3.21 More British trade unions. Consider the data on opinion polls on the power of British trade unions in Exercise 3.20.

a. Create an interaction term between Months (since August 1975) and Late. Fit the regression model that produces two lines for explaining the NetSupport, one each for the early (Late = 0) and late (Late = 1) parts of the dataset. What is the fitted model?

b. Use a t-test to test the null hypothesis that the interaction term is not needed and parallel regression lines are adequate for describing these data.

c. Use a nested F-test to test the null hypothesis that neither of the terms involving Late is needed and a common regression line for both periods is adequate for describing the relationship between NetSupport and Months.

3.22 More British trade unions. Table 3.7 in Exercise 3.20 also shows the unemployment rate in Britain for each of the months when poll data were collected.

a. Make a scatterplot of $Y = NetSupport$ versus $X = Unemployment$ and comment on what the plot shows.

b. Fit the regression of NetSupport on Unemployment and test whether there is a linear relationship between unemployment and net support for trade unions (i.e., whether the coefficient of Unemployment differs from zero) at the 0.10 level of significance.

c. Fit the regression of NetSupport on Unemployment and Months since August 1975 and test whether there is a linear relationship between unemployment and net support for trade unions, at the 0.10 level of significance, when controlling for time (Months) in the model.

d. How does the coefficient on Unemployment in the model from part (c) compare to that from part (b)? Interpret the difference in these two values.

3.23 Diamond prices. In Example 3.11, we looked at quadratic and cubic polynomial models for the price of diamonds (TotalPrice) based on the size (Carat). Another variable in the Diamonds datafile gives the Depth of the cut for each stone (as a percentage of the diameter). Run each of the models listed below, keeping track of the values for $R^2$, adjusted $R^2$, and which terms (according to the individual t-tests) are important in each model:

a. A quadratic model using Depth

b. A two-predictor model using Carat and Depth

c. A three-predictor model that adds interaction for Carat and Depth

d. A complete second-order model using Carat and Depth

Among these four models as well as the quadratic and cubic models shown in Example 3.11, which would you recommend using for TotalPrice of diamonds? Explain your choice.
3.24 More diamond prices. One of the consistent problems with models for the TotalPrice of diamonds in Example 3.11 was the lack of a constant variance in the residuals. As often happens, when we try to predict the price of the larger, more expensive diamonds, the variability of the residuals tends to increase.

a. Using the model you chose in Exercise 3.23, produce one or more graphs to examine the conditions for homoscedasticity (constant variance) and normality of its residuals. Do these standard regression conditions appear to be reasonable for your model?

b. Transform the response variable to be logPrice as the natural log of the TotalPrice. Is your "best" choice of predictors from Exercise 3.23 still a reasonable choice for predicting logPrice? If not, make adjustments to add or delete terms, keeping within the options offered within a complete second-order model.

c. Once you have settled on a model for logPrice, produce similar graphs to those you found in (a). Has the log transformation helped with either the constant variance or normality conditions on the residuals?

3.25 More diamond prices. Refer to the complete second-order model you found for diamond prices in Exercise 3.23(d). Use a nested F-test to determine whether all of the terms in the model that involve the information on Depth could be removed as a group from the model without significantly impairing its effectiveness.

3.26 More diamond prices. The young couple described in in Example 3.11 has found a 0.5-carat diamond with a depth of 62% that they are interested in buying. Suppose that you decide to use the quadratic regression model for predicting the TotalPrice of the diamond using Carat. The data are stored in Diamonds.

a. What average total price does the quadratic model predict for a 0.5-carat diamond?

b. Find a 95% confidence interval for the mean total price of 0.5-carat diamonds. Write a sentence interpreting the interval in terms that will make sense to the young couple.

c. Find a 95% prediction interval for the total price when a diamond weighs 0.5 carats. Write a sentence interpreting the interval in terms that will make sense to the young couple.

d. Repeat the previous two intervals for the model found in part (b) of Exercise 3.24, where the response variable was logPrice. You should find the intervals for the log scale, but then exponentiate to give answers in terms of TotalPrice.

3.27 First-year GPA. The data in FirstYearGPA contain information from a sample of 219 first-year students at a midwestern college that might be used to build a model to predict their first-year GPA. Suppose that you decide to use high school GPA (HSGPA), Verbal SAT score