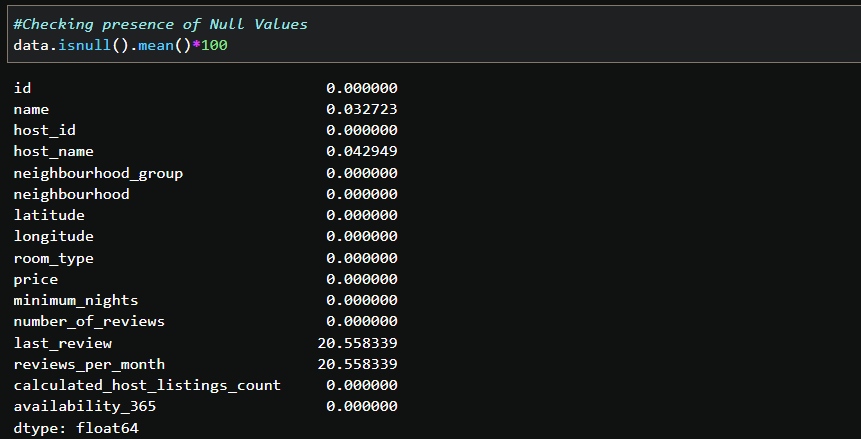
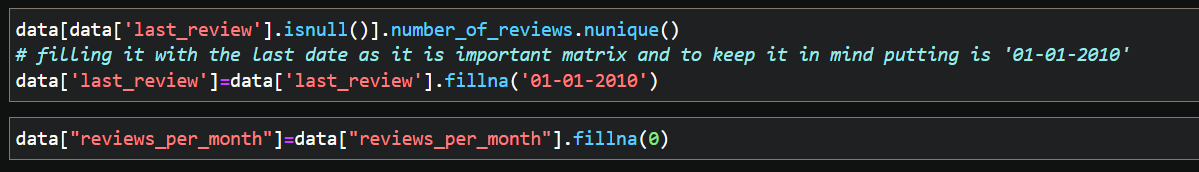
**Methodology Document For Airbnb Casestudy**

**Null Value Treatment**

To maintain the integrity of our data analysis, we have identified the presence of null values in four columns. While these values are not significantly high, we can choose to delete the corresponding rows. However, in order to retain the information and ensure accurate interpretation, we have replaced the null values with appropriate contextual information. This approach allows us to preserve the data and continue our analysis effectively.

****



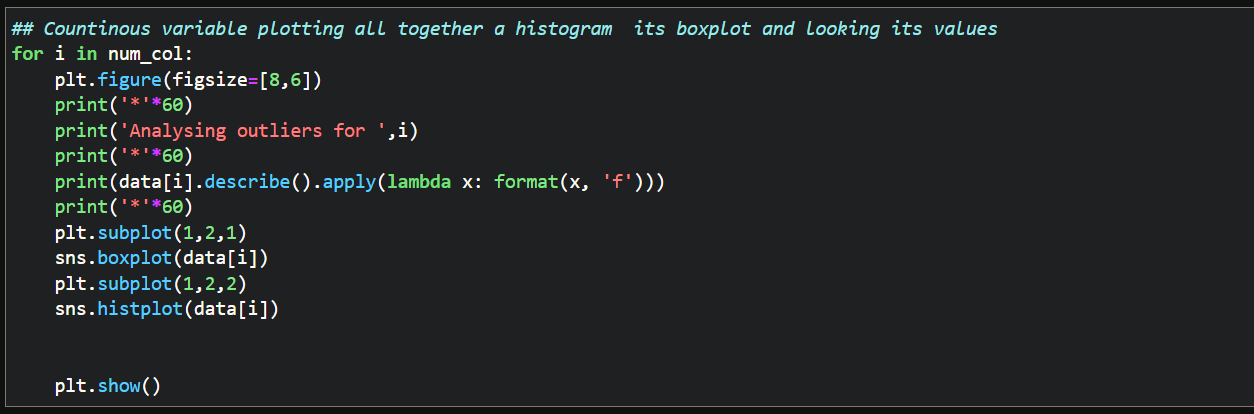
A screen shot of a computer program

Description automatically generated

**Analyzing the Outliers**

In our data analysis, we have observed the presence of outliers in the following columns: price, minimum night, number\_of\_reviews, reviews\_per\_month, and calculated\_host\_listings\_count. To analyze these outliers, we utilized a for loop to generate box plots and histograms for each column. Additionally, we calculated the relevant statistics to gain further insights into the nature and extent of these outliers. This approach allowed us to thoroughly examine and understand the impact of the outliers on our data.A screenshot of a computer program

Description automatically generated

****

**A screenshot of a computer

Description automatically generated**

After conducting a thorough analysis, we decided to explore the option of capping the variables at the 0.99 quantile. However, upon careful consideration, we have chosen not to implement this approach and have decided to retain the variables in their original form.A screenshot of a computer

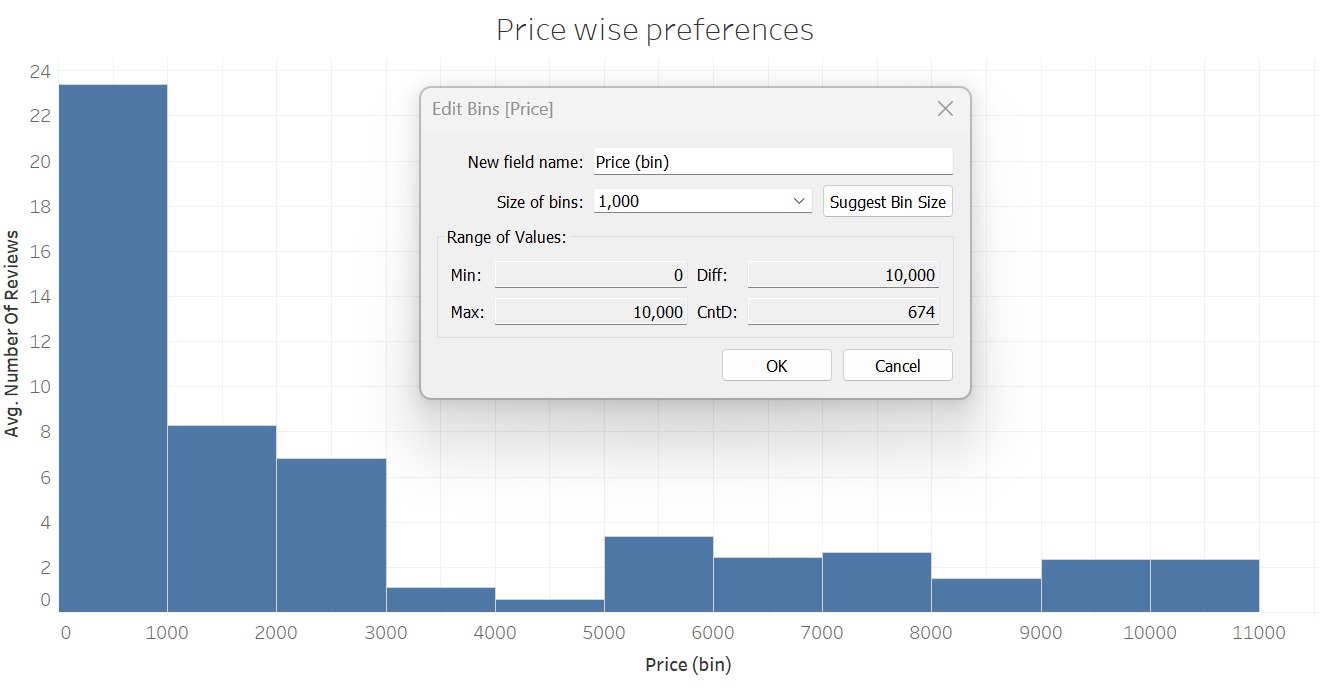
Description automatically generated

**A screenshot of a computer

Description automatically generated**

**Visualization**

We used Tableau to create most of our visualizations and created bins for the price column.



We created a filter for host\_id to show the top 10 hosts based on the Number of Reviews

A screenshot of a computer

Description automatically generated

**Tools Used:**

* Data cleaning and preparation: [Jupiter Notebook – Python](file:///E:\Data%20Science\DS%20Main%20Course\Assignments\Data%20Story%20Telling\AirBnb_case_study.ipynb)
* Visualization and Analysis: Tableau
* Data Storytelling: Microsoft PowerPoint

