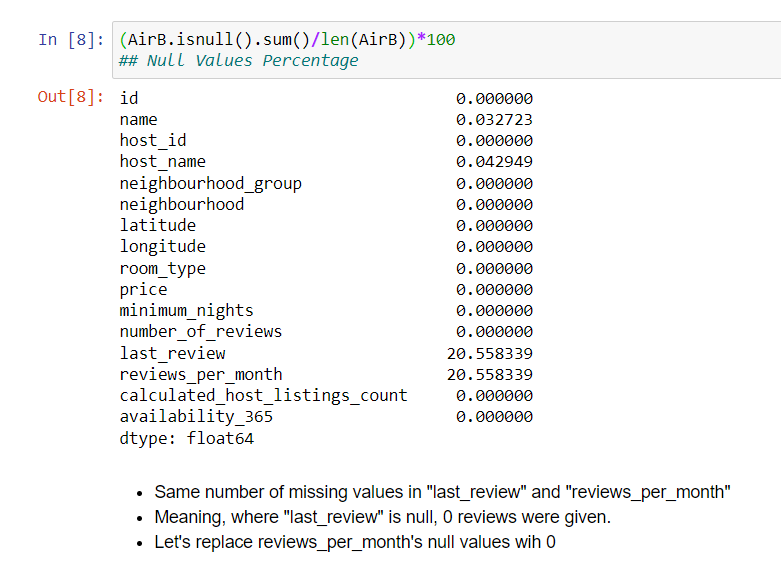
**Data Methodology 1:**

**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

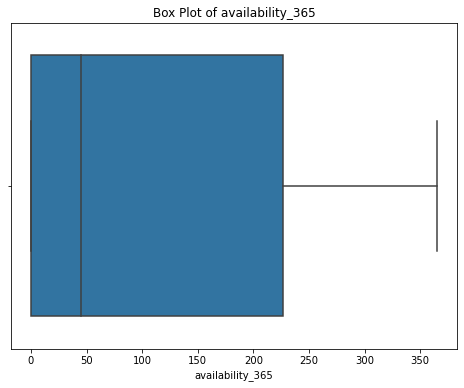
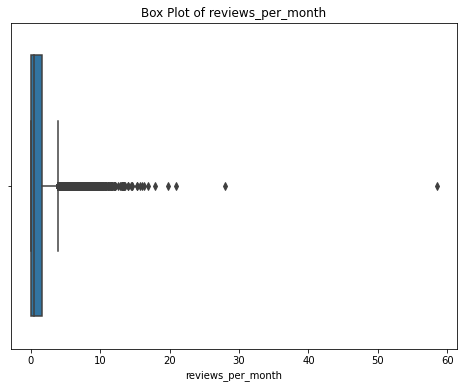
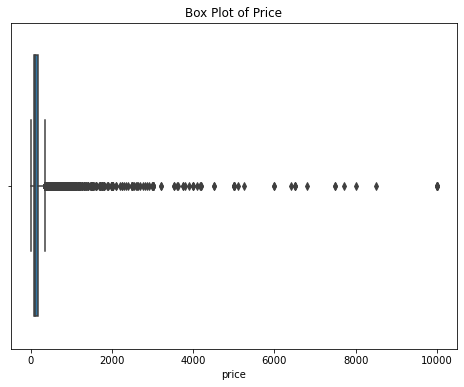
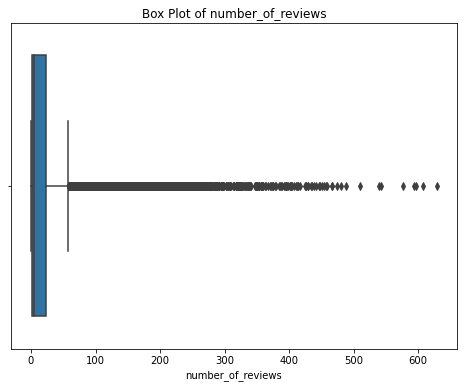
**Step 2: Data Wrangling.**

* Loaded the provided dataset into pandas and tried to understand the variables present.
* Analysed each attribute and checked the data type of each column.’
* Then calculated the null values in each column:



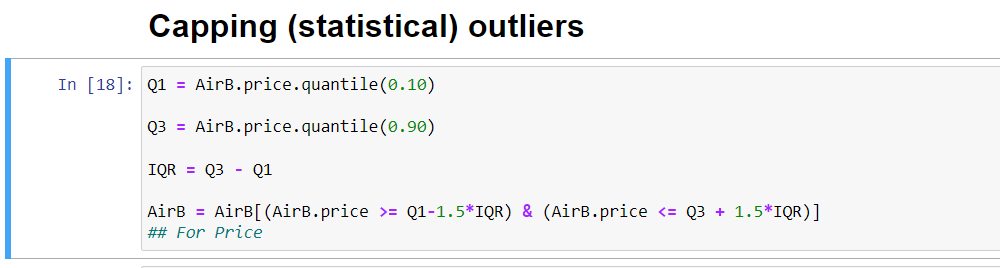
Thus, replaced the ‘reviews\_per\_month’ column with ‘0’.

* Checked the spread of the numerical variables using Box Plot:



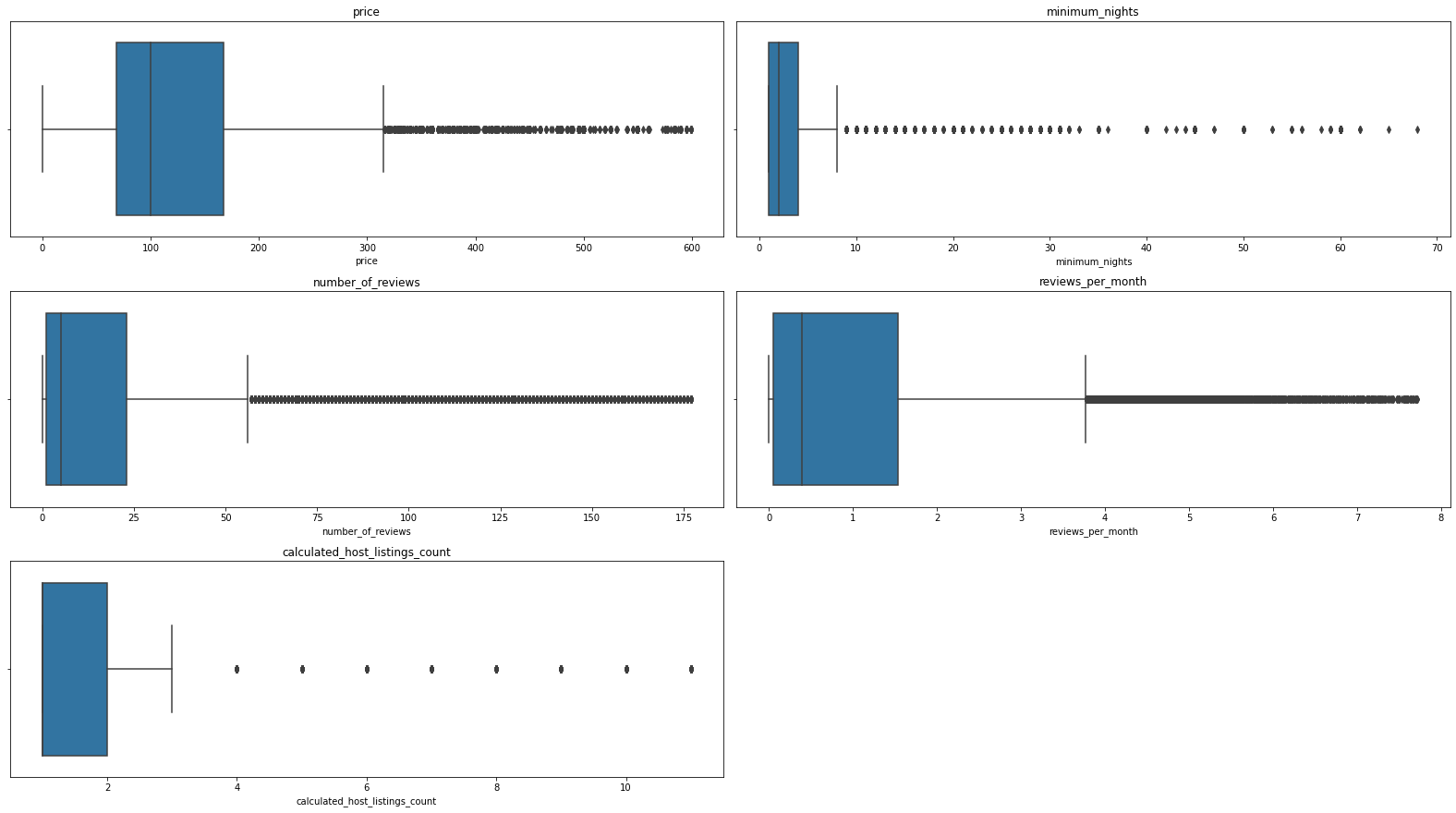
We observe in the above given box plots, that there are a lot of outliers. These can massively skew the data.

* Thus, we got rid of a few outliers using the IQR approach:



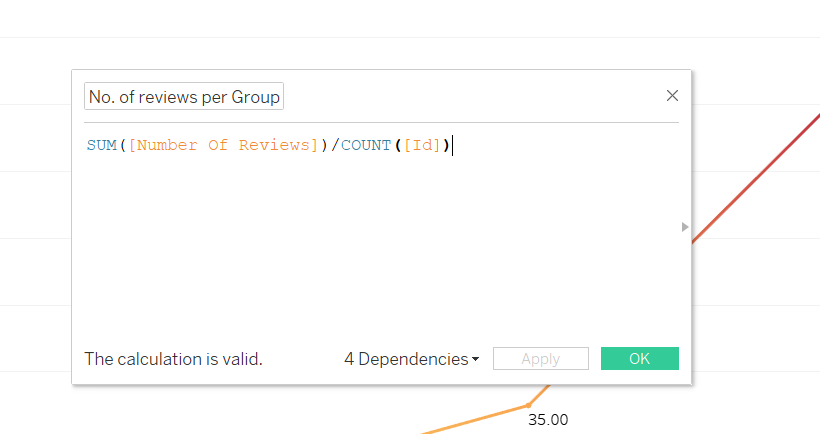
This method was done for all numerical columns (except “availability\_365” since it did not have significant outliers).

* Now, boxplot were plotted again to see the difference:

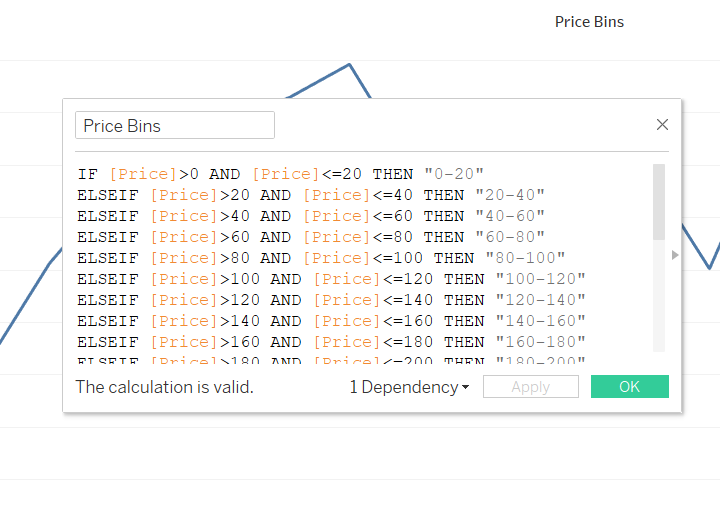


We see a significant difference in the Box Plots now.

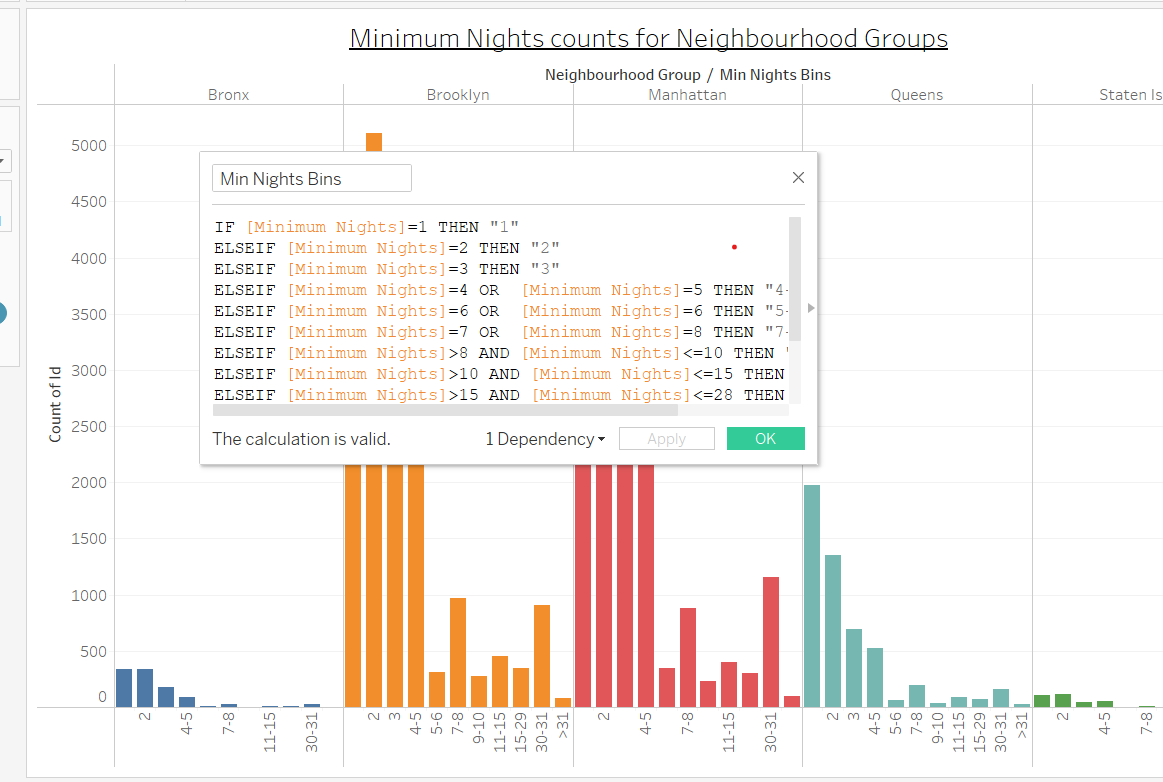
* Loaded the data in Tableau for visualization.
* Created a calculated field for **Average number of reviews**:



* Created bins for price range:

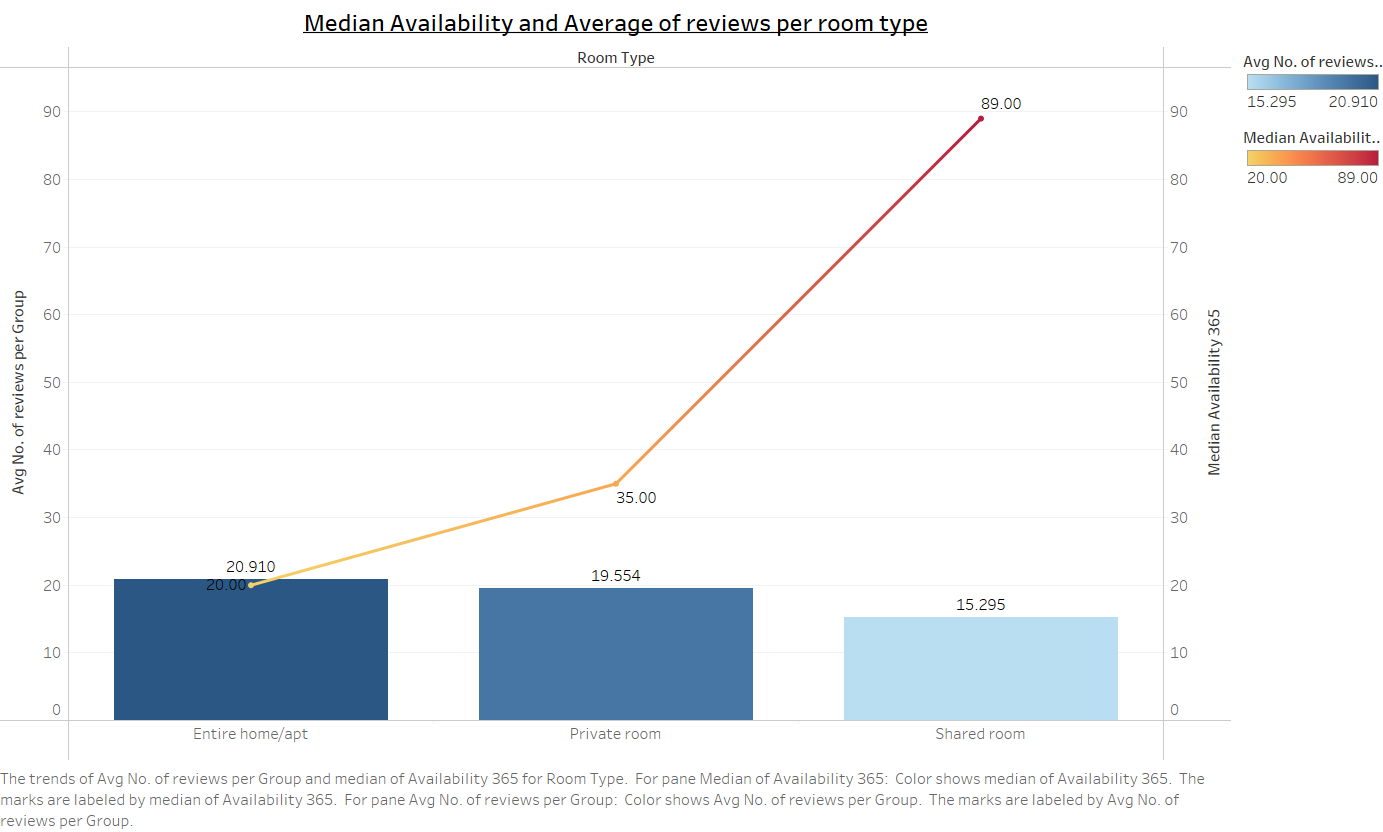


* Created Bins for Minimum number of nights offered:

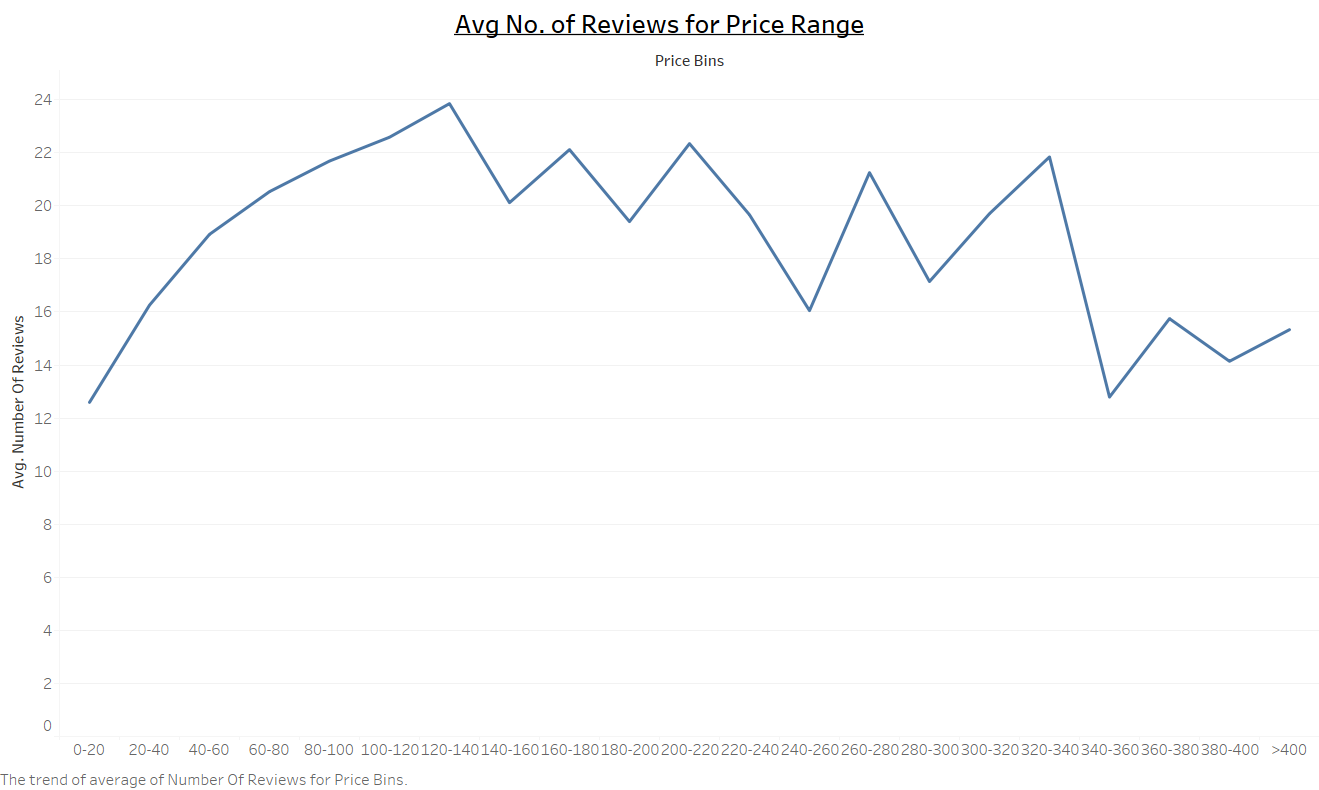


**Step 3: Data Analysis.**

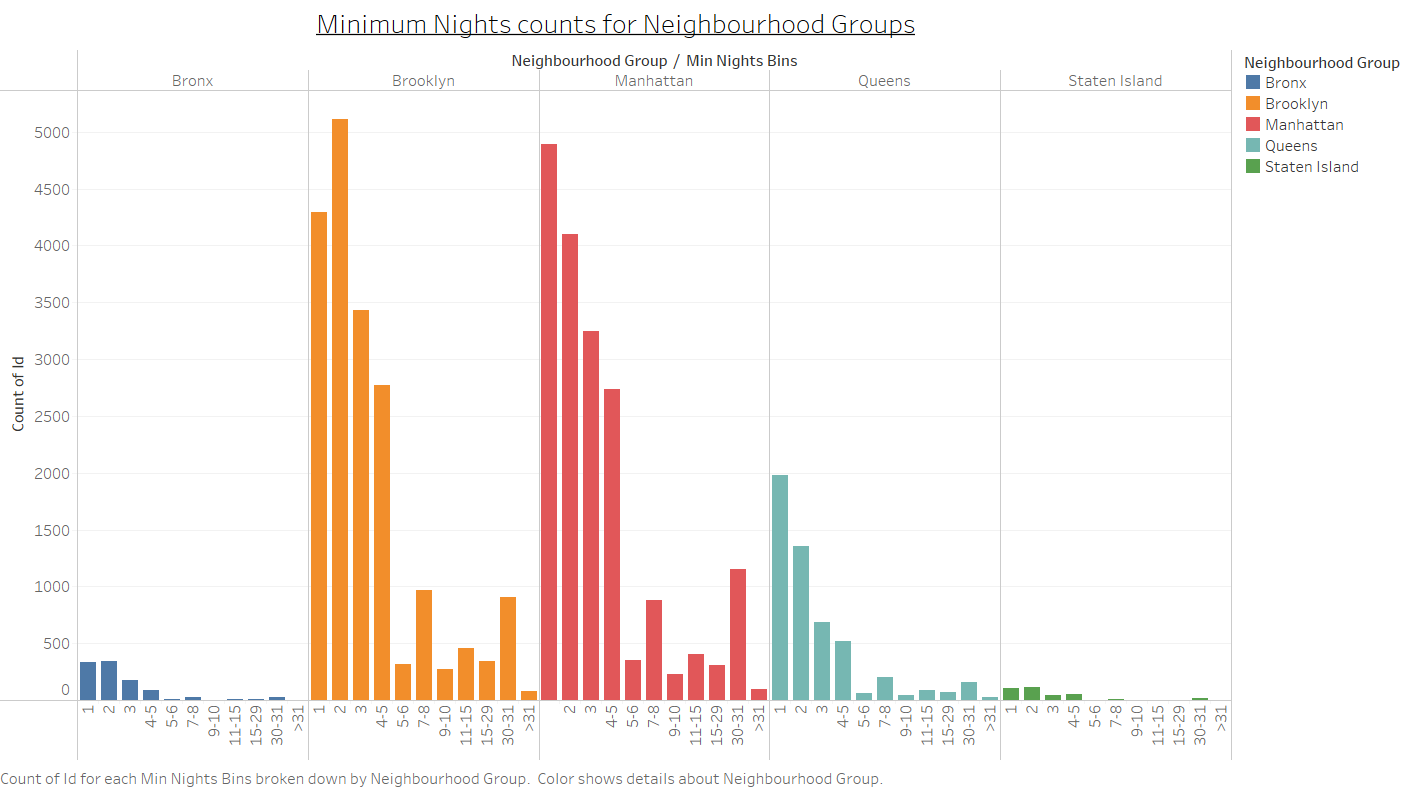
* Compared the average number of reviews (popularity measure) with the median of number of available days in a year for different room types:



* Checked the trend of average of reviews w.r.t increase in the price range:



* Compared the trend for number of room bookings w.r.t number of minimum nights stay offered in each neighbourhood:



**Step 4: Presentation.**

* Made the presentation using the above given insights and visualization using the pyramid principle and keeping the best business practices in mind.