

## Fuel Consumption Rating

The task is to predict the CO2 emission of the vehicles and to classify the vehicles based on the categorical variables, and answer the following questions:

1. Describe the steps required to train a model.
2. Attempt to use all the numerical continuous variables in the dataset provided to build a model to predict the CO2 emission. Perform exploratory data analysis to select a subset of the variables and repeat the procedure. Compare your models and report if there are any differences in the models' performances? Explain your findings.
3. Determine whether there were any noticeable improvements in the CO2 emission from year 2010 to year 2014? Explain your findings.
4. Using each categorical variable as the target variable at each instance, determine which of the variables performed best in classifying the dataset. Explain your findings.
5. How did you check whether your models did not overfit?
6. State the performance measure(s) you were most interested in and the reason(s).
7. Can your models be deployed based on their performances? Explain.
8. Use only the non-categorical variables to form groups. Using internal and external evaluation metrics, determine which categorical variable best describes the groups formed.