In-class Activity: Multiple Linear Regression February 22, 2107

Example: Consider the National Football data. The variables are

Y: number of games won

X1: team’s passing yardage

X2: the percentage of rushing plays

X3: the opponent’s yard rushing

1. Make a scatter plot and comment on the type of association between number of games won and the three covariates. Based on the scatter plot which of the covariates you think is most important to capture the linear trend in y?
2. Fit a multiple linear regression model relating the number of games won to the tree covariates.
3. Construct the analysis of variance table and test for significance of regression.
4. Calculate t statistics for testing the significance of individual regression coefficients? What conclusion can you draw about the roles the variable X1, X2, and X3 play in the model?
5. Calculate R2 and R2Adj for this model.
6. Using the partial F test, determine the contribution of X2 to the model. How is this partial F statistic related to the t-test for β2 calculated above?
7. Fit a model to these data using only X2 and X3.
8. Calculate R2Adj and compare this to that of the full-model.
9. Calculate 95% CIs for the regression coefficients
10. Construct and interpret the joint confidence region for X2 and X3. What conclusion can you draw about the roles these variables play in the model?
11. Construct and interpret the joint confidence region for X1 and X3. What conclusion can you draw about the roles these variables play in the model?