1. Methodology:
   1. Whenever there is an increase of raw fuel level as compared to raw fuel at previous timestamp, then it is considered as a refueling event else not
   2. If there is a decrease in raw fuel level as compared to raw fuel at previous timestamp and at the same time odo\_distance is not increasing then it is considered as pilferage event.
   3. All the existing records are labelled as refuel event or pilferage event as per the above mentioned logic
2. Data Preprocessing
   1. Created four new columns:
      1. **Refuel\_Event**: The column contains two types of values 1 or 0. If a refueling event has happened, it will contain 1 else 0
      2. **Refuel\_Above\_10L**: If refueling event has happened for more than 10 Litres then this column value is kept 1 else 0
      3. **Pilferage event**: This column also contains two types of values 1 or 0. If a pilferage event has happened, it will contain 1 else 0
      4. **Pilferage\_Above\_10L**: If pilferage event has happened for less than 10 Litres then this column value is kept 1 else 0
   2. Checked missing values, duplicate values, outliers and dropping irrelevant columns
      1. One of the outlier value in **temperature** is replaced with 1.5\*IQR rule
   3. Implemented Standardization to bring the column values to same scale
   4. Converted **ts** column from **object** type to **datetime** so that respective visualizations can be created
3. EDA:
   1. Univariate Analysis is performed by using countplots
   2. Bivariate Analysis is performed using scatterplots, and boxplots
   3. Correlation matrix is created for all the numerical columns
4. Findings:
   1. Most of refueling and pilferage events happened in the month of Jan and February
   2. Most of the refueling and pilferage events happened when the temperature is between 20 and 30
   3. Most of the refueling and pilferage events happened when the Raw Fuel Level is between 125 and 150 Litres
5. ML Algorithms:
   1. Logistic Regression, Gradient Boosting, and Random Forest techniques are trained on existing data to predict refueling and pilferage events