

# Vehicle Analysis Project

## About Project:

The data set in this project is a vehicle analysis report where it is collecting data to monitor the operations in terms of fuel specifics, vehicle health status, route details, and driver behaviour.

## Objective:

The task is to conduct exploratory data analysis on the provided vehicles dataset. By analysing the dataset, aim is to gain insights into various aspects of vehicles attributes and trends. The analysis will help to understand the data better and identify potential areas for further investigation.

## Task To Perform:

- Data loading and preparation.
- Data cleaning and preprocessing.
- Exploratory data analysis.
- Feature Engineering.
- Insights and Recommendations.

## Approach For Solution:

The vehicle data has total of 1834 rows and 144 columns at start and they are further broken down as float-122, int-13, object-16.

To get insights from the data we have used the following approach to do the analysis.

### Step1: Data loading and preparation.

- To start with the analysis first we have imported the important libraries into our project. Like- Numpy, Pandas, Seaborn, Matplotlib and SkLearn.
- Then we have read the data set into the jupyter notebook and assign it to the variables.
- Basic Exploratory work is performed like-first few rows of the data set, identifying types of variables, size of the data set, missing values and irregularities.

### Step2: Cleaning Of Data Set.

- Missing values were identified and were imputed with median values.
- Columns having all the **NAN** were removed to reduce the memory space of the data set and total Of 6 columns having same values throughout were also removed.
- All the categorical column which was not useful was dropped and our total columns were then reduced from 144 to 24.

- Then we have removed the outliers from the column 'Fuel Efficiency' as we have taken it as a dependent variable in our analysis.
- Duplicates were checked and found no duplicate values.
- Further we have done a correlation heatmap on the numerical features of our data set to get a deep knowledge about how they are correlated the different features.

### **Step3: Exploratory data analysis (EDA):**

- First, Univariate analysis was done on categorical features where we have created a count plot with different categorical features such as – torque Mode, Break Clutch.
- Then we have done univariate analysis on different numerical features and made some of the plots like – kde plot, histogram plot, boxplot. With that we were able to find the extreme values, outliers and their distribution.
- After, that we have done bivariate analysis on different features taking categorical and numerical features and plotting various charts such as – swan plot with torque mode and fuel economy, barplots, scatterplots, and Line graph.

### **Step4: Feature Engineering:**

In this we have created a new feature out of current feature that is Percentage change in engine torque and demand of driver torque with respect to engine torque

### **Step5: Standardization:**

As we could see our data was not distributed normally which result in lot of skewedness to deal with it we have standardized our data so that our data can get fit with different machine learning models properly if needed.

### **Conclusion:**

We can further conclude after performing EDA and feature engineering that our variables or features are very much correlated while the fuel efficiency and fuel rate feature are very much dependent on other features such as engine Torque, Driver demanded Torque, engine load, Accelerator Pedal Position, rpm etc.

Accordingly, feature Engineering was done and new column was added which shows driver pattern of driving the vehicle plays a huge role in fuel efficiency, keeping vehicle at a desired rpm increases the fuel efficiency of the vehicle, higher the engine load lowers the fuel efficiency, fuel efficiency and fuel rate has an inverse relationship.

### **Recommendations:**

To prove our conclusion, we should do further model testing and should conclude how much reliable our model is with further vehicle maintenance also plays a very important role in keeping the fuel efficiency at the lower levels, while proper driver hygiene should be followed to be at optimum level of fuel efficiency and potential modification or tuning should be done as per the matrix.