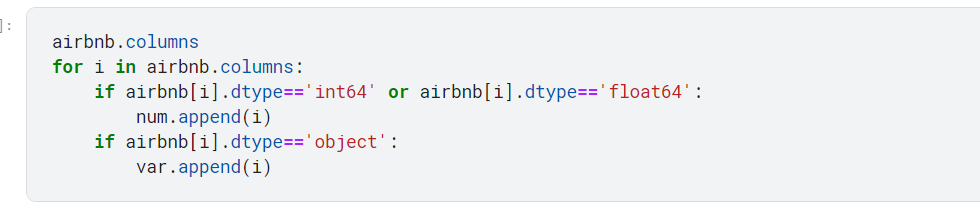
AirBnB – NY Case Study

Technology / Tools used :

Pandas, Numpy, Pyplot, Tableau, Microsoft power point.

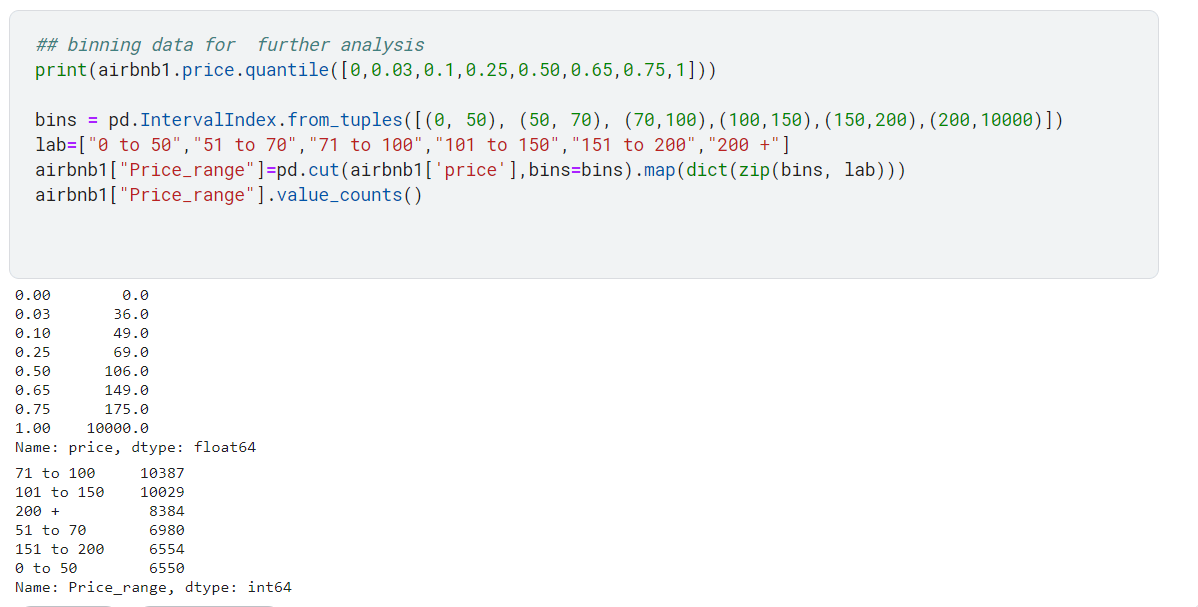
Steps followed

1. Based on case study description, noted down the important questions that needs to be answered for this analysis.
2. Data analysed using python , pandas , numpy and pyplot.
3. Unimportant columns like Id ,Host Id, Name etc are removed before analysing the data. Then divided the columns to 2 types , objects and numeric to further analyse it.



1. Null and outliers were expected and handled based on usability of columns.
2. For null, its replaced with 0 and kept as is for total reviews and last review date column respectively.
3. For measures there were outliers. As all outliers can be completely valid and useful. We have kept it as it is and wherever required binning is used to group the data. Also, instead of mean / average , median is used in analysis for rest of the fields.
   1. Price: Based on quantiles and outliers , bins are created.
   2. Minimum nights: Based on quantiles , bins are created.
   3. Reviews/ Reviews per month: Medians are used in analysis.
   4. Calculated host listing and availability 365: We used them as dimensions for better analysis.

Price:



Minimum nights



1. Then, clean data imported to tableau for further analysis and visualization
2. In tableau, stories are created using components like bars, bubbles, map, pie charts, Gantt charts.
3. All important visualizations are then added to PPT to create the data story for stakeholders.