

Lab 7: Automatic Model Selection

We will be using the cars2010 data set within the library AppliedPredictiveModeling (recall that cars2010 is the training data set).

Potential predictors include: EngDispl, NumCyl, Transmission, AirAspirationMethod, NumGears, TransLockup, TransCreeperGear, DriveDesc, IntakeValvePerCyl, ExhaustValvesPerCyl, CarlineClassDesc, VarValveTiming and VarValveLift.

The outcome variable is a measure of Fuel Efficiency or FE, which is the unadjusted highway data.

- a. Using the cars2010, perform a *forward selection* (using the p-value criteria at 0.10). What is the final model?
 - i. What was the first variable added?
 - ii. What was the last variable added?
- b. How many **variables** (not parameters or estimates) would result in the final model from using *stepwise selection* with the *BIC* criteria?
- c. Are the two models from a and b the same?