Problem 4.48

(a) The $R^2 = 0.76$ indicates that approximately 76% of the variability observed in SAT-Math scores can be attributable (or explainable) by the two predictors in the model: PSAT scores and coaching.

(b) A 95% CI for $\beta_2$ is given approximately by

$$b_2 \pm 2\hat{\sigma}_{b_2} \approx (13, 23).$$

We can say that the interval (13, 23) covers the true value of $\beta_2$ with high probability.

(c) Students who were coached are expected to obtain about 19 more points in the SAT-Math that students who were not coach, for any PSAT score. Since the model does not have an interaction, the expected value of SAT-Math scores is:

$$E(y) = \beta_0 + \beta_1 \text{ PSAT score} \quad \text{students with no coaching}$$

$$E(y) = (\beta_0 + 19) + \beta_1 \text{ PSAT score} \quad \text{students with coaching}.$$ 

Problem 4.49

(a) We need two dummies to represent the three political systems. If we define:

$$x_1 = 1 \text{ if country is communist}, \quad x_1 = 0 \text{ otherwise}$$

$$x_2 = 1 \text{ if country is a democracy}, \quad x_2 = 0 \text{ otherwise}$$

then we can use the following model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon,$$

with $\epsilon \sim N(0, \sigma^2)$ and $\epsilon_i$ independent of $\epsilon_j$ for all $(i \neq j)$. 
(b) If $\beta_1$ is positive (negative) then we would conclude that the expected assassination risk is higher (lower) in communist countries than in countries governed by a dictator. If $\beta_2$ is positive (negative) we say that expected assassination risk is higher (lower) in democracies than in dictatorships. If we wished to compare expected assassination risks in democracies and communist nations, we would do so by computing the difference $\beta_1 - \beta_2$.

Problem 5.52

(a) We need only one dummy because there are two types of repellents: lotion and spray. The model is:

$$y = \beta_0 + \beta_1 d_1 + \epsilon,$$

where $y$ is cost per use, $d_1 = 1$ if the product comes in the form of a spray and $d_1 = 0$ if it comes in the form of a lotion, and errors are independent and $\epsilon \sim N(0, \sigma^2)$.

(b) Use JMP.

(c) Since we have only one predictor, in this special case we can use either the $t$-test for $\beta_1$ or the $F$-test for the utility of the model. In fact, with one predictor, $t^2 = F$. Using the $t$-test:

$$H_0 : \beta_1 = 0, \text{ versus } H_a : \beta_1 \neq 0.$$

(d) The $t$-statistic for the test is (from JMP output) 0.24, and the $F$-ratio is $0.24^2 = 0.057$. The $p$-values for both tests are 0.81. Since $0.81 > 0.10$ we fail to reject $H_0$. That is, we conclude that product type is not a useful predictor of cost per use.

(e) Same as before but now we use the maximum hours of protection as the response variable. We find that product type is not useful for predicting effectiveness of the product either.

Note from AC: If these data were presented to me for analysis, I would have most likely fitted two models, one with cost and another one with effectiveness as responses, and in both I would have used the indicator for
type, the remaining continuous variable (either cost or effectiveness) and an interaction between the two.

Problem 4.53

(a) For each of the tests that we need to conduct, we will take the shortcut route that follows: compute $b_j/\hat{\sigma}_{b_j}$ and compare the value to 2. If ratio is bigger than 2, we reject the null hypothesis of a null $\beta_j$.

1. Model 1: Vintage year is associated to log price because the ratio $0.0354/0.0137 = 2.58 > 2$.

2. Model 2: Vintage year, Average growing season temperature and rainfall in August and September appear to be significantly associated to the log price of the wine. Precipitation prior to vintage, however, does not seem to be associated to the price of the wine.

3. Model 3: When Average Sep. temp is added to the model, rainfall in months preceding vintage becomes significant, even though Ave Sep temp itself is not significant. All other predictors continue to be significantly associated to the response variable.

From the results above, it appears that there is little multicollinearity between predictors, with the exception of Sept temperature and rainfall in months preceding vintage. In the presence of collinearity, the parameter estimates would vary noticeably between models as more predictors are added or deleted from the model.

(b) In Model 1, we can conclude that the expected log price value of a case of wine will increase by 0.0354 when vintage year increases by 1. In Models 2 and 3, each of the $\beta$'s must be interpreted relative to the others. For any of the other $\beta$'s we would say: the expected log price of a case of wine increases/decreases by an amount equal to $b_j$ when $x_j$ increases by one and when all other $x$'s remain fixed at some level.

(c) Model 2 should be used because it has the highest $R^2$, the lowest RMSE and relatively few predictors. Model 3 is a close choice, but it has one more predictor and a smidgeon of a higher RMSE value.

Problem 4.54
(a) \( \hat{y} = -4.3 - 0.002x_1 + 0.336x_2 + 0.384x_3 + 0.067x_4 - 0.143x_5 + 0.081x_6 + 0.134x_7 \).

(b) The model appears to fit the data reasonably well. The \( R^2 \) of 0.712 indicates that approximately 71% of the variability observed in the logarithm of audit fees charged to auditee can be explained by the predictors in the model. Further, the significant \( F \)-ratio indicates that the model is useful for predicting fees.

(c) When the number of subsidiaries of auditee increases by one and all other variables remain constant at some level, the log of fees charged to auditees can be expected to increase by \( \log 0.384 \).

(d) Since the \( p \)-value for the hypothesis test for \( \beta_4 \) is \( 0.079 > 0.05 \) we fail to reject \( H_0 \) and conclude that the data do not provide evidence for there being an expected increase in the log of fees when auditee receives an audit qualification.

(e) There is no evidence in this dataset to support the hypothesis that new auditors charge less than the old auditors in any given year. While the estimate of the regression coefficient \( \beta_1 \) is in fact negative, we cannot conclude that it is different from zero.

**Problem 4.59**

Pairs of models that are nested include:
- a nested within d and e.
- b nested within a, c, d and e.
- c nested within e.
- d nested within e.

**Problem 4.63**

(a) There are two predictors, one quantitative \( (x_1, \text{ age of beanie baby in months}) \) and one dummy to represent the two-level qualitative predictor
$(x_2 = 1$ if baby retired and $x_2 = 0$ if baby still in production. The complete second order model is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_2 x_1 x_2 + \beta_4 x_1^2 + \beta_5 x_1^2 x_2 + \epsilon$$

The textbook includes the term $\beta_5 x_1^2 x_2$ as a second order term. I disagree with calling that a second order term. We are multiplying a second order term by a first order term and thus this interaction is a third order term. If you included the term in the model (following the book) we will not take points off, but if you did not include the term (as I think would be more appropriate) you do not have points off either. I will answer the rest of the questions including that third order term.

(b) $H_0 : \beta_4 = \beta_5 = 0$ versus $H_a$ : at least one of the two $\beta$’s different from 0.

(c) $H_0 : \beta_2 = \beta_5 = 0$ versus $H_a$ : at least one of the two $\beta$’s different from 0.

(d) We use the $F$-tests for nested models. The general form of the $F$-ratio is

$$F = \frac{(SSE_R - SSE_C)/p}{SSE_C/(n - k - 1)}$$

where $p$ is the number of parameters in the null hypothesis (the difference in number of parameters between models R and C), $k + 1$ is the total number of parameters in the complete (C) model, and $SSE_c$ and $SSE_r$ are the error sums of squares for the complete and reduced models, respectively.

To test (b):

$$F = \frac{(3689526 - 3618994)/2}{3618994/(50 - 5 - 1)} = 0.43.$$ 

Such a low statistic leads to failure to reject the null hypothesis, so data appear to indicate that the two second order terms are not really contributing information about the value of the babies.

To test (c):

$$F = \frac{(3723332 - 3618994)/2}{3618994/(50 - 5 - 1)} = 0.63.$$ 

Same conclusion: interaction terms are not contributing information beyond what the first order and quadratic terms contribute.
Problem 4.66

The profitability hypothesis is not supported. The hypothesis says that only industry-wide conditions would be related to profitability. Yet, while industry market share appears to be strongly associated to profitability of the airline, the $F$-test to assess the usefulness of the airline-level dummies indicates that at least some of those dummies are contributing information about airline profitability. Because the dummies represent unchanging characteristics of each airline (such as public perception and preference, quality of service, location of hubs, flight routes, etc), we conclude that profitability appears to be associated not only to industry-wide characteristics but also to airline-level attributes summarized in the dummies.