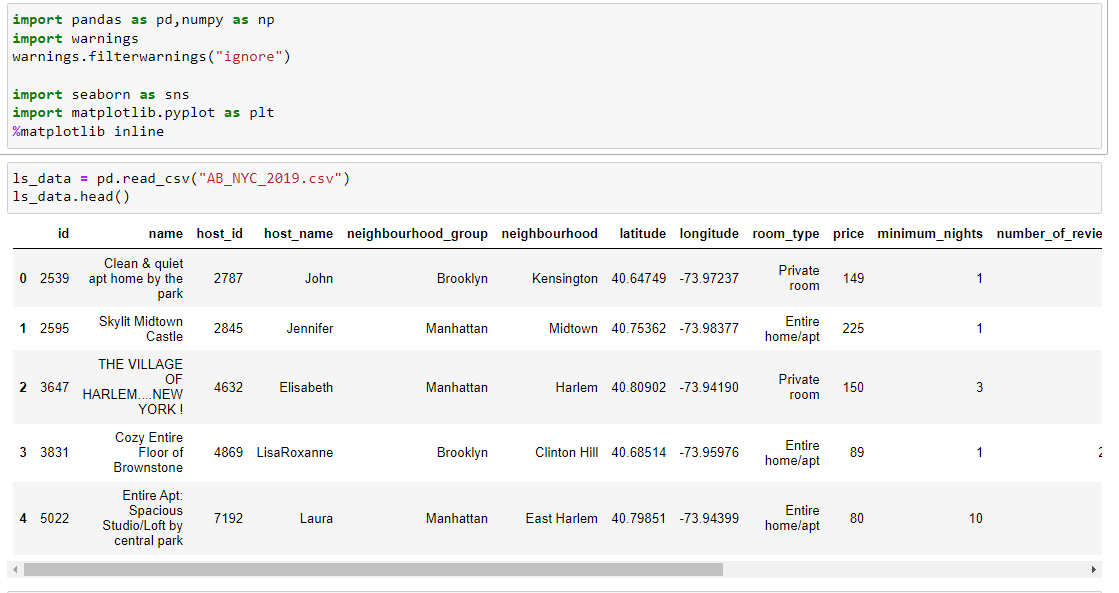
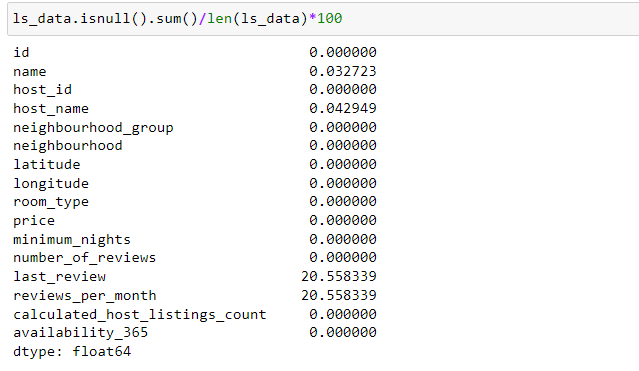
**Methodology document for ppt 1**

In this case study I used Jupyter notebook to perform data cleaning and for data analysis. Tableau was used for data visualization and check various graphs.

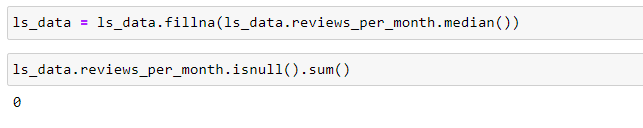
1. Imported necessary libraries and imported the data set in jupyter notebook.



2. Checked for any null values in the data set.



3. “reviews\_per\_month” column: replaced the rows containing the null values with median since there are outliers present in it and median will not affect the distribution.



4. Similarly filled the rest of the null values with their respective data.

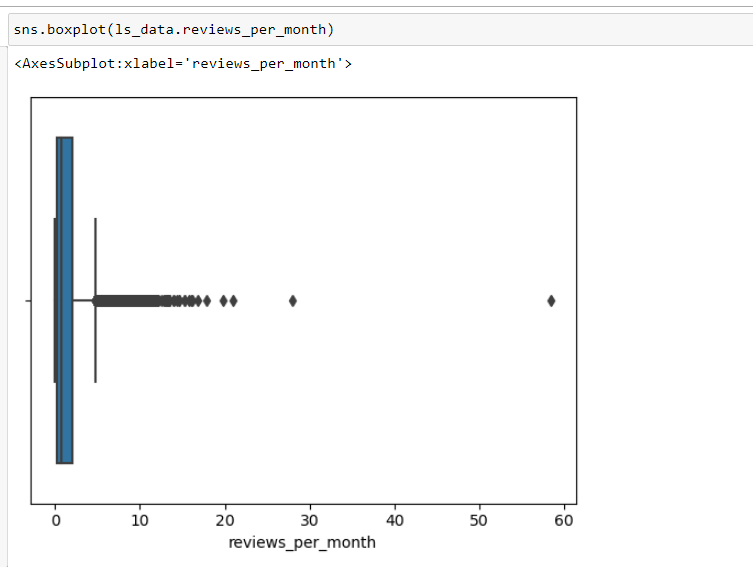
**Data Wrangling:**

1. In data wrangling process we usually collect raw data from different sources and collect it in a common location and then convert that raw data into readable data which can be analyzed.

2. Since the data source is already provided in Excel format so we need not extract any raw data or merge any different format.

3. Data cleaning was done by imputing null values, deleting columns with extremely high null values.

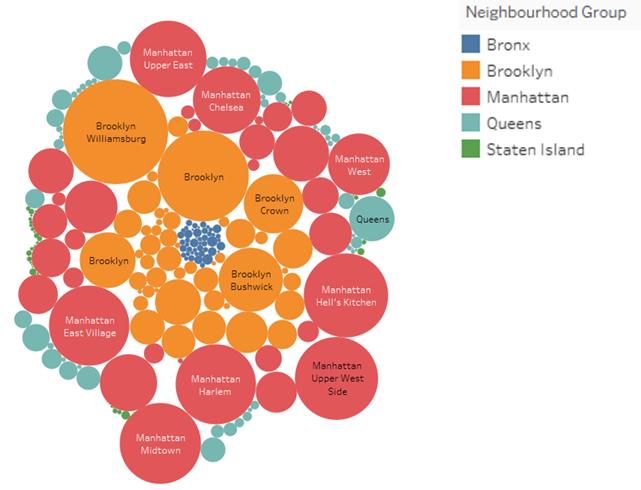
4. Checked for Outliers in the data and there were outliers in the price and reviews columns.



**Data visualization using tableau:**

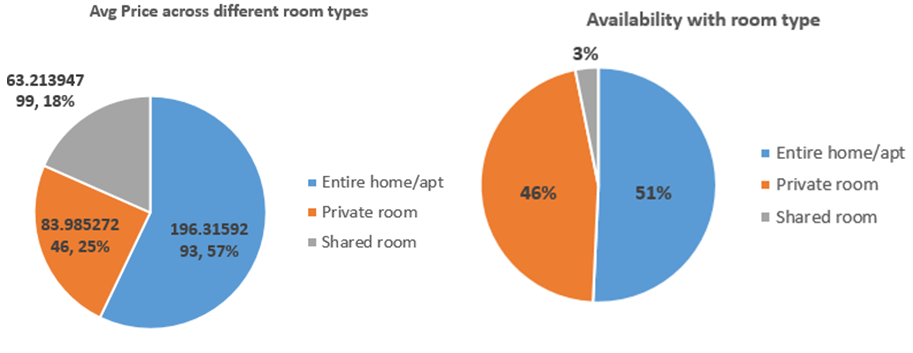
1. *Prices across areas:*

First we calculated prices smaller area-wise and plotted a bubble chart where prices resembled the size of the bubble and color represented the top 5 major regions.



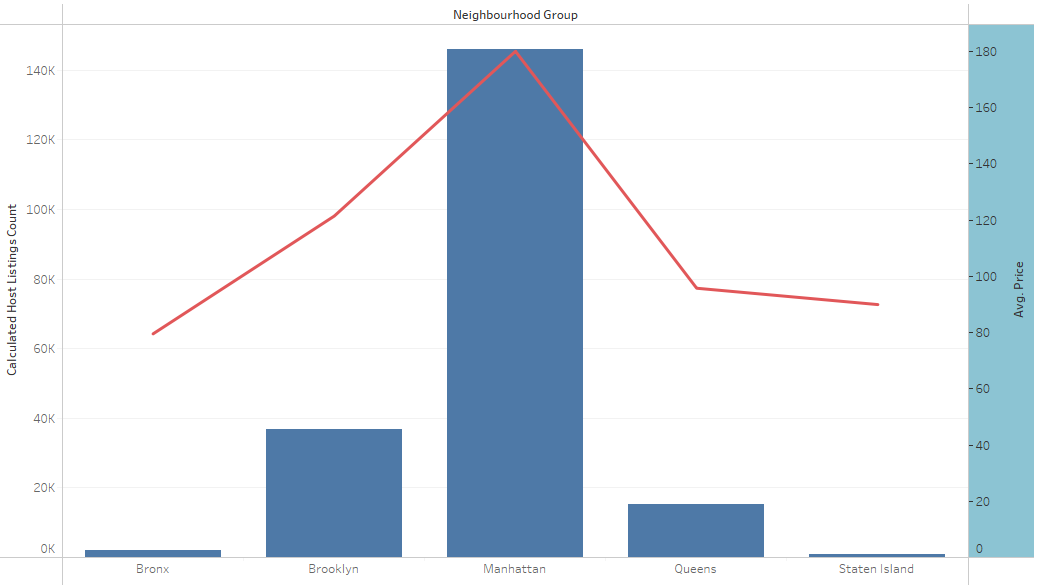
2. *Room type pie chart:*

We calculated average price and availability of different room types with the help of pie chart, we tried showing the share in percentage of each room type across all the listings.



3. *Prices across different neighbourhood group:*

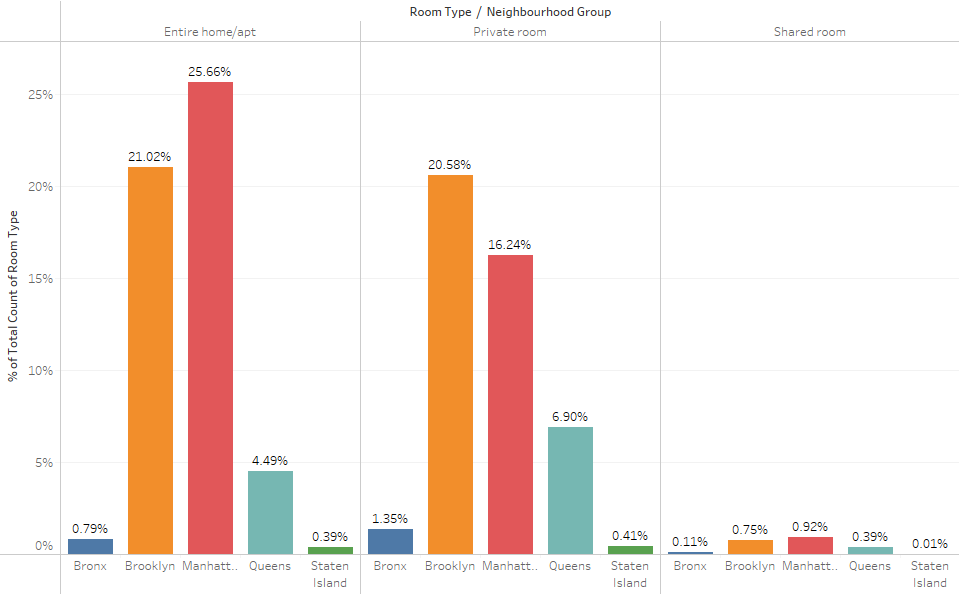
We tried to check different prices across different locations where it clearly shows manhattan at the top. We used dual axis charts to get the results.



**Methodology document for ppt 2**

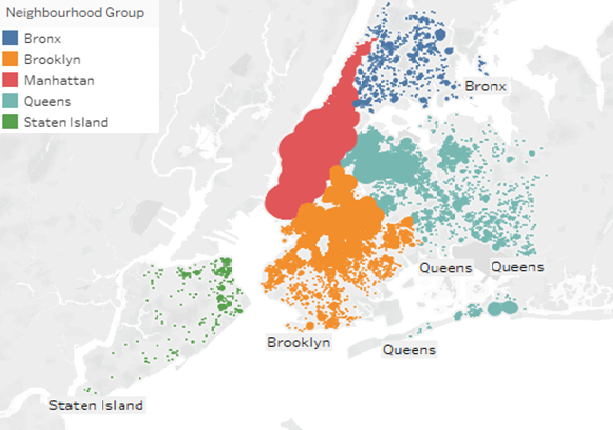
1. *Room type in %wise:*

In this visualization we plotted a bar chart across different room types and across different neighbourhoods. We checked percent wise share of different room types which help to calculate the share of each.



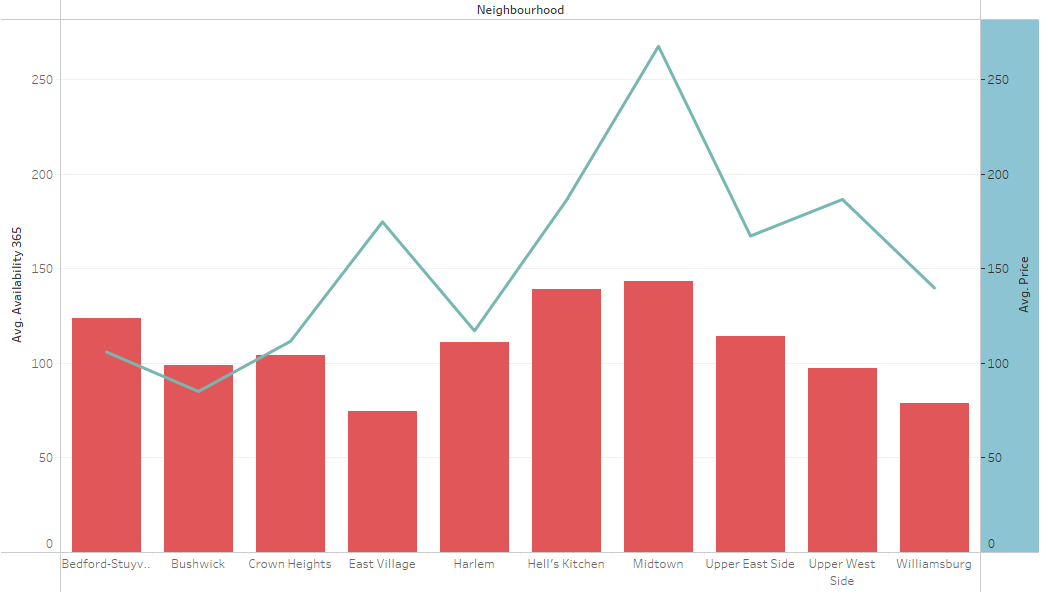
2. *Listings on neighbourhood:*

Listings count according to different neighbourhood areas was plotted with the help of latitudes and longitudes in tableau tool. It provided plottings on actual map.



3. *Neighbourhood vs availability:*

We plotted a Dual axis chart here to compare the availability of room’s vs price per night column. We synchronized the axis so it doesn’t mislead us with the readings.



4. *Listings count vs Number of reviews different neighbourhoods:*

In this bar chart we calculated listings count across different neighbourhoods in percentage and also number of reviews.

