**Technical Task Summary**

1. I did an investigation of the data provided like dataset tendencies (mode, mean) and distribution. After finding out there is data missing it was time to understand the structure of data to determine suitable cleaning and processing methods.
2. Filled the missing data using mean for numerical missing values and mode for categorical missing values.
3. Used Pearson Correlation, because it measures the linear relationship between two continuous variables.

Correlation Coefficient:

0.6155274811861676

P-Value:

0.0 < 0.05

This meant the correlation is statistically significant.

1. Used “Group By”, because it helps in analyzing data across different categorical segments and then Unstack to transform the grouped data into a tabular form.
2. Created a binary indicator for the presence of the string 'Ribo' in column X\_8.
3. Counted the occurrences of the string 'Ribo' in column X\_8.
4. Visualized the relationship between X\_6 and X\_7, showing counts of categories.

I used:

* Pivot Table – created a table to aggregate data.
* Plot - used matplotlib to plot a stacked bar chart.
* Annotate - counted labels to each segment for clarity.