

Data Methodology:

Step1: Story Boarding

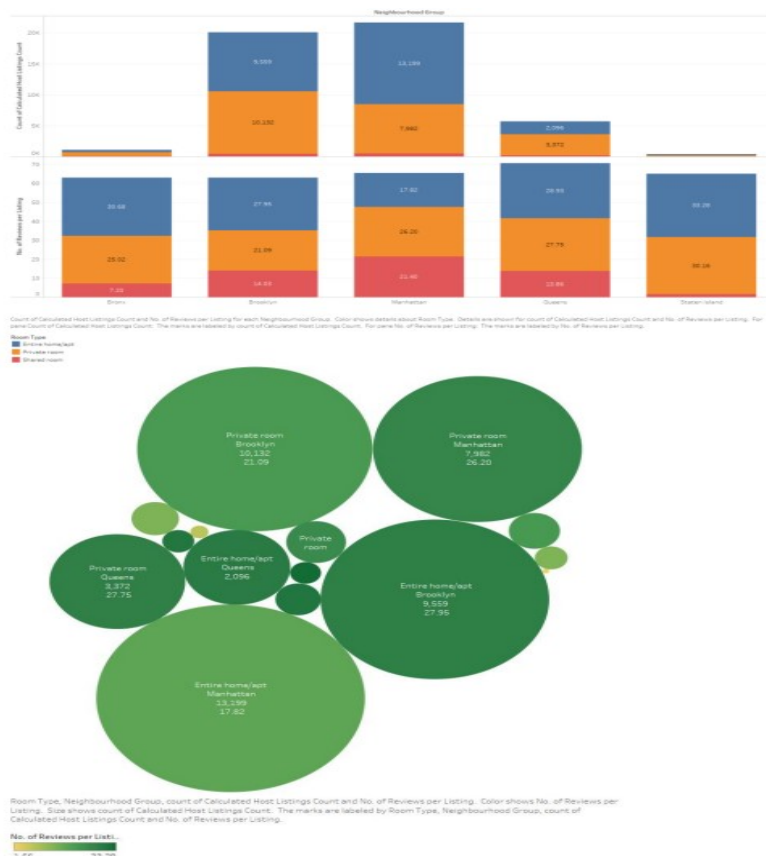
- Gone through dataset and got noted down the important fields
- Made a mind map of the various slides of the presentation
- Made a rough template based on this mind map.

Step 2: Data Wrangling

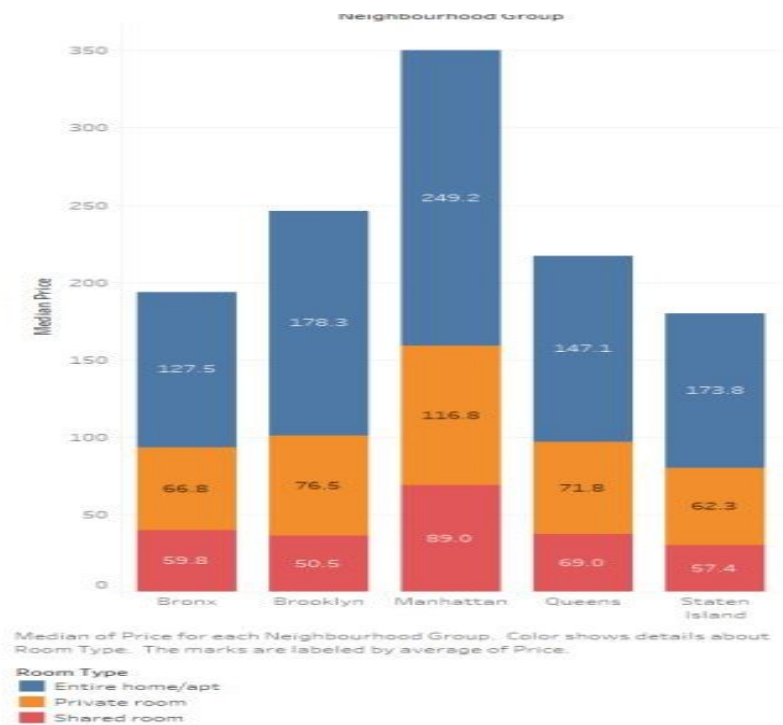
- Did univariate analysis using Tableau on the fields to see their distributions, the unique values in a field, the missing values and to check for outliers if any
- There was a small proportion of null values which would not affect my analysis so let them stay as it is
- Price was highly positively skewed so median was very close the lower quartile with some outliers as seen in the boxplot below
- Since price has outliers, used median instead of mean as the measure for price
- Host Listings count is maximum for entire apartment and private room and is very small for shared room as seen below
- Created a grouped field for Minimum Number of Days assuming null values belonged to the category
- Created a calculated field of number of reviews per listing.

Step 3: Data Analysis

- Checked neighbourhood grouped wise distribution of price and room type.



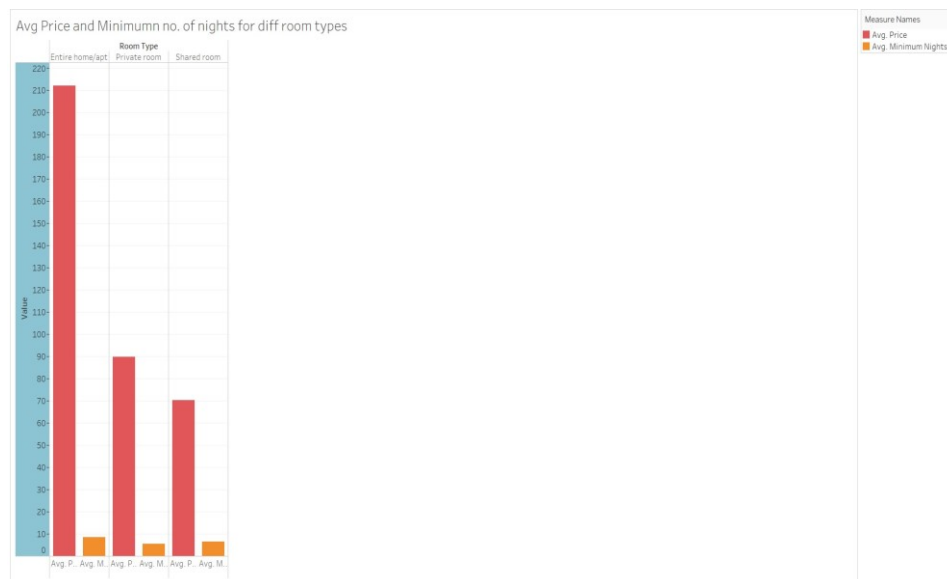
- Entire apartments are popular and so are private rooms
- Except in Manhattan, shared rooms are very unpopular
- Entire room's reviews per listing for Manhattan is 35% lower than the overall average and number of listings is the highest



- Manhattan is costliest overall and the cheapest are:
- Entire apt: Bronx
- Private room: Staten Islands
- Shared room: Brooklyn
- Found difference in prices between shared rooms, private rooms and entire apartments for each neighbourhood as shown
- Entire home/apt's price is ~100% more than private rooms except Brooklyn and Staten Island with ~150% - Private room's price is ~10% more than shared rooms except Brooklyn and Manhattan with ~40%

Step 4: Presentation

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



Conclusion:

For our data exploration purpose, it would have been nice to have couple of additional features like positive and negative reviews. Depending upon that we could have gauged the popularity of each neighbourhood.