Project Summary: Mercedes-Benz Greener Manufacturing

Objective:

The goal of this project is to **optimize the testing time** of Mercedes-Benz cars using **machine learning**. This helps the company **reduce resource consumption**, **increase efficiency**, and contribute to a **greener manufacturing process** by minimizing unnecessary test cycles.

Dataset Overview:

The dataset contains:

- **ID:** Unique identifier for each car.
- **y:** Target variable (test time in seconds).
- **X0–X385:** Features representing various car specifications, some categorical and some numerical.

Project Workflow:

1. Data Preprocessing

- **Handling Missing Values:**
  - Used **median imputation** for numerical data.
- **Feature Encoding:**
  - Categorical features were converted into numerical form using **Label Encoding**.
- **Removing Zero-Variance Columns:**
  - Dropped columns where all values were identical (no variation).
- **Feature Alignment:**
  - Ensured train and test datasets had the same features.

---

2. Feature Reduction (PCA - Principal Component Analysis)

- Applied **PCA** to reduce the number of features while retaining **95% of the variance**.
- Helps in **reducing dimensionality** and improving model efficiency.

---

3. Model Training & Evaluation

- **Split** the dataset into **training (80%)** and **validation (20%)** sets.
- **Trained machine learning models** (e.g., **XGBoost, Linear Regression, Random Forest**) to predict y (test time).
- **Evaluated model performance** using **R^2 score and RMSE (Root Mean Squared Error)**.

---

4. Prediction on Test Data

- Used the trained model to predict test times for new Mercedes-Benz cars.

Expected Impact:

✅ **Reduced testing time** → Faster production cycle.
✅ **Lower energy consumption** → Cost-effective and eco-friendly.
✅ **Optimized resource allocation** → Better manufacturing efficiency.

Conclusion:

This project demonstrates how **machine learning** can help **automate and optimize testing processes**, making **Mercedes-Benz manufacturing more sustainable and efficient**.