**Mercedes-Benz Greener Manufacturing**

**ABSTRACT**

In this project, I was asked to experiment with a real-world dataset of Mercedes-Benz and to explore how machine learning algorithms can be used to find patterns in data, explore and analyze the data. The requirement was to optimize the speed of testing using powerful algorithmic approach. The train dataset is then loaded for the model to train the algorithm. The datasets are explored and the total number of rows and columns are extracted. The given dataset is then plotted on a graph as time series to check for outlier values and variable distribution. The distinct values in a dataset are returned and the features are extracted. Check for null and zero values in the dataset and it is removed. It is also required to apply label encoder using ord and perform dimensionality reduction. Install XGboost package and perform principal component analysis to predict test\_df values.