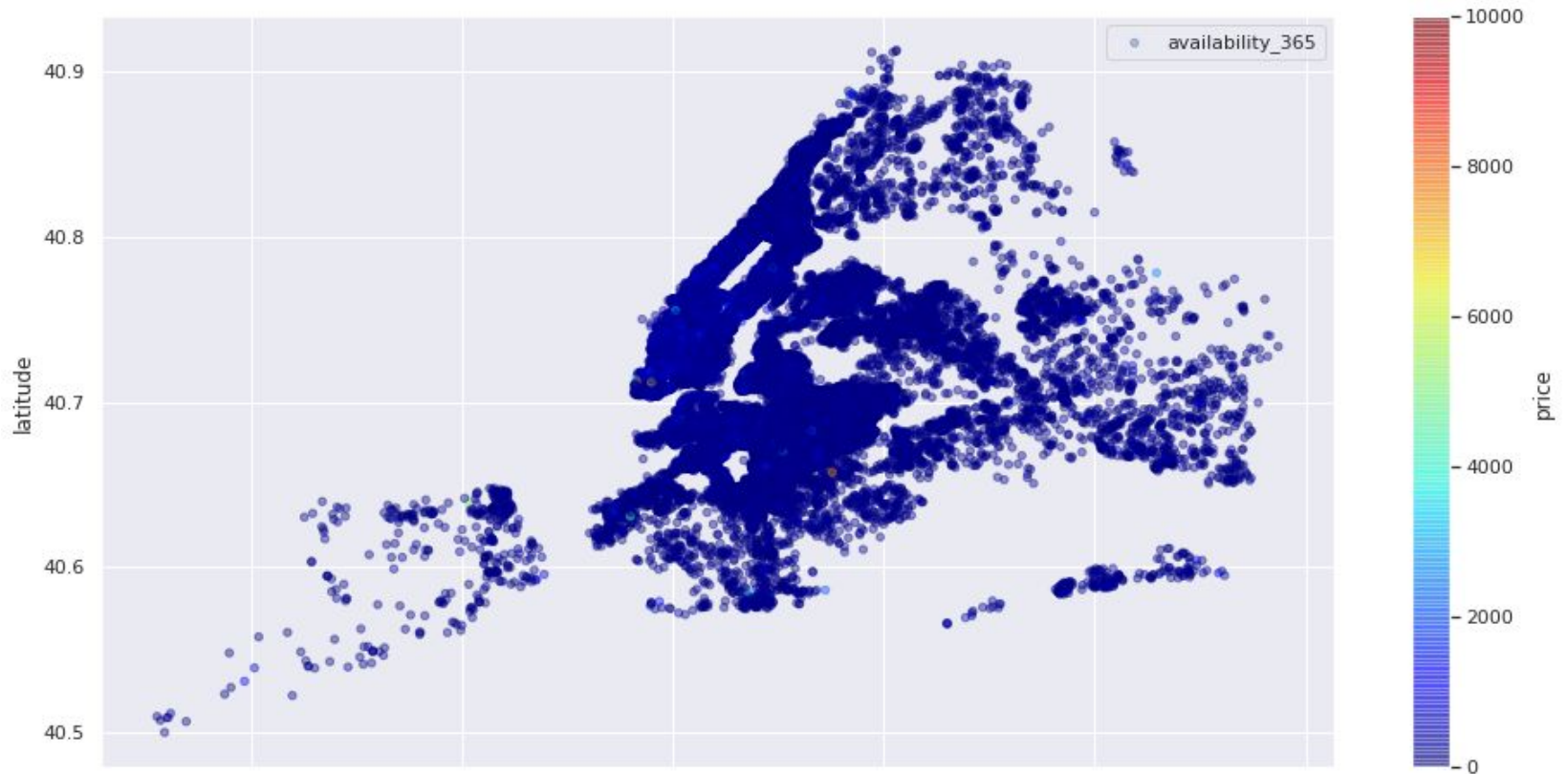
Availability of Airbnb Rentals
As a % of the Calendar Year 2019



Availability of Airbnb Rentals
As a % of the Calendar Year 2019

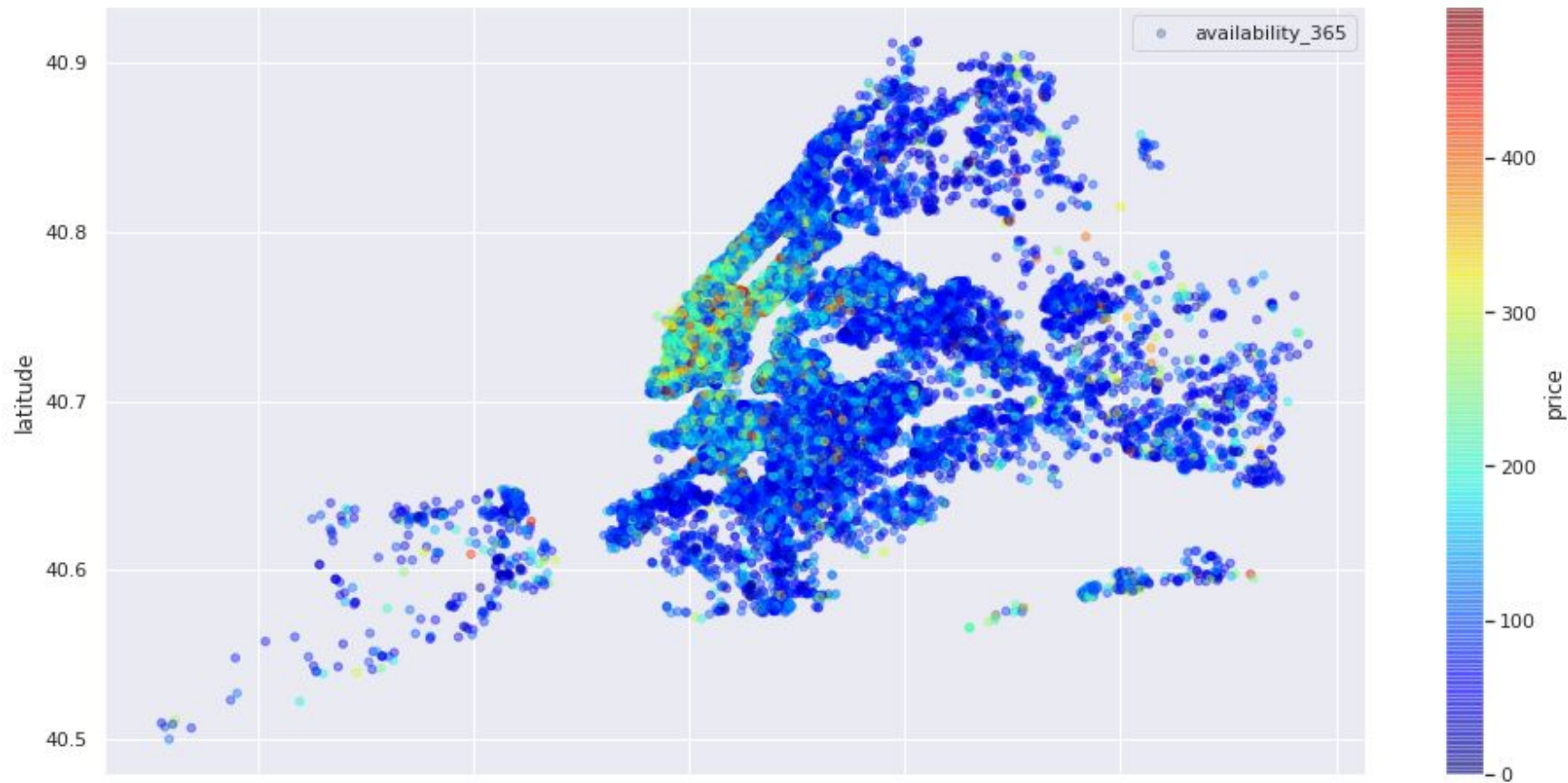


Plotting the entire Dataset based on Longitude & Latitude on basis of Availability and Prices

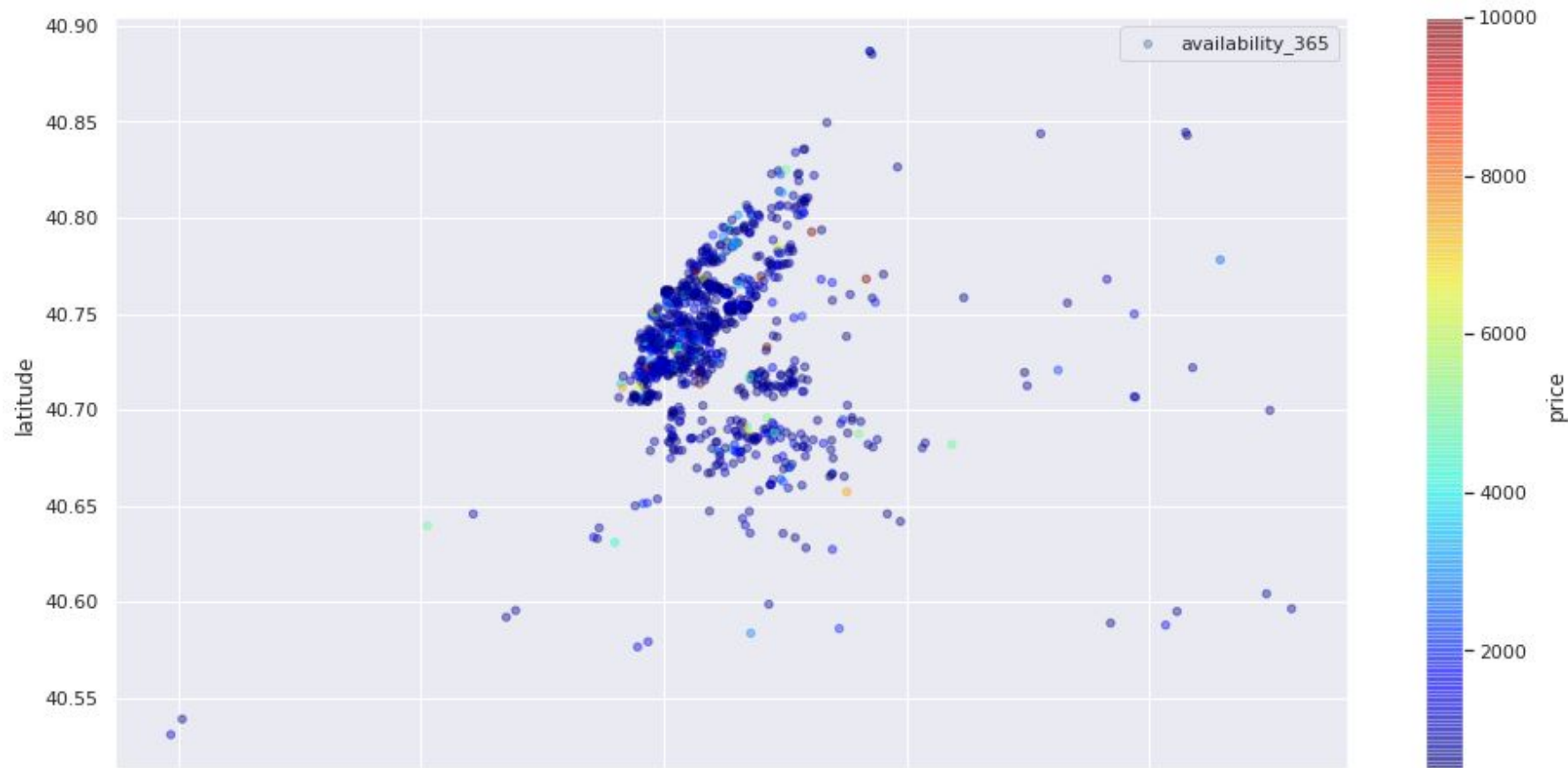


- When we plot all the coordinates i.e., all the listed rentals in the New York City altogether we rather saw a blue dominance in the price range, this tells about the abundance of small price range rentals listing in the New york City.
- For further more clarity we will map them in separate categories i.e., i) with price range < 500
ii) with price range > 500

Rentals below 500 dollars/night in New York City



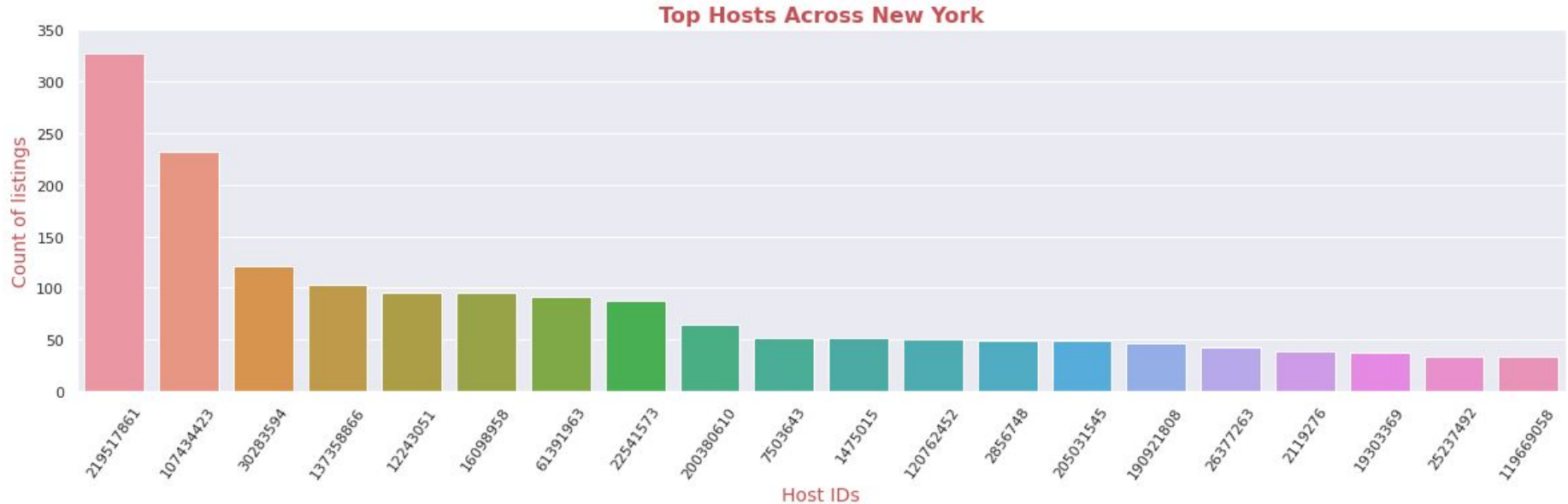
Rentals above 500 dollars/night in New York City



Conclusion

- From the above mapping now we have a more clearer picture about the abundance prices range.
- Most of the listing were near to 100 dollars.
- Only very few listed rentals cross the mark of 500 dollars mark.

Top Hosts across the New York City



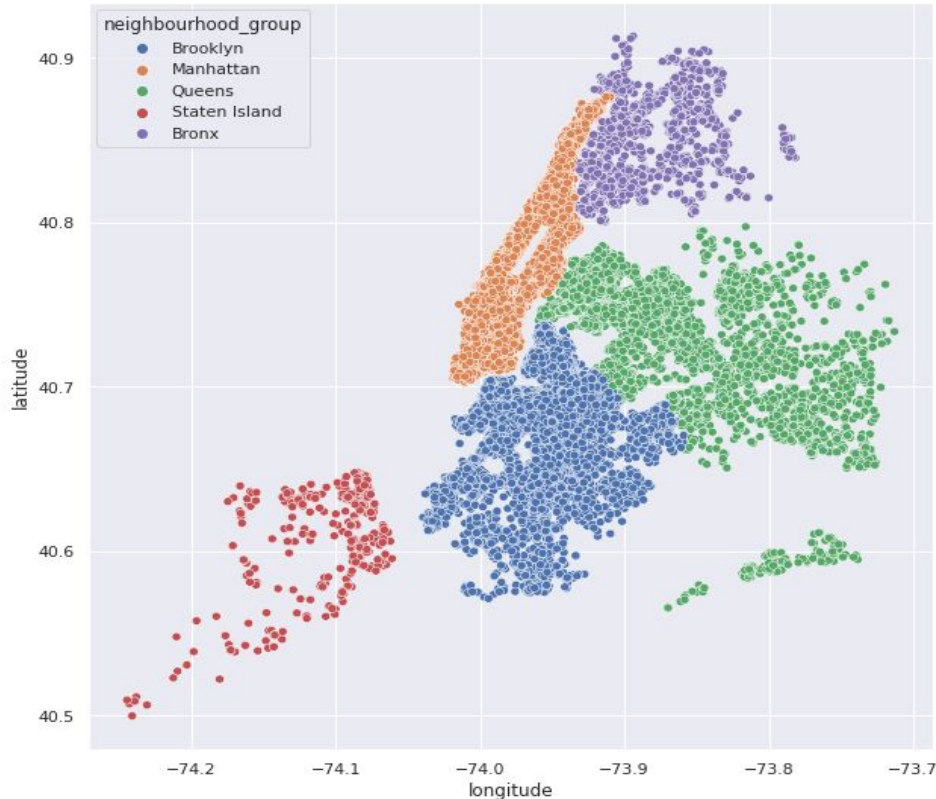
- We came to a finding that all the neighbourhood groups/locations are part of the [New York City](#).
- The above bar-graph represents the top-hosts based on the number of listing they have in the Airbnb(no.of rentals listing).
- The x-axis represents the particular host Id's which are specific to one person.
- The y-axis features the count of the listing, i.e., the number of rentals owned/registered under a host on the Airbnb.

Best and Top of the Class Neighbourhood In the New York City



- The graph above features the top 10 neighbourhoods in the [New York City](#).
- The x-axis features the neighbourhood.
- The y-axis features the number of listing/rentals in the neighbourhood.
- Leading the list is 'Williamsburg'.

Plotting all the listed Rentals in the New York City using Latitude & Longitude



- The mapping on the left represents the five different neighbourhoods of the New York City with different colour codes.
- Namely - 'Brooklyn' in blue, 'Manhattan' in orange, 'Queens' in green, 'Staten Island' in red and 'Bronx' in magenta
- We can have a guesstimate and conclude that 'Brooklyn' & 'Queens' have a greater number of rental listing.
- 'Staten Island' seem to be either a small neighbourhood or less listings on Airbnb, as we can have a clear picture of the total rentals.

Conclusion

- We found out about the number of listed rentals in the five neighbourhoods of the New York City.
- We have then categorised the different room types i.e., (i) Entire Home/Apt (ii) Private Rooms (iii) Shared Rooms.
- We then calculated the average price of the listed rentals/night among different neighbourhoods in the New York City.
- We next faced an obstruction as all the listed rentals were not available throughout the year, So we calculated and displayed the percentage of rentals operation/functioning throughout the year.
- Next we have some scatter plots of the New York City's given neighbourhoods to give us an idea of the ranges and abundance of the various price ranges and we came to a conclusion that most of the rentals cost around 100-130 dollars/night.
- Our next task was to find the top hosts of all the neighbourhoods, so we took the number of listings as the base, as the most successful and popular hosts will only have more listing rentals.
- Next we have gone through the process to find the sub-locations among all the neighbourhood that had most of the listings.
- At the end we finished it off with a scatter plot of all the listings in all the five given neighbourhood with each having a separate colour code of themselves for easier Identification and Analysis.