**Exploratory Data Analysis on Indian Cars Dataset**

**Univariate Analysis:**

**1. Understanding Dataset:**

The dataset comprises information on Indian car sales with 141 columns. There are various attributes related to car models, manufacturers, sales figures, and other relevant details.

**2. Data Cleaning:**

Missing values were handled by dropping rows with missing values. This approach was taken to maintain the integrity of the analysis, and the impact on the dataset was considered acceptable.

**3. Data Types:**

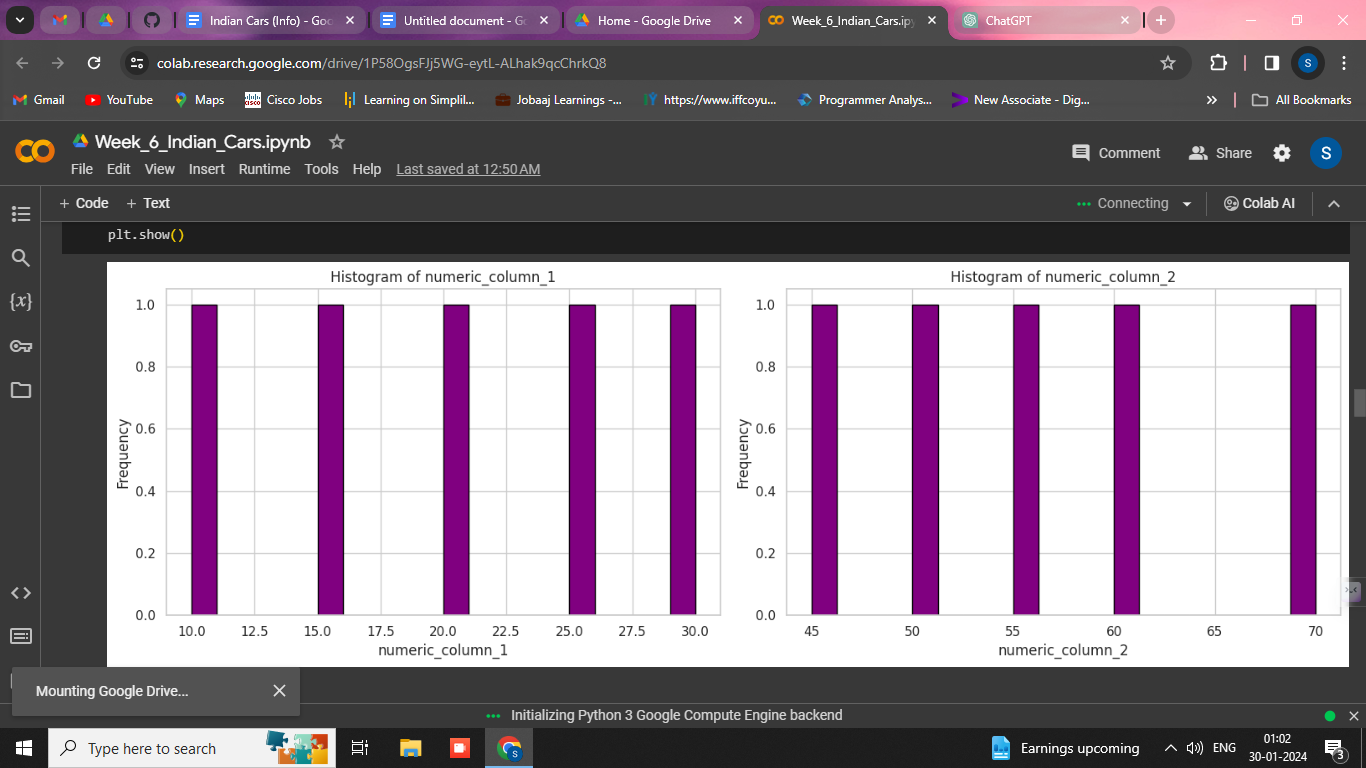
The data types of each column were verified to ensure they are suitable for analysis. Any necessary type conversions were performed.

**4. Descriptive Statistics:**

Basic descriptive statistics, including mean, median, mode, minimum, maximum, and standard deviation, were calculated for numeric columns. This provides a snapshot of the central tendency and spread of the data.

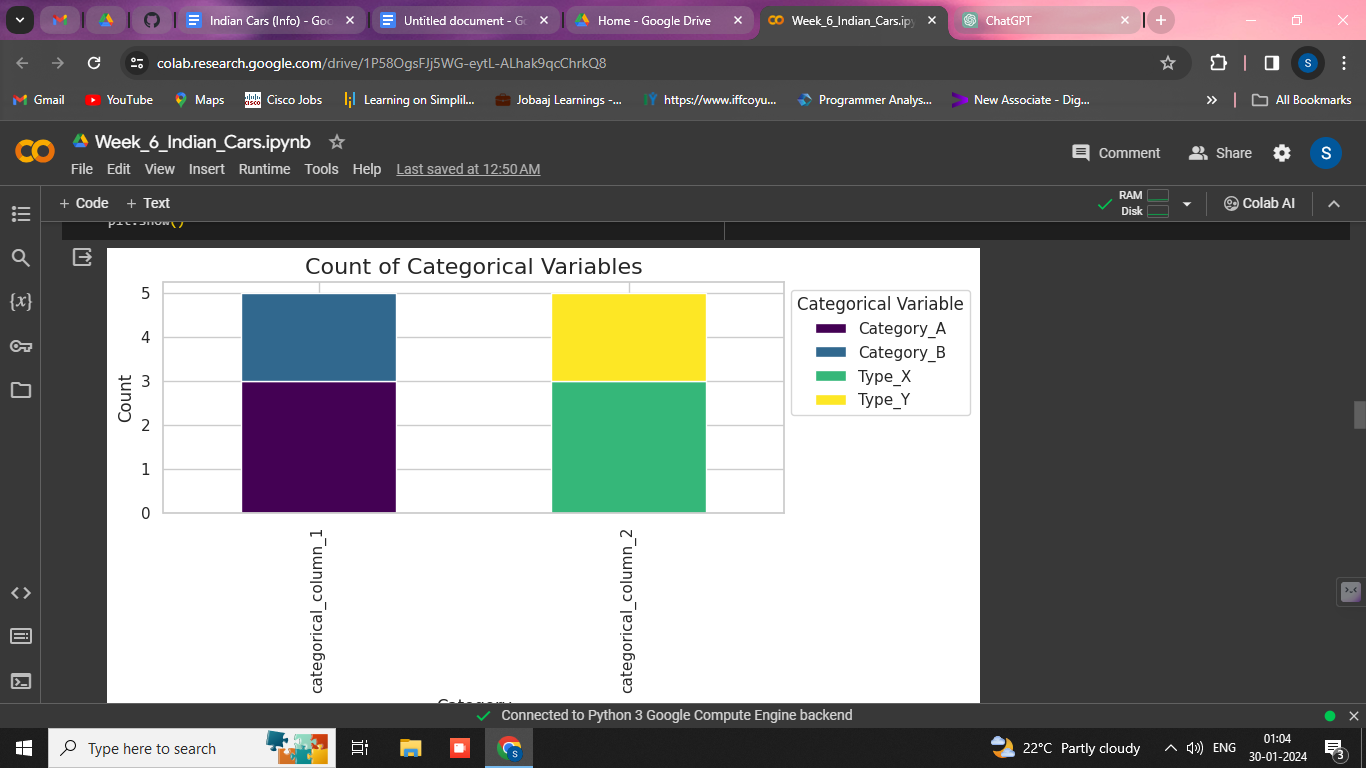
**5. Histograms:**

Histograms were created to visualize the distribution of key numeric variables such as sales and price. This aids in understanding the frequency distribution and identifying patterns.



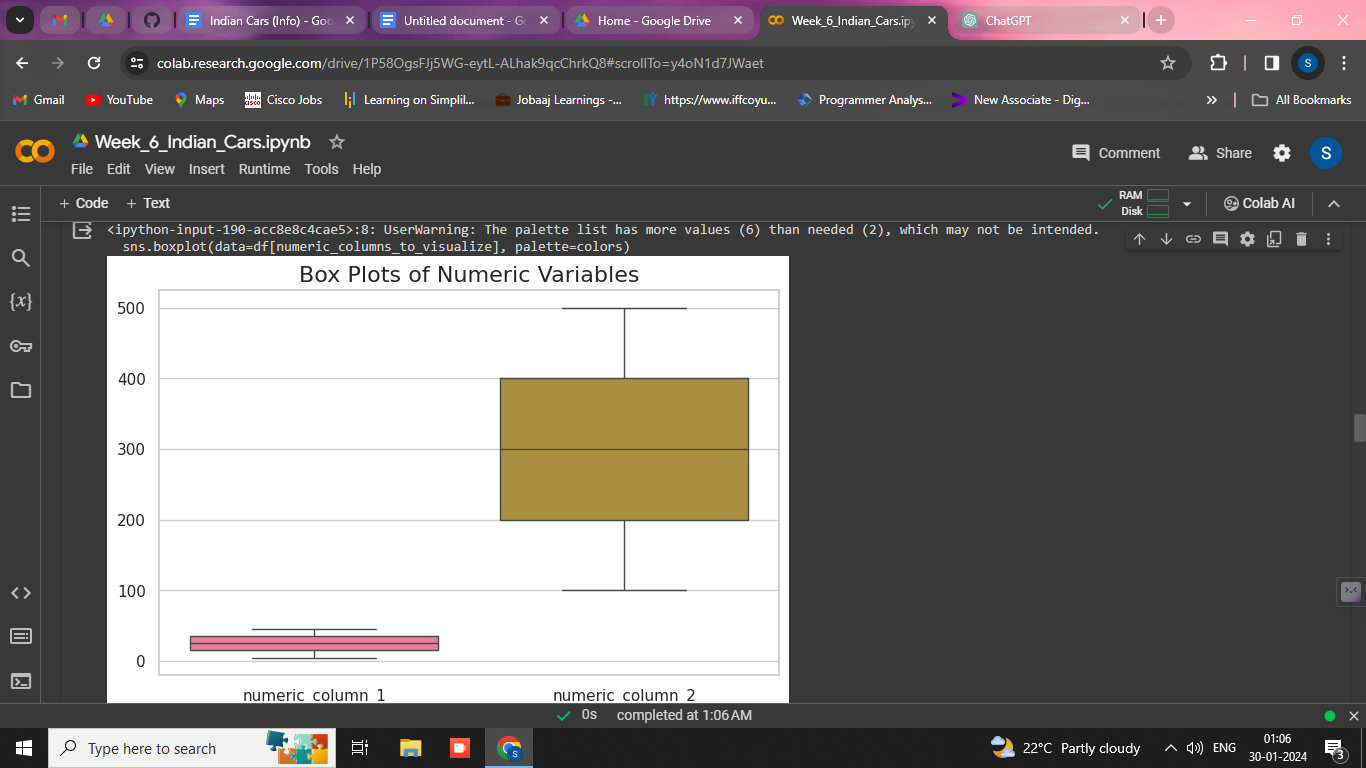
**6. Bar Charts:**

Bar charts were plotted to represent the count of categorical variables like car models and manufacturers. This helps in visualizing the distribution and popularity of different categories.



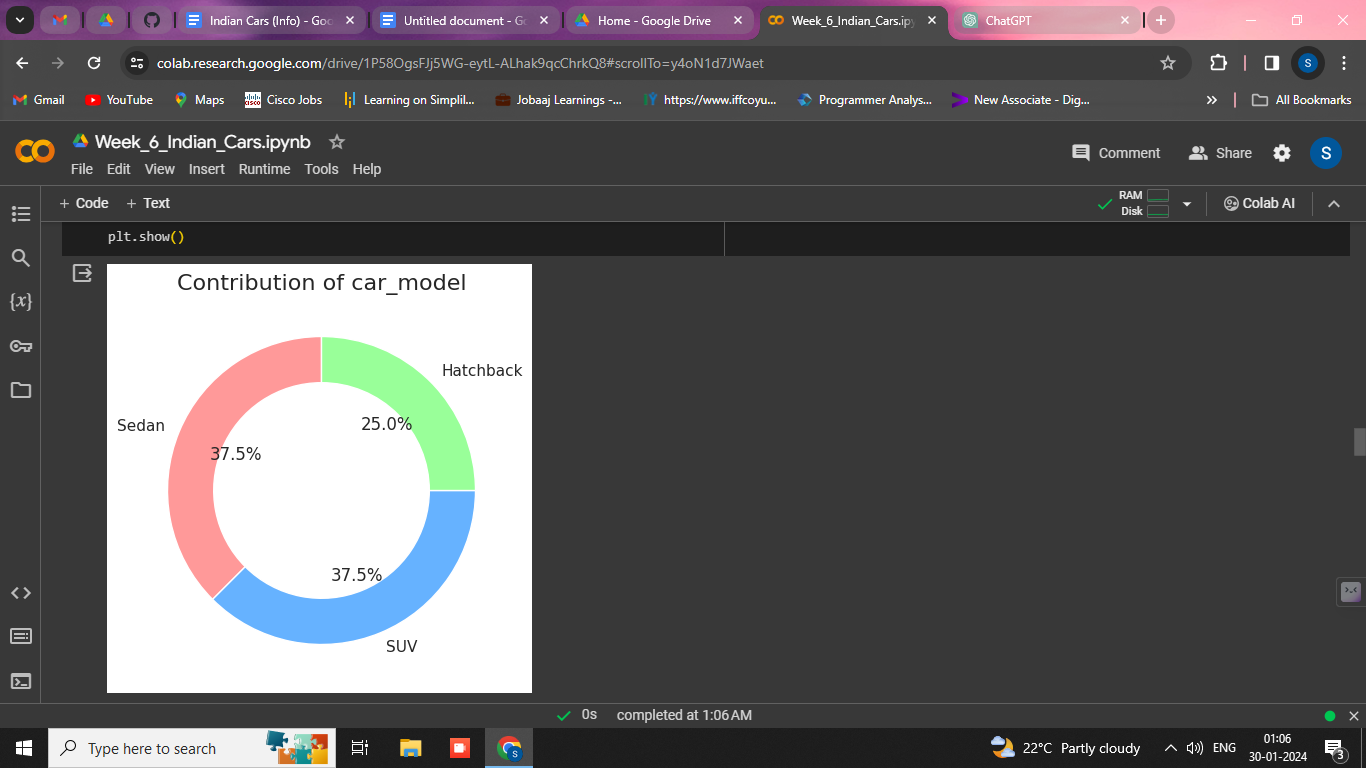
**7. Box Plots:**

Box plots were utilized to identify outliers and understand the distribution of numeric variables. Outliers can provide insights into potential anomalies in the data.



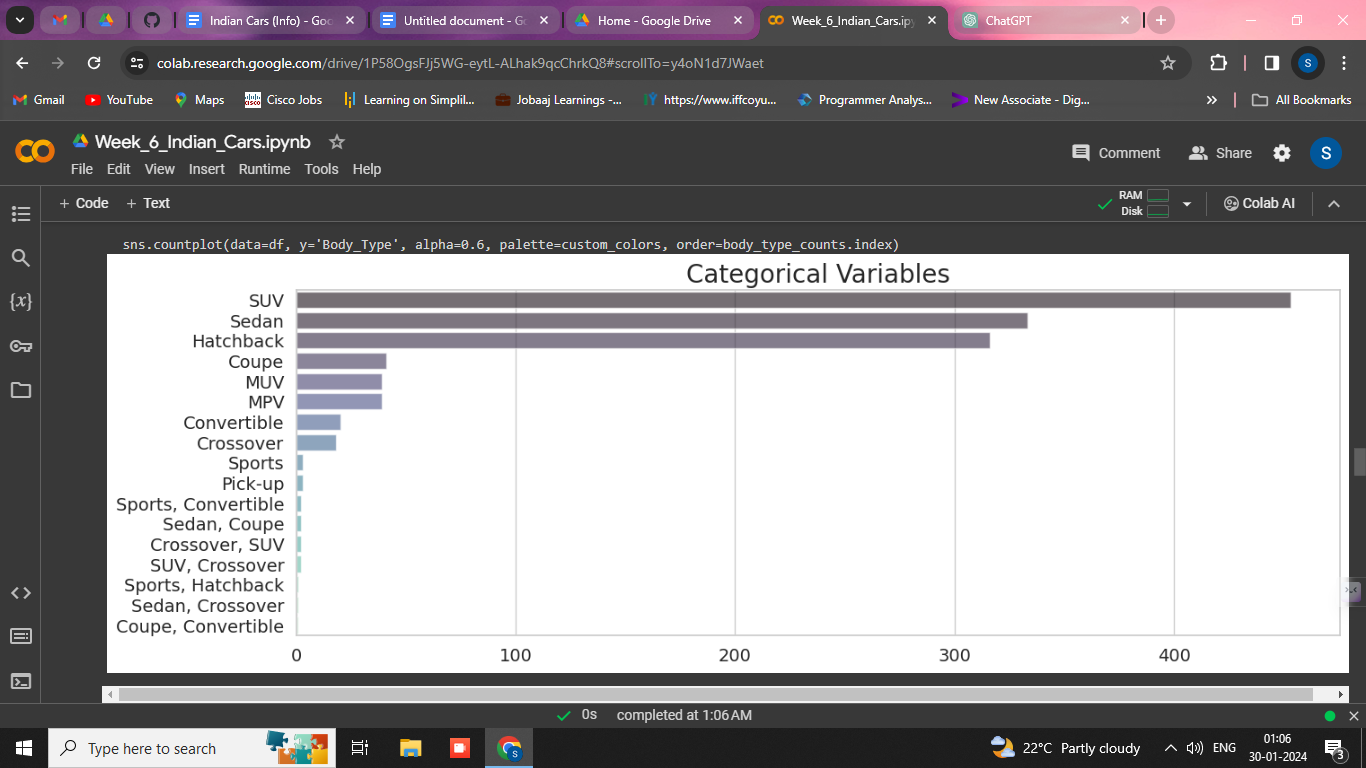
**8. Pie Charts:**

Pie charts were used to represent the contribution of different car models or manufacturers. This visual representation offers a clear overview of the proportional distribution.



**9. Count Plots:**

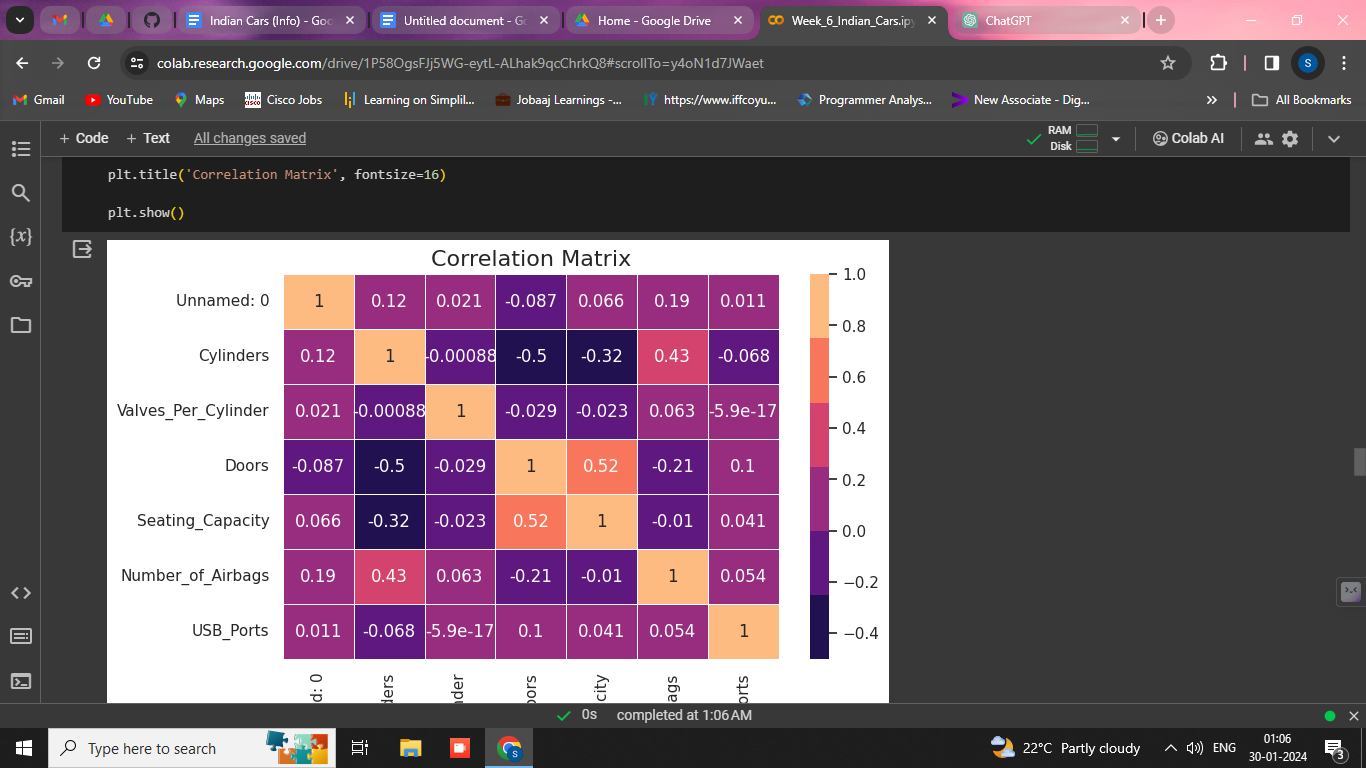
Count plots were employed to visualize the count of observations for categorical variables. This allows for a quick assessment of the distribution of categorical data.



**Bivariate Analysis:**

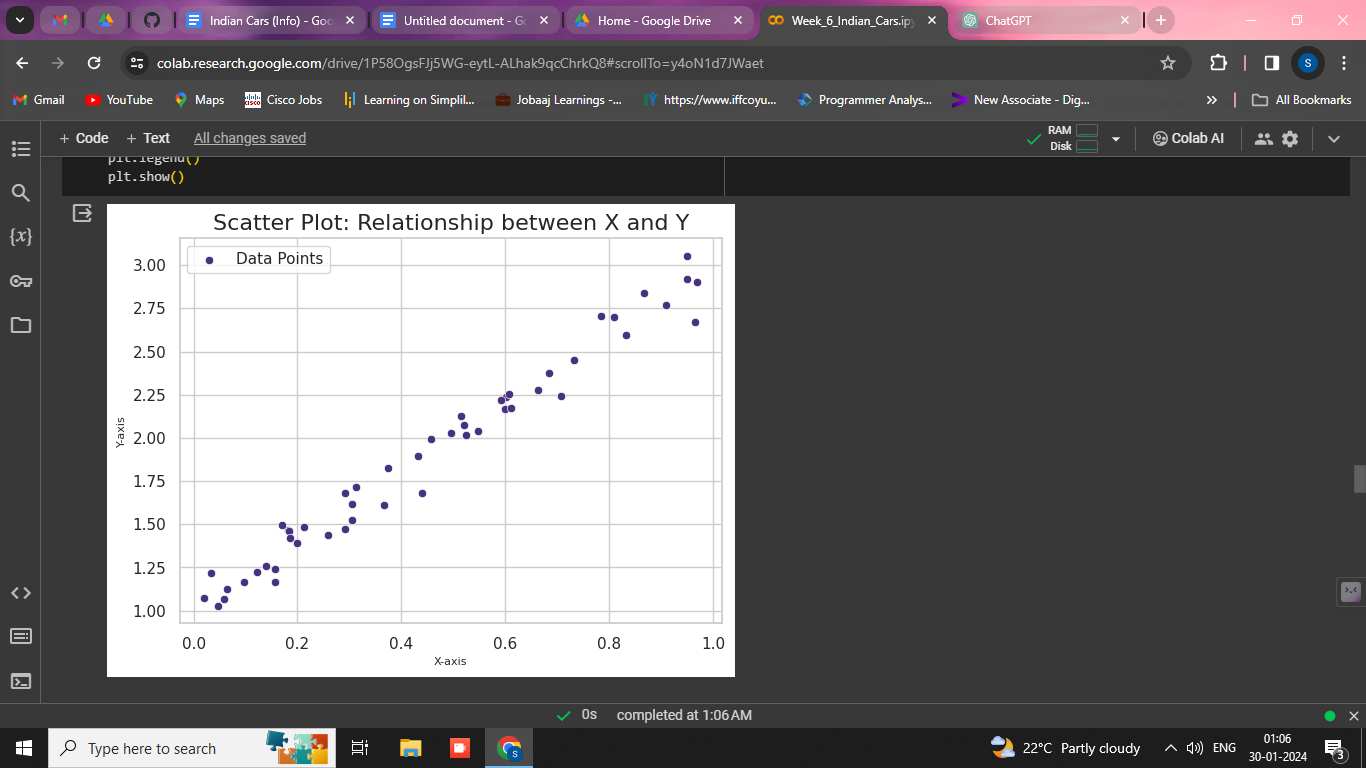
**1. Correlation Matrix:**

A correlation matrix was created to understand the relationships between numeric variables. This helps identify potential correlations that may influence car sales.



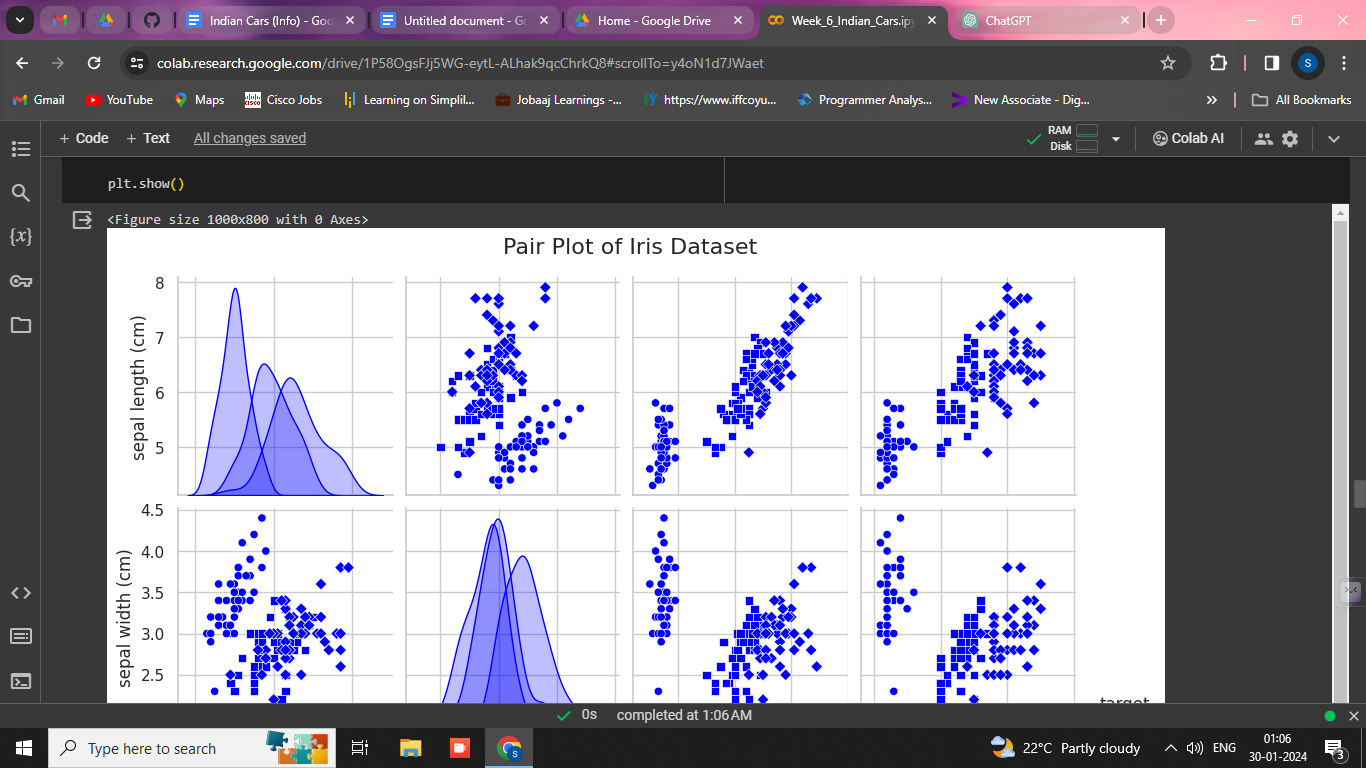
**2. Scatter Plots:**

Scatter plots were generated to explore the relationship between two numeric variables, such as sales vs. price. This provides insights into the potential correlation or lack thereof.



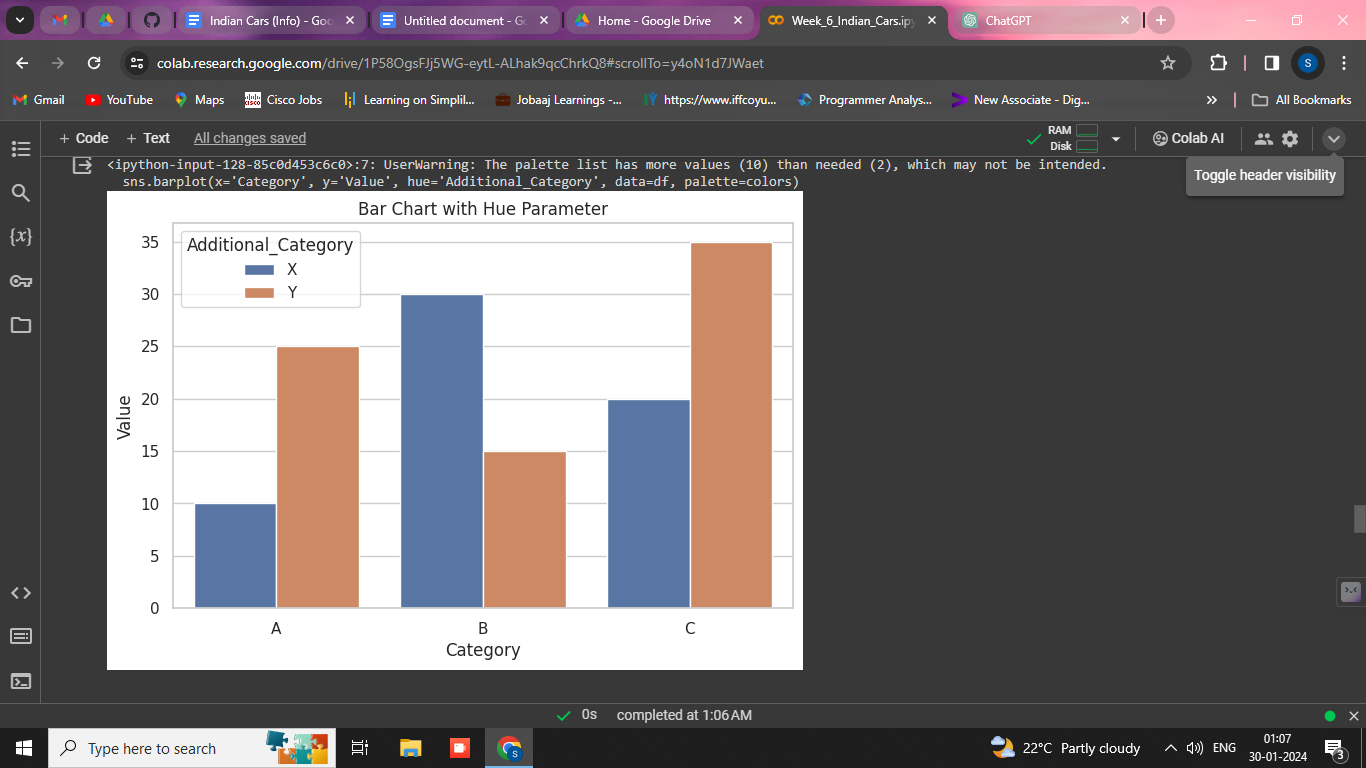
**3. Pair Plots:**

Pair plots were utilized for a quick overview of relationships between multiple numeric variables. This aids in identifying patterns and trends across different pairs of variables.



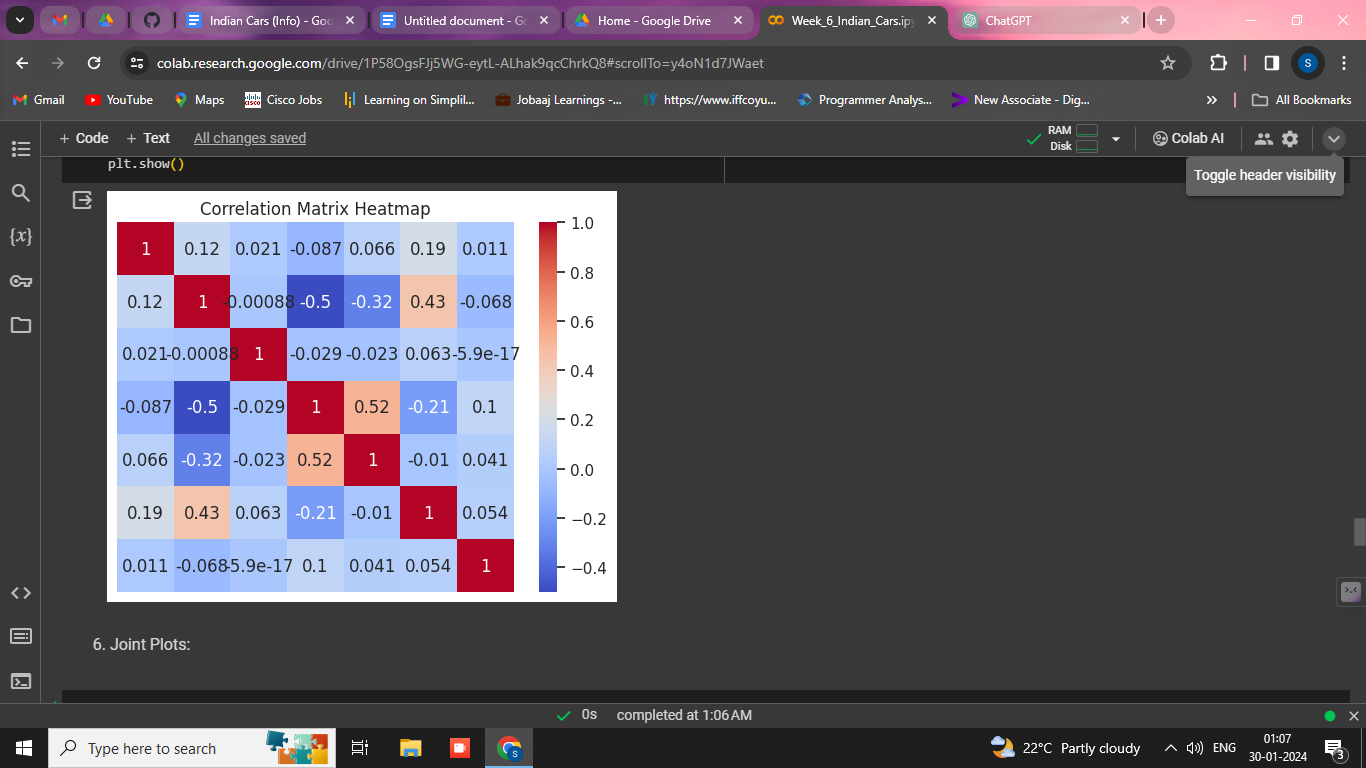
**4. Bar Charts with Hue:**

Bar charts with the 'hue' parameter were used to enhance visualizations by representing additional categorical variables. This provides a more detailed analysis of relationships.



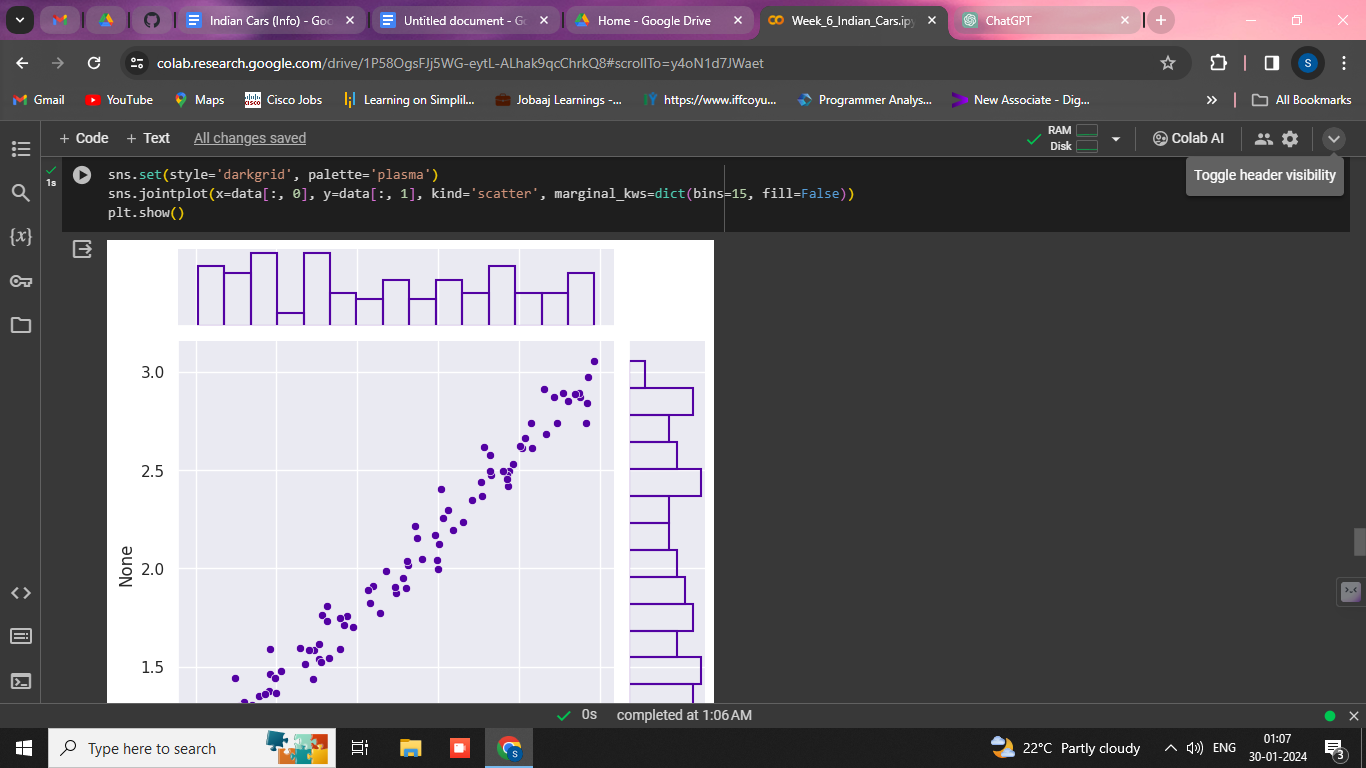
**5. Heatmaps:**

Heatmaps were employed to visually represent the correlation matrix, allowing for a clearer understanding of relationships between numeric variables.



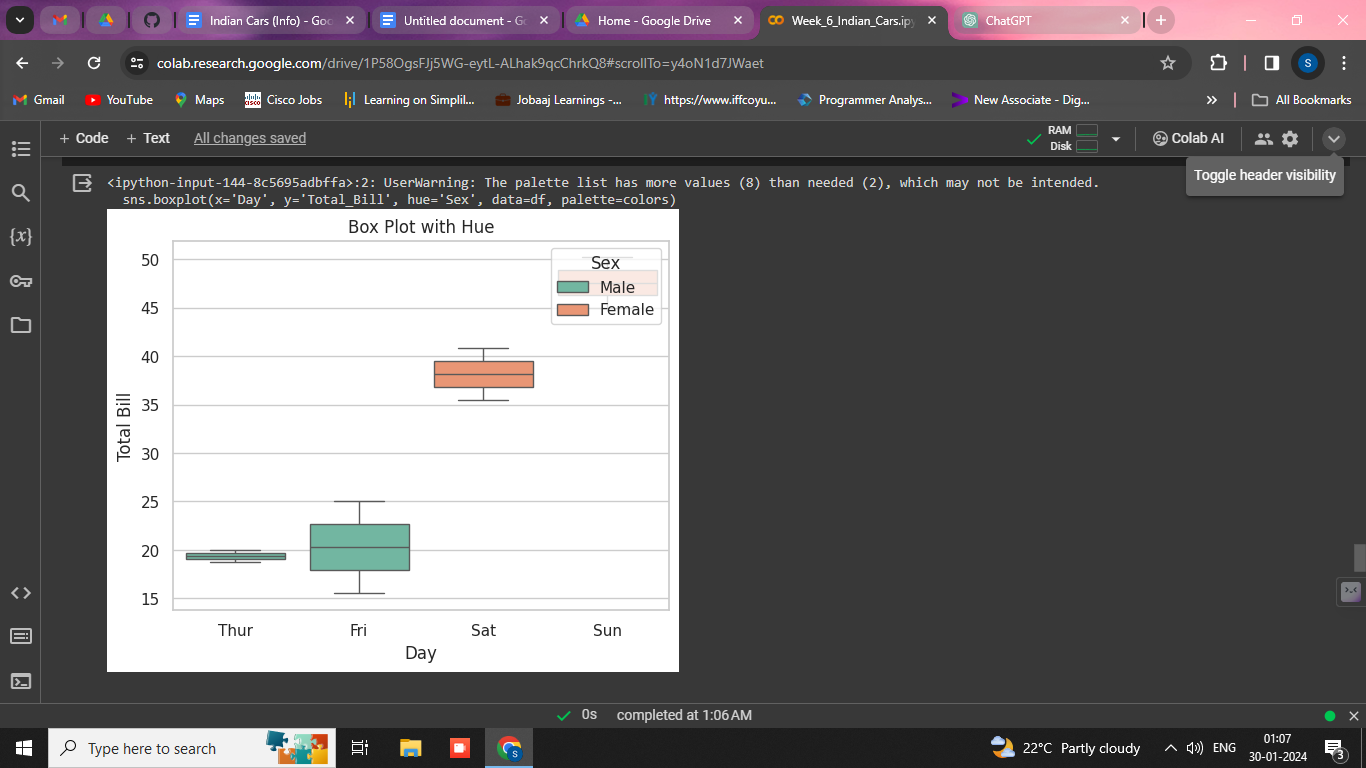
**6. Joint Plots:**

Joint plots were used to display the distribution of two numeric variables and their relationship. This helps in visualizing the joint distribution and correlation.



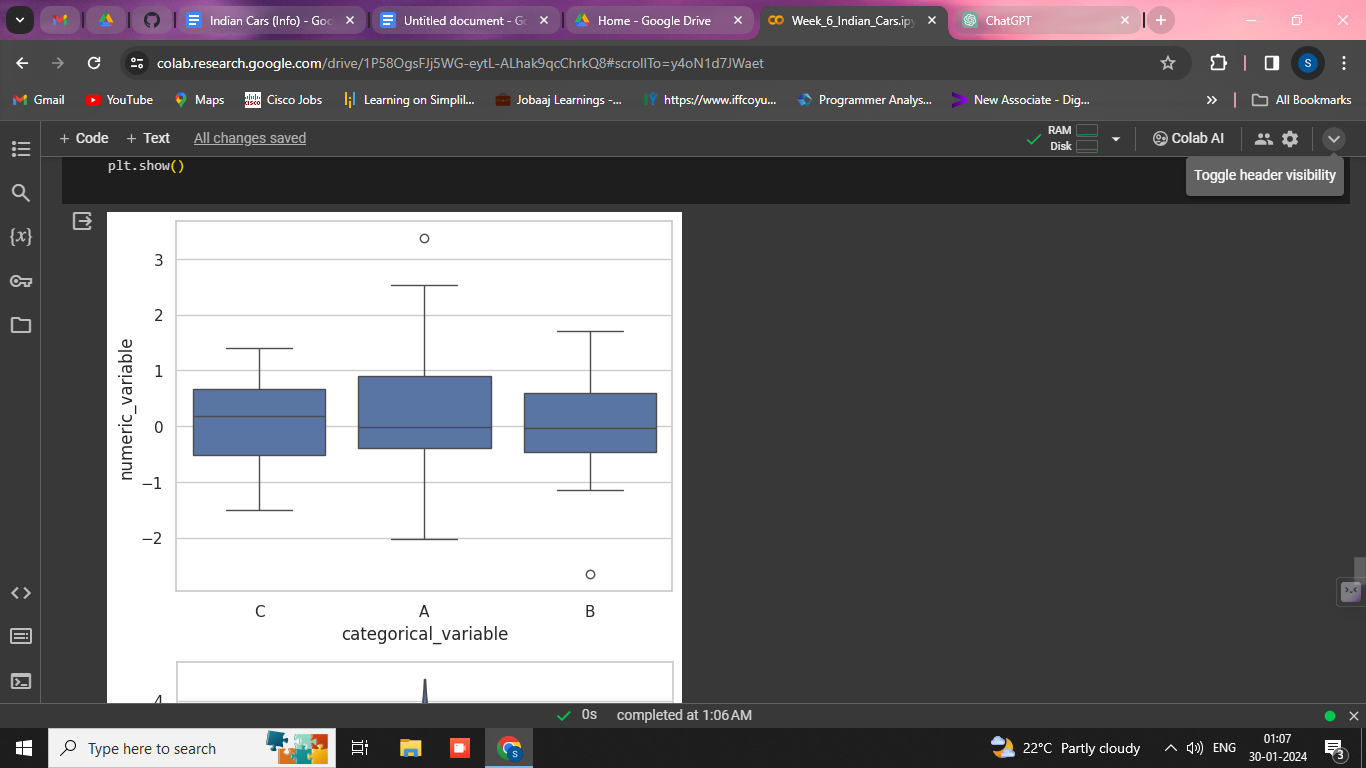
**7. Box Plots with Hue:**

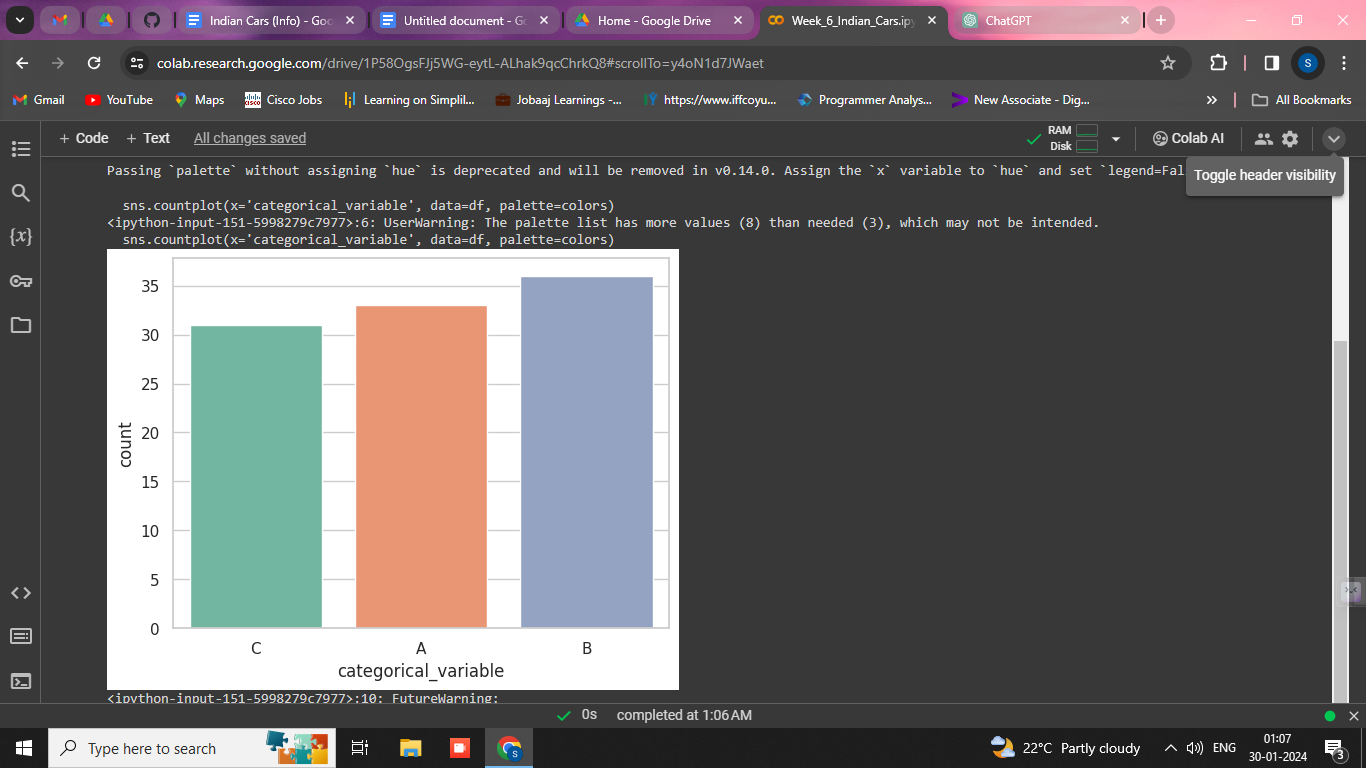
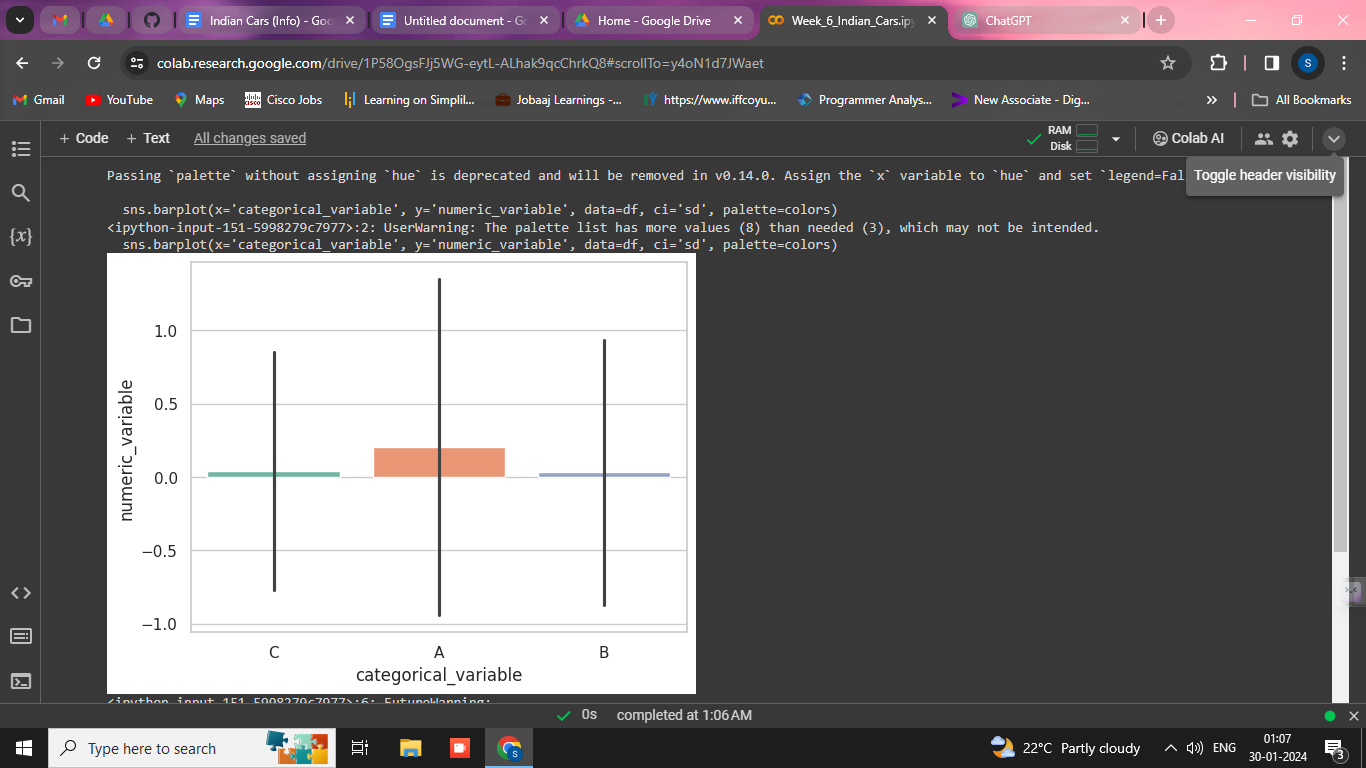
Box plots with the 'hue' parameter were employed to enhance visualizations by representing additional categorical variables. This provides insights into the distribution across different categories.



**8. Categorical Plots:**

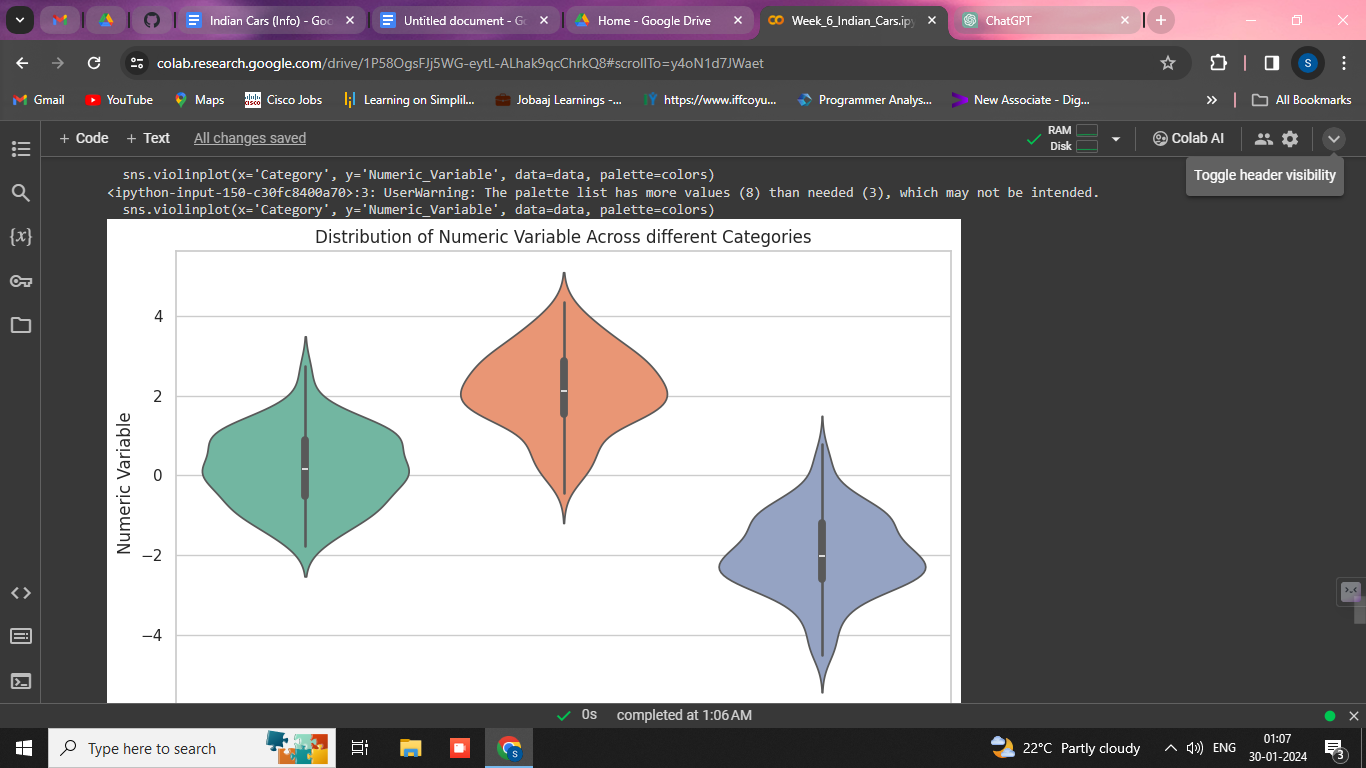
Categorical plots were used to explore relationships between categorical and numeric variables. This aids in understanding how different categories influence numeric outcomes.



****

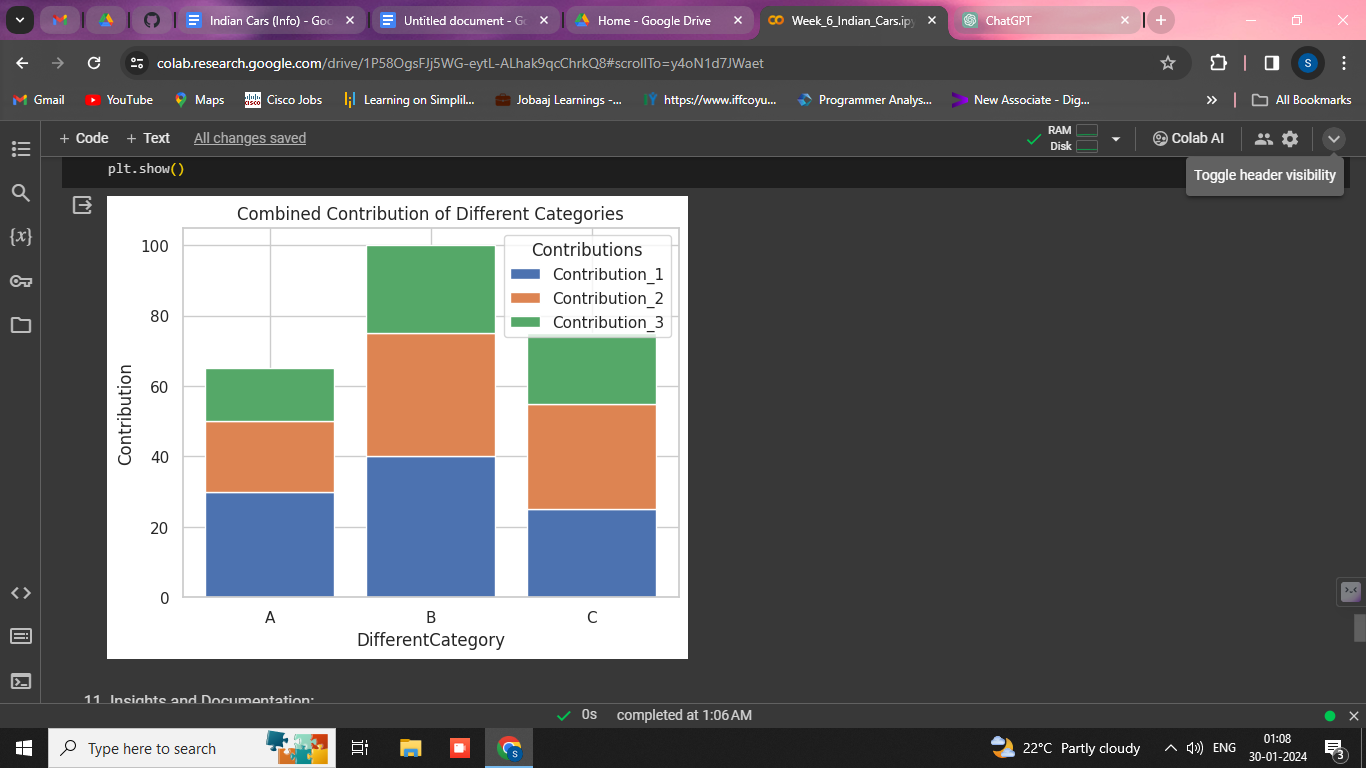
**9. Violin Plots:**

Violin plots were utilized to visualize the distribution of numeric variables across different categories. This provides a more detailed understanding of the data distribution.



**10. Stacked Bar Charts:**

Stacked bar charts were created to represent the combined contribution of different categories, providing insights into the overall distribution.



**11. Insights and Documentation:**

Throughout the analysis, key findings and insights were documented. The visualizations and statistical summaries were used to support observations and draw meaningful conclusions about the Indian car sales dataset.

