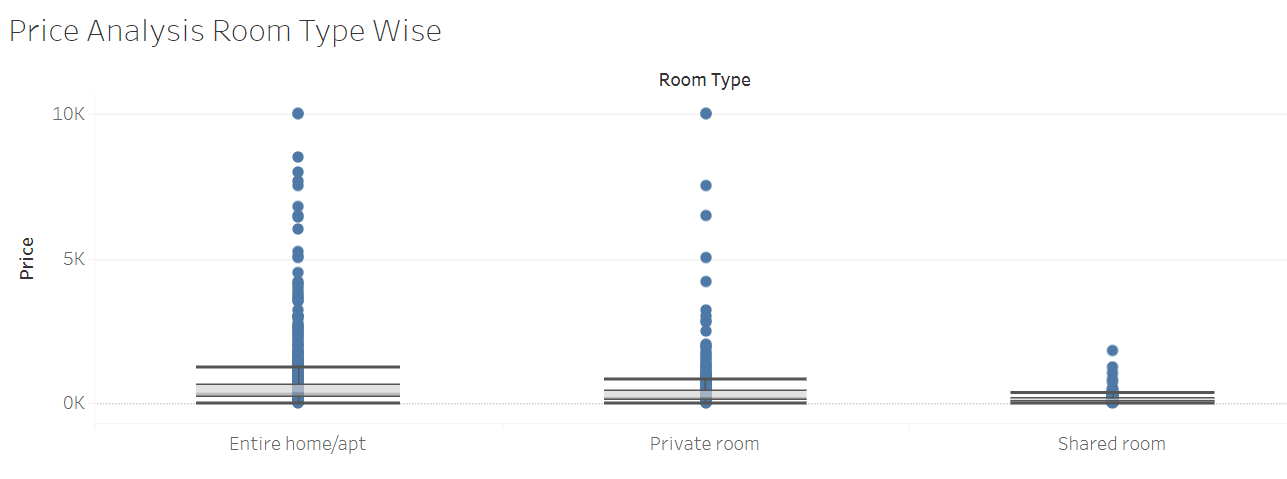
**DATA METHODOLOGY**

**Step 1: Storyboarding**

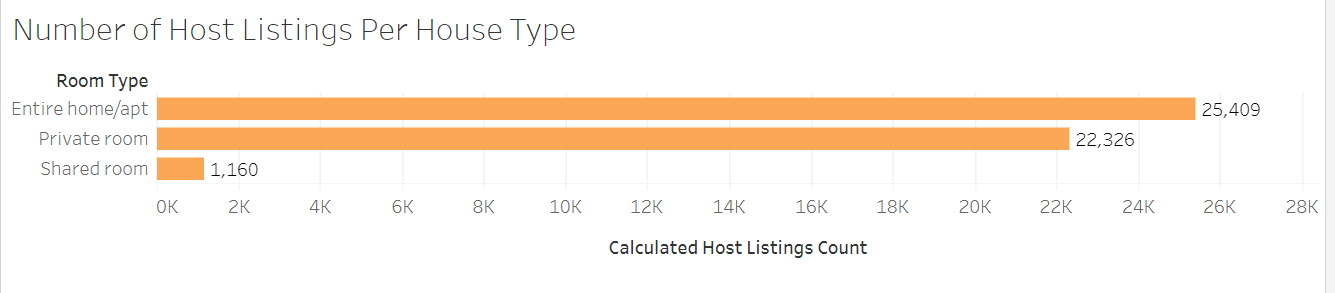
* Went through the data to get familiarized with it and noted down important fields.
* Made a mind map of the various slides of the presentation.
* Made a rough template based on this mind map.

**Step 2: Data Wrangling**

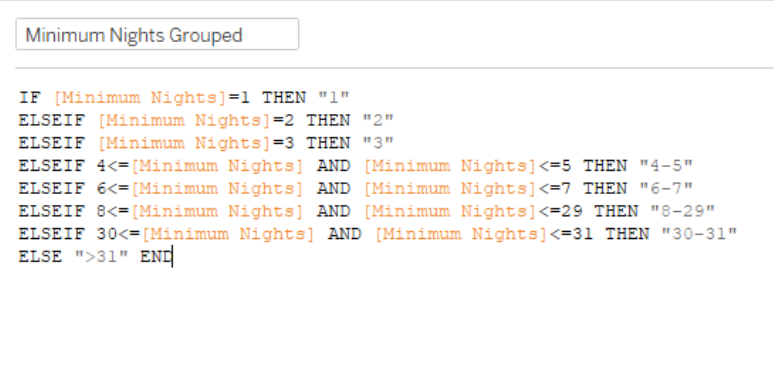
* Did univariate analysis using Tableau on the fields to see their distributions, the unique values in a field, the missing values and to check for outliers if any.
* There was a small proportion of null values which would not affect my analysis so let them stay as it is.
* Price was highly positively skewed so median was very close the lower quartile with some outliers as seen in the boxplot below



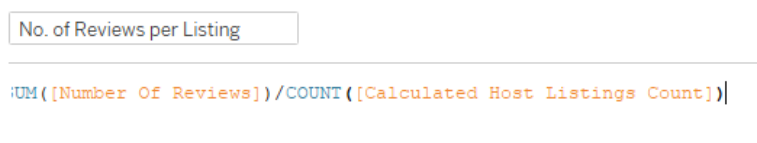
* Since price has outliers, used median instead of mean as the measure for price
* Host Listings count is maximum for entire apartment and private room and is very small for shared room as seen below



* Created a grouped field for Minimum Number of Days assuming null values belonged to the category

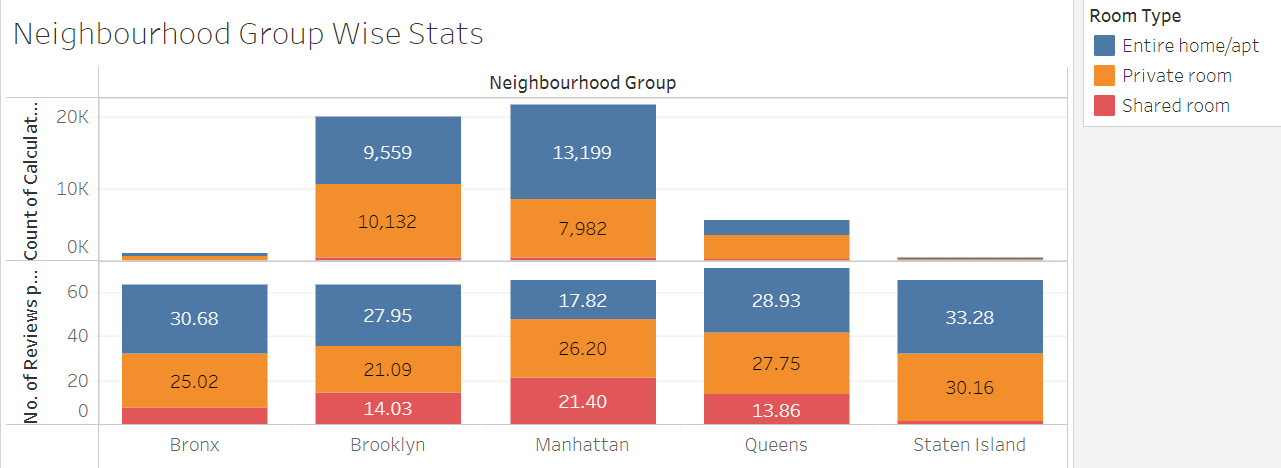


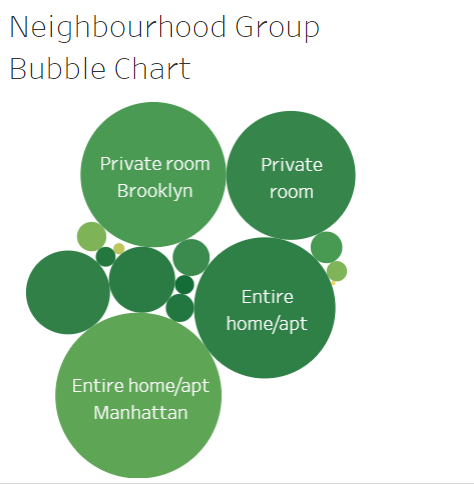
* Created a calculated field of number of reviews per listing

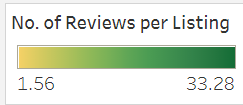


**Step 3: Data Analysis**

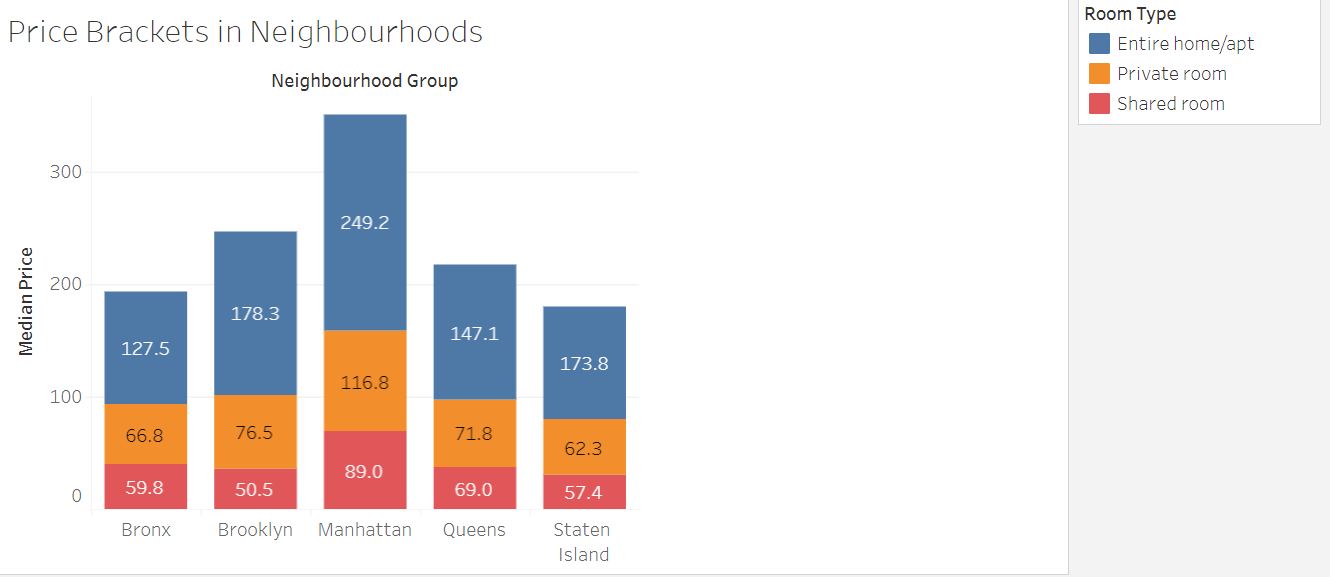
* Checked neighbourhood grouped wise distribution of price and room type



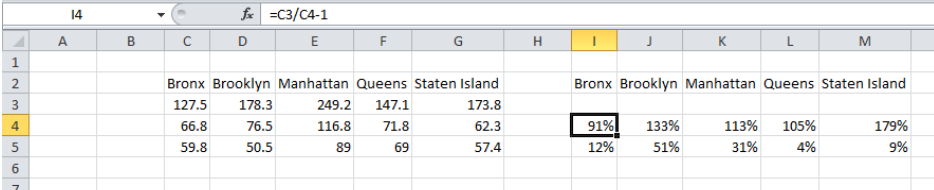




* Through these we clearly understand:
  + Entire apartments are popular and so are private rooms
  + Except in Manhattan, shared rooms are very unpopular
  + Entire room’s reviews per listing for Manhattan is 35% lower than the overall average and number of listings is the highest



* 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 extire apartments for each neighbourhood as shown below



* Through this clearly, we can say that:
  + 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

**Inference after analyzing the data -**

* We saw that people like to visit the center of New York from where they can see the beauty of the city.
* Number of listings of shared rooms are limited but their average price is placed less, and availability is high.
* Number of reviews and reviews per month are more at less price than the higher price as there is less chance of people going for a high price room.
* Manhattan and Brooklyn are very costly neighborhood groups.
* People show interest in the host Blueground and spend more nights here.
* Minimum number of nights to stay reduces with increase in price.
* Focus on prime locations like Manhattan and Brooklyn where people show interest.