

Logistic Regression & LDA

Using the Auto data, create a 0/1 variable, **mpg01**, as an indicator of whether mpg is *greater than the third quartile* (mpg01=1) or *less than or equal* to the third quartile (mpg01=0). You can use the quantile() function for this.

```
source("http://www.uvm.edu/~rsingle/stat3880/data/scripts-3880.R")
library(ISLR)
data(Auto)
```

- 1) Create a visualization of the relationships among mpg01 and the other variables that helps identify which variables might be good predictors for mpg01. Explain why this is a helpful visualization.
- 2) Fit a logistic regression model predicting mpg01 as a function of the **weight** and **year** of manufacture.
 - a. Create a confusion matrix for the predictions using the 0.5 probability threshold for predicting an outcome of mpg01 equal to 1. Use the confusion matrix to compute the accuracy rate for the model predictions.
 - b. State an interpretation of the regression coefficients using appropriate units of measurement.
 - c. Compute the log likelihood for the model and show its relationship to elements from summary(mod).
 - d. Compute McFadden's pseudo- R^2 for the model and interpret it.
 - e. Test your model for goodness of fit and state a conclusion.
- 3) Fit an LDA model predicting mpg01 as a function of the **weight** and **year** of manufacture.
 - a. Create a confusion matrix for the predictions using the 0.5 probability threshold for predicting an outcome of mpg01 equal to 1. Use the confusion matrix to compute the accuracy rate for the model predictions.