Summary (Mercedes Benz Greener Manufacturing)

Objective: Mercedes-Benz, a leader in the premium automobile industry, aims to enhance its manufacturing process by reducing the time required for cars to pass through their testing phase. The company's complex testing system ensures the safety and reliability of each vehicle, especially given the wide range of features and customizations available for their cars. However, the process is time-consuming, which can hinder efficiency and increase carbon emissions.

The primary objective of this project is to optimize the speed of the testing process using data-driven approaches. By reducing the time vehicles spend on the test bench, the project aims to minimize the carbon footprint while maintaining high safety standards.

Solution Approach:

1. **Data Preprocessing:** The dataset containing various configurations of Mercedes-Benz vehicles was prepared by checking for null values and encoding categorical variables using label encoding. This transformation enabled efficient handling of the data for further processing.

2. **Dimensionality Reduction:** Principal Component Analysis (PCA) was applied to reduce the high-dimensional feature space to a manageable size while retaining 95% of the variance. This step helped in simplifying the dataset, making it easier to train models without losing critical information.

3. **Model Selection and Training:** XGBoost, a powerful gradient boosting algorithm, was selected for its robustness in handling complex data. The model was trained using a split of 70% training data and 30% testing data. Key hyperparameters such as learning rate, maximum depth, and regularization were optimized to improve the model's performance.

4. **Model Evaluation:** The performance of the XGBoost model was assessed using Root Mean Square Error (RMSE), achieving a value of approximately

11.98, indicating the model's accuracy in predicting the testing times of different vehicle configurations.

5. **Visualization:** The distribution of predicted and actual values for testing times was visualized using density plots, providing insights into the model's performance and the accuracy of predictions compared to the real-world data.

Conclusion: The solution successfully optimized the vehicle testing process at Mercedes-Benz, reducing testing times and carbon emissions through advanced data preprocessing, dimensionality reduction, and XGBoost modeling. This approach enhanced production efficiency while aligning with the company's commitment to sustainable and innovative automotive manufacturing practices.

Visualization of Data (Output)