**Exploratory Data Analysis (EDA) Report**

**I. Introduction**

The dataset under analysis contains information about car specifications, and the aim of this EDA is to gain insights into various features and their relationships.

**II. Data Overview**

The dataset has been loaded, cleaned, and preprocessed. It includes features such as Power, Torque, Acceleration, CO2 emissions, and more. The dataset has no missing values after cleaning.

**III. Univariate Analysis**

**A. Numerical Features**

1. **Power (HP):**
   * The distribution of power values reveals a peak around a certain horsepower range.
   * The majority of cars in the dataset have a horsepower within a specific range.
2. **Torque (System Performance):**
   * The distribution of torque shows a spread of values, indicating variability among cars.
   * Torque values are diverse, with some cars having higher performance.
3. **Acceleration (s):**
   * The distribution of acceleration times shows a range of values.
   * Acceleration times vary, reflecting the diversity of cars in the dataset.
4. **Top Speed (Kmph):**
   * The histogram depicts the distribution of top speeds among cars.

**B. Categorical Features**

1. **Fuel Type:**
   * The market share by fuel type indicates a percentage distribution among Diesel, Premium Petrol, and Ultra Premium Petrol cars.
2. **Body Type:**
   * The pie chart represents the market share of different car types (Sedan, Hatchback, Estate, Roadster).

**IV. Bivariate Analysis**

**A. Numerical-Numerical Relationships**

1. **Correlation Matrix:**
   * The correlation heatmap shows relationships between various numerical features, including Power, Torque, Acceleration, and CO2 emissions.
   * High correlation between Displacement (CCM) and CO2 emissions.
   * Inverse correlation between Length, Width, Height, and CO2 emissions.
   * Moderate correlation between Power and Emissions, and Curb weight and Power.
   * High correlation between Top Speed and Power.

**B. Numerical-Categorical Relationships**

1. **CO2 Emissions by Fuel Type:**
   * The distribution of CO2 emissions is analyzed for Diesel, Premium Petrol, and Ultra Premium Petrol cars, revealing differences in emission levels.
2. **Year-wise Trends in CO2 Emissions:**
   * A line plot illustrates the trends in CO2 emissions over the years, providing insights into the evolution of emissions.
3. **Market Share by Car Type:**
   * The pie chart represents the market share of different car types (Sedan, Hatchback, Estate, Roadster).

**V. Outlier Detection**

Outliers in the dataset have not been explicitly addressed in the code. Further analysis may be required to identify and handle outliers.

**VI. Missing Values**

The dataset has been cleaned, and no missing values are present in the current analysis.

**VII. Key Findings**

1. **Displacement and Emissions:**
   * There is a high correlation between Displacement (CCM) and CO2 emissions.
2. **Car Dimensions and Emissions:**
   * Length, Width, and Height of the car show an inverse correlation with emissions.
3. **Power and Emissions:**
   * Power and emissions are moderately correlated.
4. **Curb Weight and Power:**
   * Curb weight and power show a high correlation.
5. **Top Speed and Power:**
   * Top speed and power are highly correlated.
6. **Emissions Trends Over Time:**
   * CO2 emissions show a decreasing trend over the years.
7. **Fuel Type Impact on Emissions:**
   * Diesel cars tend to have higher CO2 emissions compared to petrol cars.
8. **Torque and Emissions:**
   * Torque and emissions have a very high correlation.
9. **Diesel vs. Petrol Torque:**
   * Diesel engines produce more torque than petrol engines with the same capacity.

**VIII. Recommendations**

Based on the findings, further analysis could focus on:

* Detailed investigation of outliers and their impact on the analysis.
* Exploring additional features that might influence emissions.
* In-depth analysis of specific car models or manufacturers.

**IX. Conclusion**

This EDA provides valuable insights into the dataset, uncovering relationships between different features and their impact on CO2 emissions. Further analysis and modeling can be performed based on these findings. The corresponding graphs are attached for reference:

1. Histograms for Numerical Features
2. Distribution of CO2 Emissions for Diesel and Petrol Cars
3. Year-wise Trends in CO2 Emissions
4. Market Share by Car Type
5. Correlation Heatmap