

The sixth assignment includes problems from Chapters 11-12 on the areas of regression diagnostics and variable selection.

1. Chapter 11, problem 10 (page 332). This problem asks you to fit a regression model, to assess if there are influential points, and if so, to do something about it.
2. Chapter 11, problem 20 (page 334). This problem looks at data from dinosaur bones. There are two outliers. The issue is to understand that the usual influence statistics examine changes one point at a time, which may not capture points that are influential in pairs.
3. Chapter 12, problem 10 (page 368). This problem asks you to apply the formulas for C_p , adjusted R^2 , and BIC. In addition to the questions asked, answer these questions.
 - (a) Calculate the AIC for each model.
 - (b) Would the C_p , AIC, and BIC procedures result in the same best model?
 - (c) If you use the AIC criterion for either forward selection or backward elimination, would you find the same model? Is this model the same as the overall best AIC model?
 - (d) If you use the BIC criterion for either forward selection or backward elimination, would you find the same model? Is this model the same as the overall best BIC model?
 - (e) In general, is it possible for the best model under the BIC criterion to have more parameters than the best model under the AIC criterion? Explain. *Hint: Each possible model has a score that is made up of a sum of squares part that is the same for both AIC and BIC plus a penalty. The BIC penalty per parameter is always larger than the penalty per parameter under AIC.*
4. Chapter 12, problem 22 (page 372). Revisiting Bush-Gore 2000 using additional explanatory variables.