Airbnb, NYC – Data Story Telling Case Study

In this Case study we considered Data set with 48,897 rows and 16 columns, for better and faster Visualizations to reduce memory occupied we have removed columns which wouldn't be necessary, thereafter considered few outliers and tried to remove those for a clear picture of insights then we finally arrived at a Data Set with 45,920 rows and 13 columns

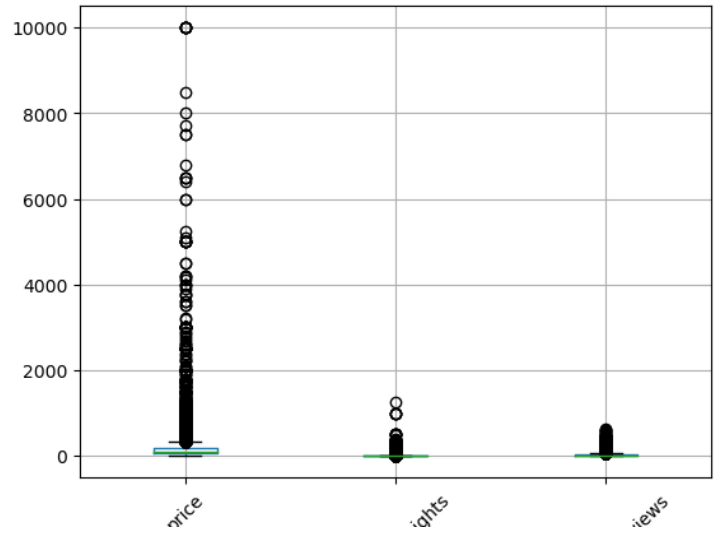
To obtain above said we have followed few procedures of Data Wrangling in Jupiter Notebook:

```
air =pd.read_csv("AB_NYC_2019.csv")   #Importing Data Set
air.dtypes                            # Checking the data types of the columns
air.isna().sum()                      # Check for Null values in  columns
air.drop(['id','last_review','name'], axis=1,inplace=True)      # Dropping columns which are not
                                                                  relevant for Analysis
```

We have checked for Outliers for Numerical columns and found that Price Column has Outliers, treating Outliers can give a better view of Visualizations

```
In [7]: air1=air[['price','minimum_nights','number_of_reviews']]
        air1.boxplot(rot=45)
```

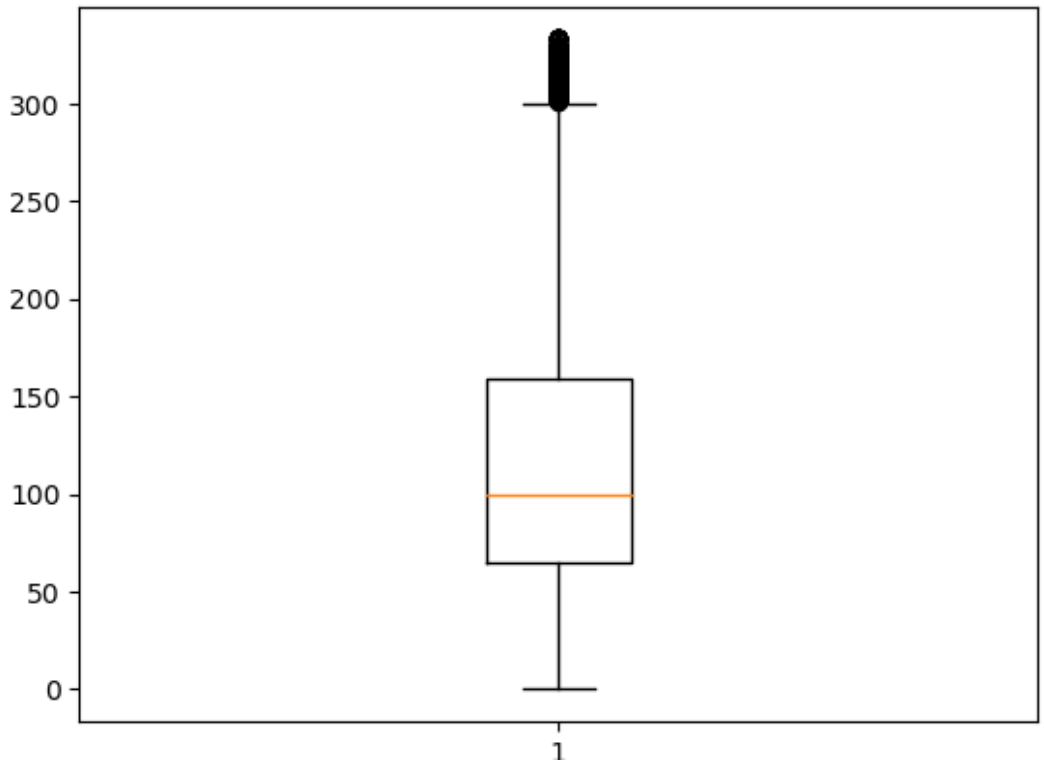
```
Out[7]: <Axes: >
```



Finding Interquartile Range (IQR) :

```
Q1=air.price.quantile(0.25)
Q3= air.price.quantile(0.75)
IQR=Q2-Q1
UE=Q2+1.5*IQR        #UE= Upper Extreme
LE=Q1-1.5*IQR        # LE = Lower Extreme
air2=air[(air.price>LE) & (air.price<UE)]
```

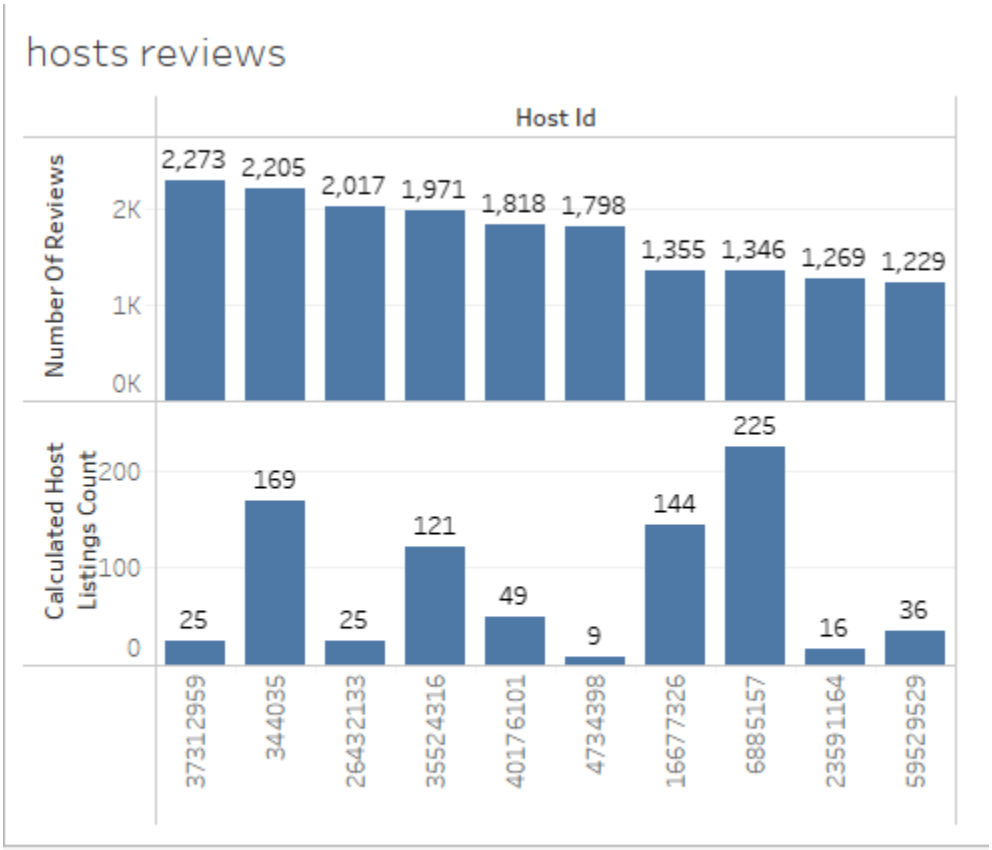
```
: plt.boxplot(air2['price'])
  plt.show()
```




```
air2.drop(['host_name'],axis=1,inplace=True)        # Dropping another irrelevant Column
air2.reviews_per_month.fillna(0)                    # Replacing Null Values with 0 for ease of Analysis
air2.to_csv('airbnbjup.csv')                        # Converting Data Frame to Excel File, to import into
                                                      Tableau for Visualizations
```

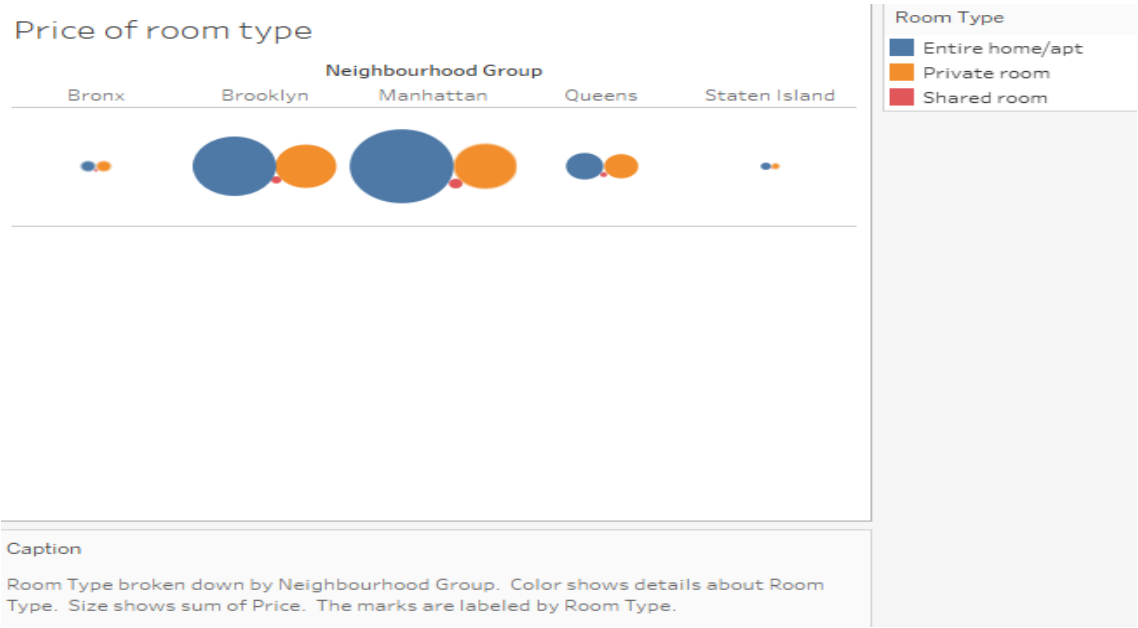
Data Analysis and Visualizations Using Tableau:

Top 10 Hosts (host id's) who have higher customer reviews and their respective count of listings



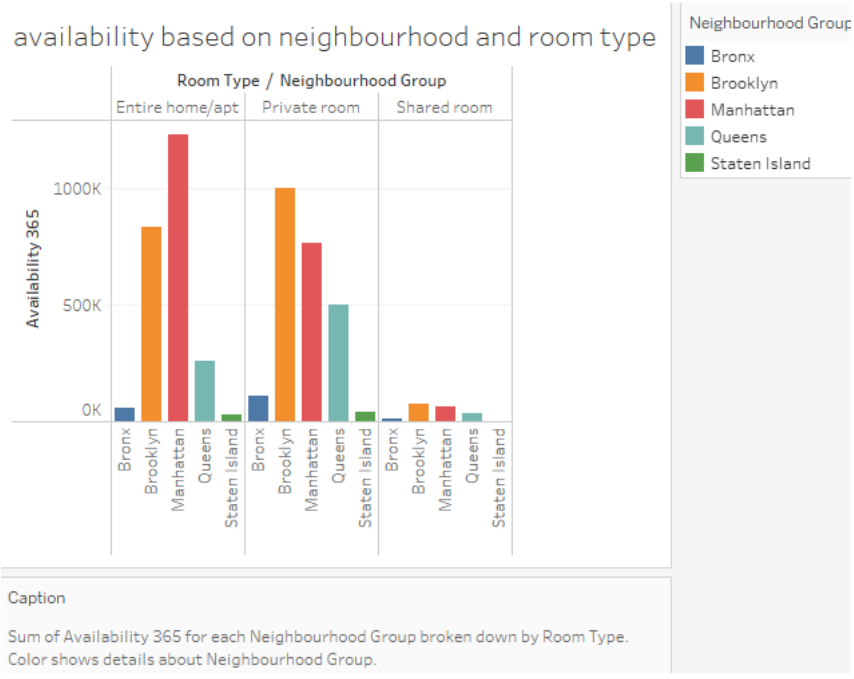
- Here we can see that few Hosts though they have less no of listings have more no of reviews which states they provide very good service at their property

Price of Room Type in each Neighborhood Group:



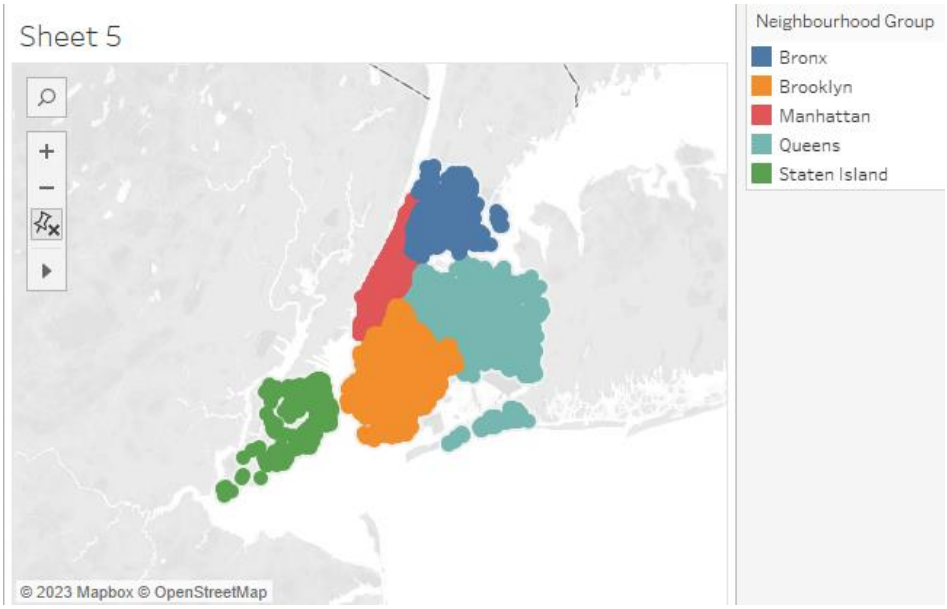
- Prices are higher in Manhattan and Brooklyn where Entire home/apt followed by Private room

Availability of Rooms based on Neighborhood Group and Room type:

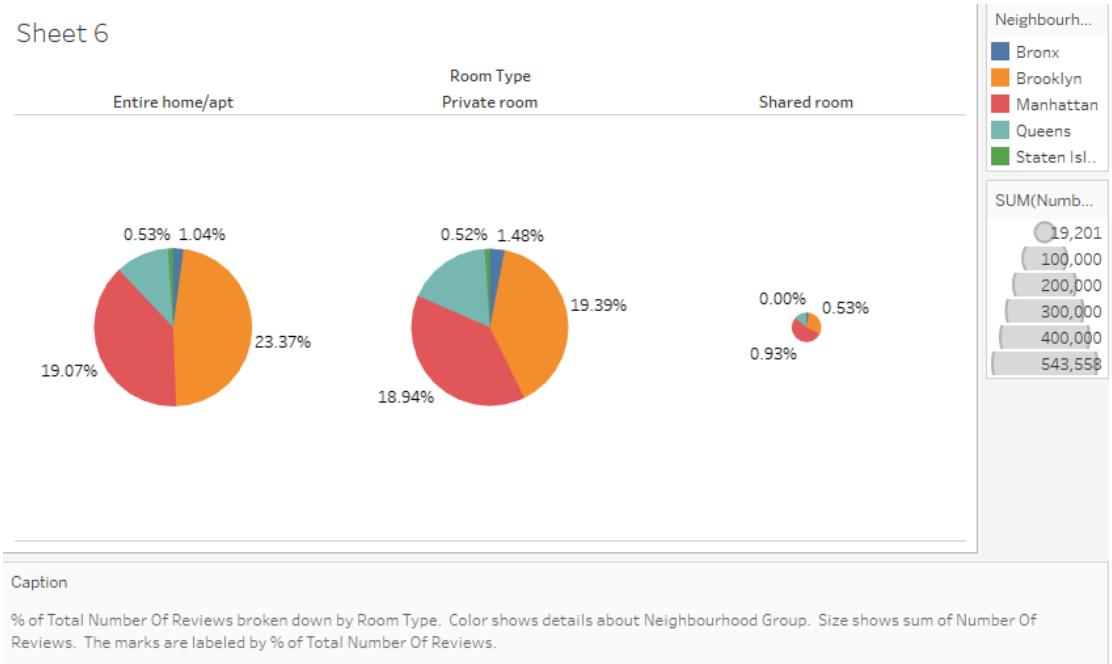


- Manhattan and Brooklyn has more no of Rooms availability with max share in Entire Home/apt next comes the Private Room

Distribution of Rooms based on the Neighborhood Group:

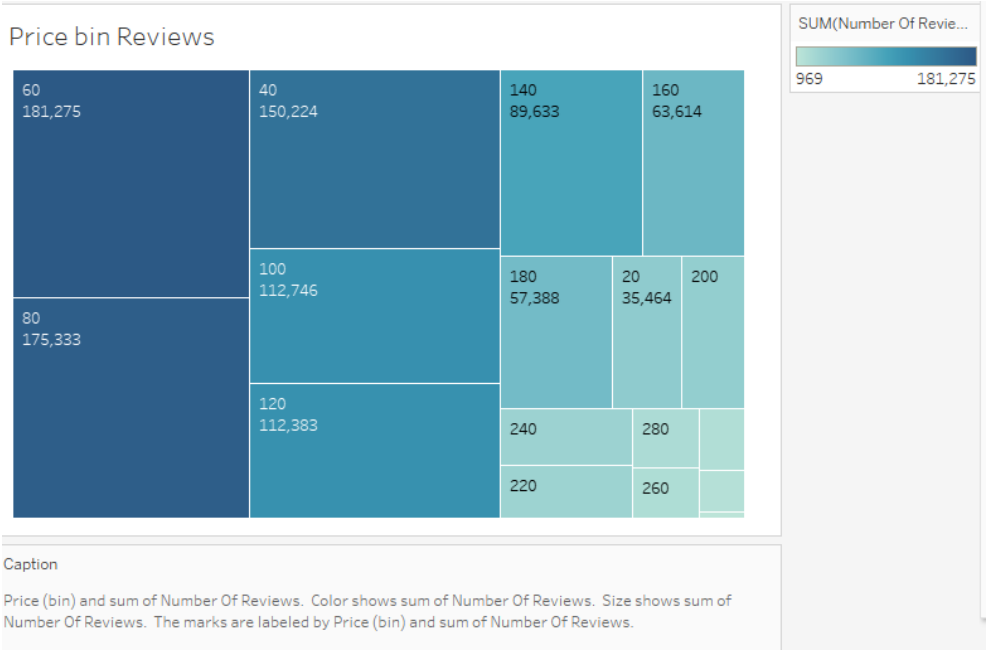


Percentage of Reviews on Each Room Type & Neighborhood Group :



- Entire Home/apt was most attracted by Customers, Manhattan was best Neighborhood Group among them

Most Interested Price Range of the Customers based on Reviews:

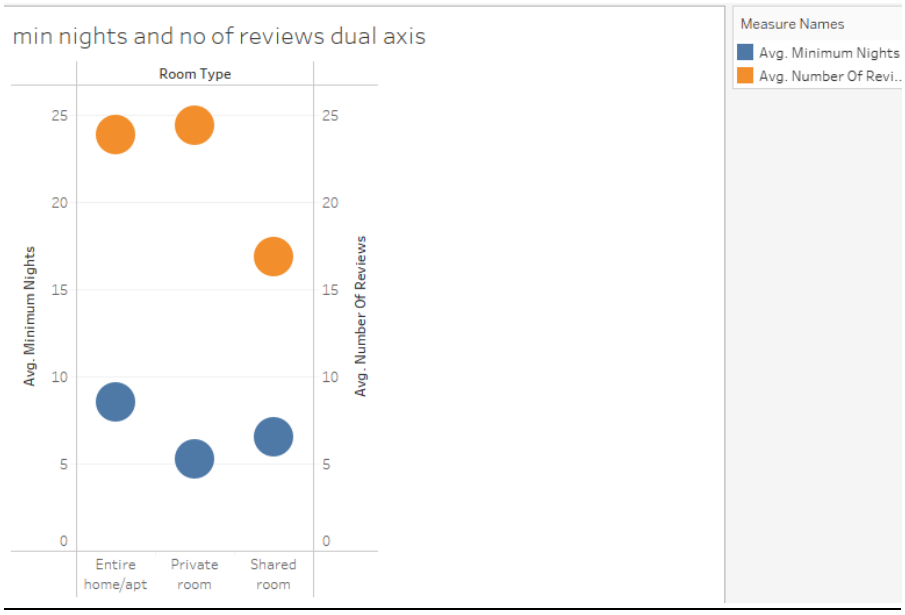


- 40 – 120$ Price Range rooms were most picked by the customers

Max No of Listings:

- Manhattan has the maximum no of Listings Compared to other neighborhood groups

Minimum Nights and Average No of Reviews for a room type:



- For Entire home/apt the avg min nights lie at 8-10 with the Avg no of reviews lies at 23-25
- For Private Room the avg min nights lie at 5 with avg no of reviews at 25
- For Shared room the avg min nights lie at 7-8 with avg no of reviews at 16-17