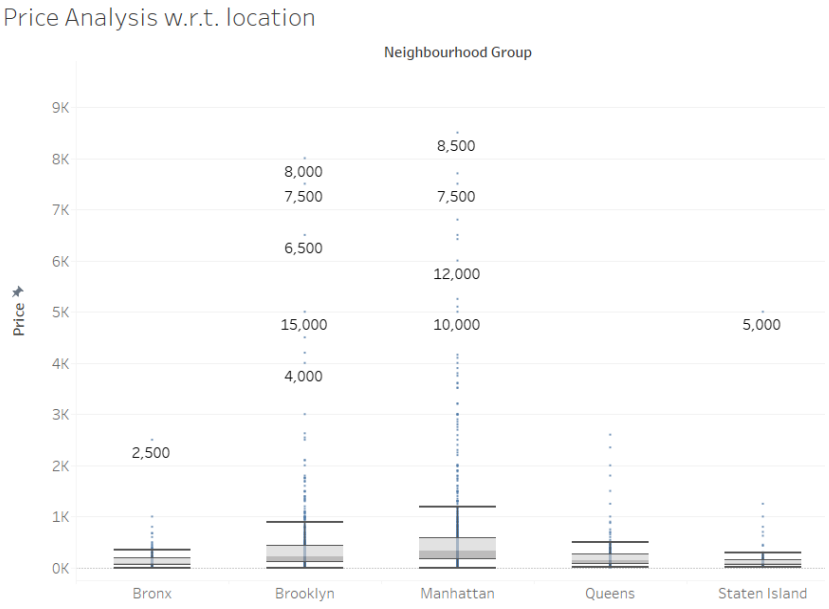
Data Methodology

Step 1: Storyboarding

- Went through the data to get familiarized with it and noted down important fields
- Made a mind map of the various slides of the presentation
- Made a rough template based on this mind map.

Step 2: Data Wrangling

- Performed univariate analysis using Tableau on the fields to see their distributions, the unique values in a field, the missing values and to check for outliers if any
- There was a small proportion of null values in “price” column (0.02%) which was very insignificant, so left them as it is.
- In “No. of Reviews”, “Last Review” and “Review_per_month” column there were almost 20% missing values. So, imputed with 0 in “No. of Reviews” column and left them blank “Last Review” and “Review_per_month”column.
- Price was highly positively skewed so median was very close the lower quartile with some outliers as seen in the boxplot below



- Found out neighbourhood group wise 99th percentile price and created a new column with price capped at 99th percentile.

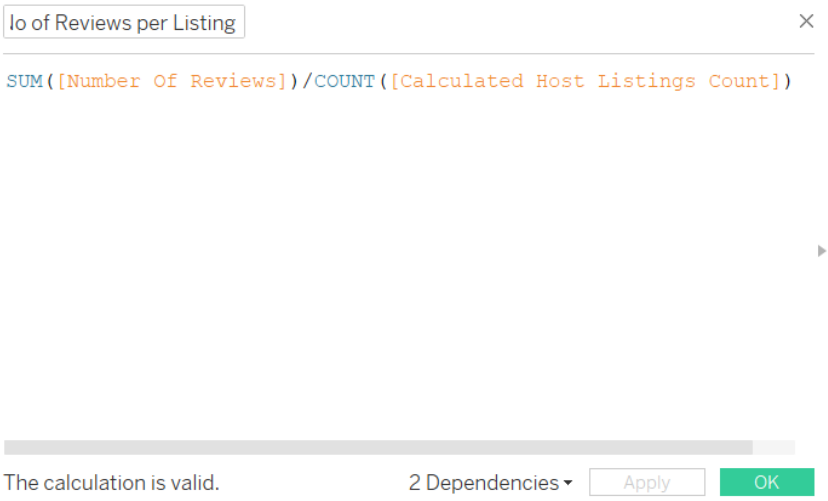
Neighbourhood ..	
Bronx	450.0
Brooklyn	550.0
Manhattan	999.0
Queens	400.0
Staten Island	728.0

- Created Price segment based on below calculation, by creating a calculated field.

Percentile	Price Segment	Price
25	Affordable	Less than equal to $69
50	Reasonable	More than $69, Less than equal to $106
75	Premium	More than $106, Less than $175
Above 75	Luxurious	More than $175

- Created a calculated field for Minimum Nights.

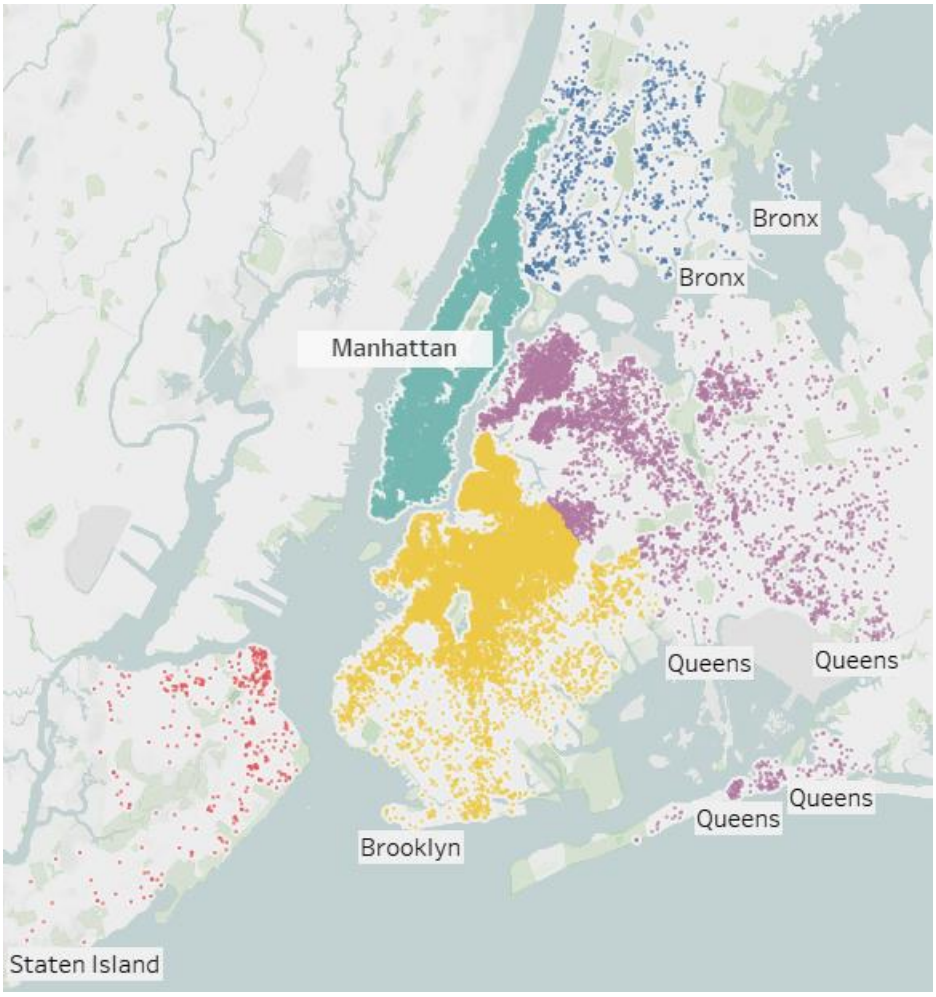
- Created a calculated field of number of reviews per listing



- Created a calculated field as “days since last review” for recency analysis. The days difference is calculated by considering latest review date as reference date.

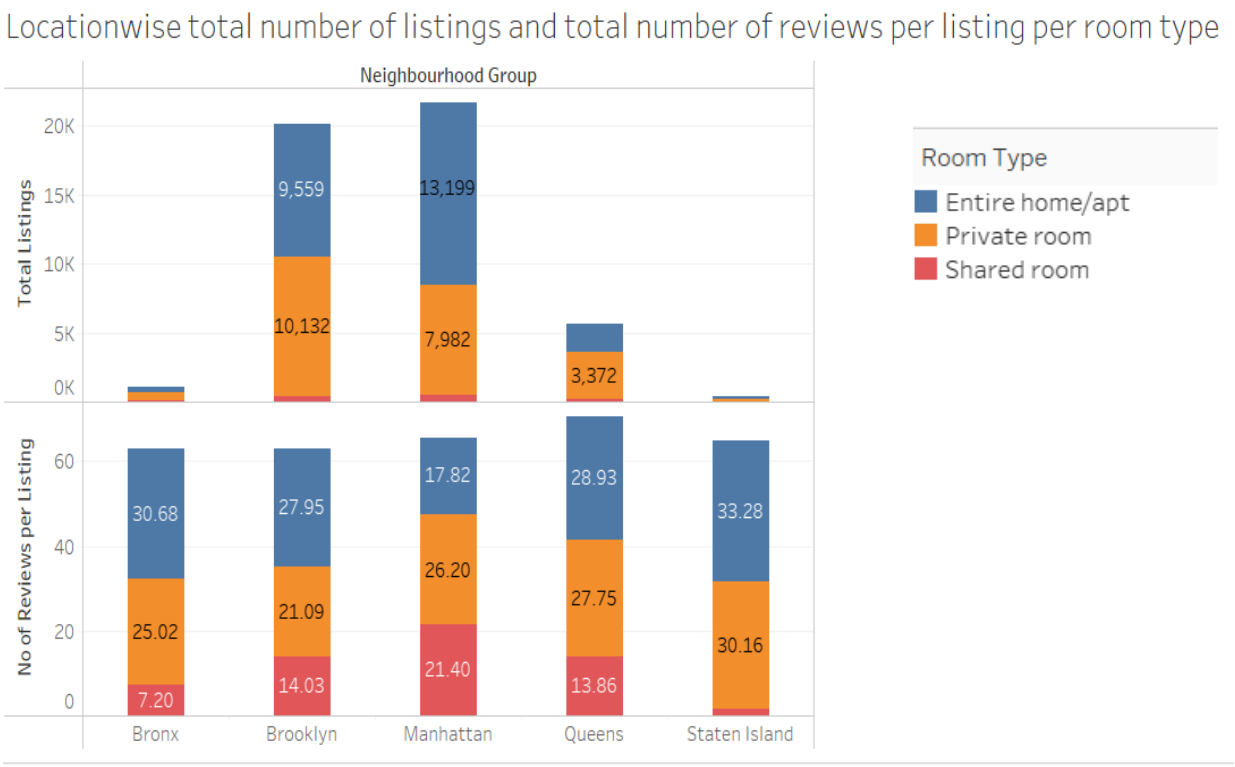
Step 3: Data Analysis

- Checked for approx. market penetration using latitude and longitude.



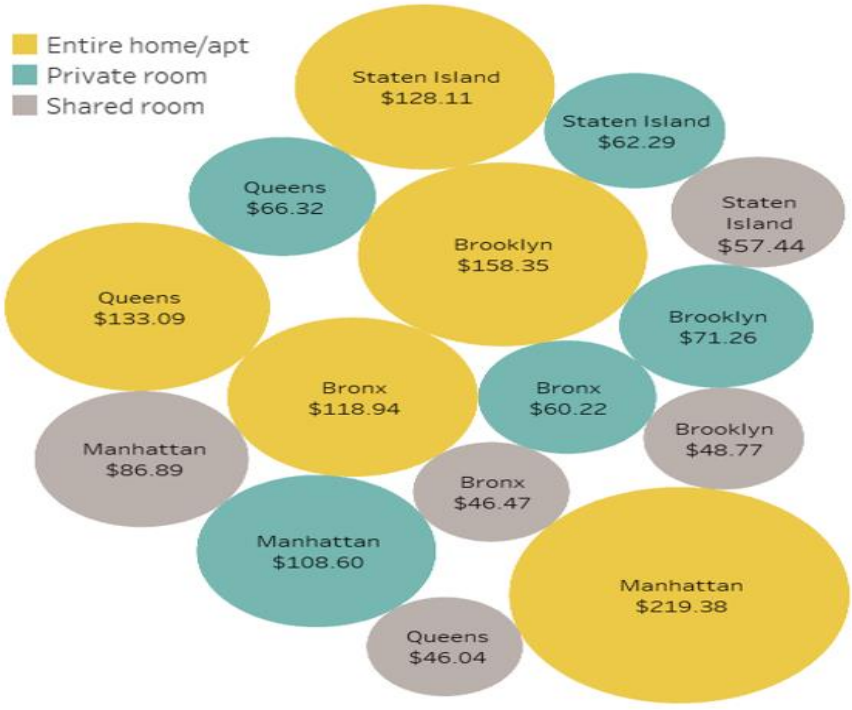
- Manhattan has the highest market penetration, followed by Brooklyn
- Queens and Bronx still have some untapped areas
- Staten Island has the lowest market penetration

- Checked neighbourhood grouped wise distribution of listings, reviews, price w.r.t room type



- In Manhattan, Entire home/apt are preferred over private room by the customers, whereas, in rest of the locations, customers prefer private rooms.
- For Manhattan, entire room's reviews per listing is 35% lower than the overall average and total number of listings is the highest

- Checked for average price of each room type and percentage increase on the basis of location.

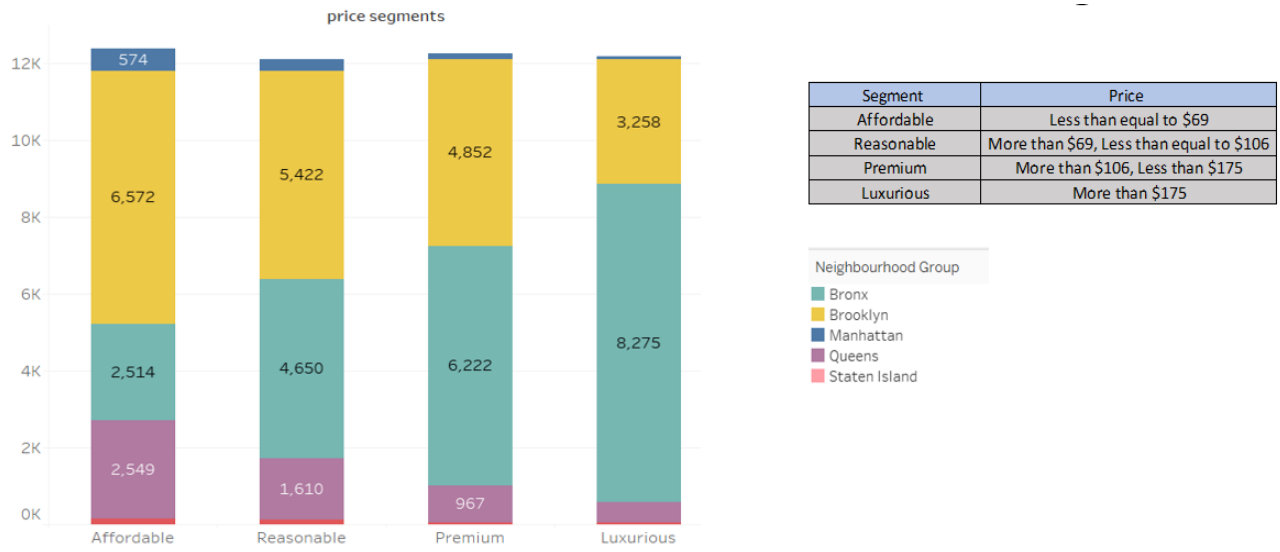


Neighbourhood Group

Room Type	Bronx	Brooklyn	Manhattan	Queens	Staten Island
Shared room	46.5	48.8	86.9	46.0	57.4
Private room	60.2 29.6%	71.3 46.1%	108.6 25.0%	66.3 44.0%	62.3 8.4%
Entire home/apt	118.9 97.5%	158.4 122.2%	219.4 102.0%	133.1 100.7%	128.1 105.7%

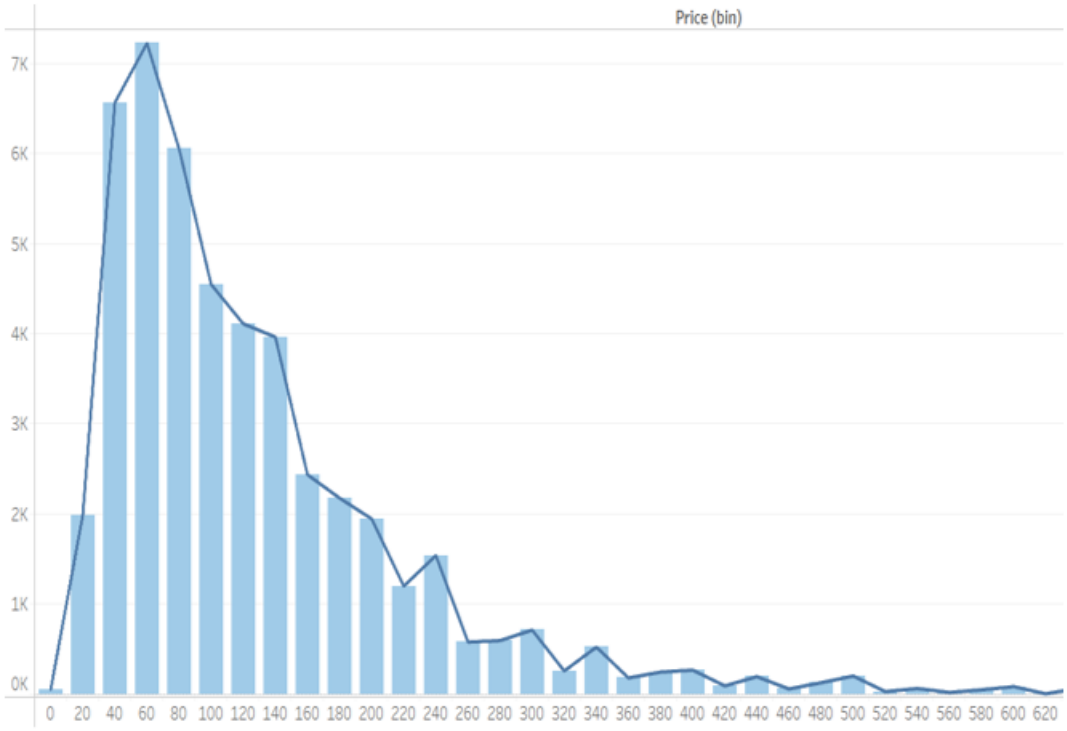
- Considering price variation, Manhattan and Brooklyn comparatively caters to wide range of customers
- At all locations, on an average, Entire home/apt is roughly 2 times costlier than private room
- In Staten Island, private rooms are only 8.4% costlier than shared rooms.

- Checked for location wise variation in count of rooms types based on price segment.



Segment	Price
Affordable	Less than equal to $69
Reasonable	More than $69, Less than equal to $106
Premium	More than $106, Less than $175
Luxurious	More than $175

- Checked for price preferences



Step 4: Presentation

- Made the presentation adhering to best practices and pyramid principle
- Added recommendations for the respective departments.