1. Data were collected from a random sample of 220 home sales from a community in 2003. Let $Price$ denote the selling price (in $1000), $BDR$ denote the number of bedrooms, $Bath$ denote the number of bathrooms, $Hsize$ denote the size of the house (in square feet), $Lsize$ denote the lot size (in square feet), $Age$ denote the age of the house (in years), and $Poor$ denote a binary variable that is equal to 1 if the condition of the house is reported as “poor.” An estimated regression yields

$$\hat{Price} = 119.2 + 0.485BDR + 23.4Bath + 0.156Hsize + 0.002Lsize + 0.090Age - 48.8Poor,$$

$R^2 = 0.72$, $SER = 41.5$

a. Suppose that a homeowner converts part of an existing family room in her house into a new bathroom. What is the expected increase in the value of the house?
b. Suppose that a homeowner adds a new bathroom to her house, which increases the size of house by 100 square feet. What is the expected increase in the value of the house?
c. What is the loss in value if a homeowner lets his house run down so that its condition becomes “poor”?
d. Compute the $R^2$ for the regression.

2. A researcher plans to study the causal effect of police on crime using data from a random sample of U.S. counties. He plans to regress the county’s crime rate on the (per capita) size of the county’s police force.

a. Explain why this regression is likely to suffer from omitted variable bias. Which variables would you add to the regression to control for important omitted variables?
b. Use your answer to (a) and the expression for omitted variable bias given in Equation (6.1) to determine whether the regression will likely over- or underestimate the effect of police on the crime rate. (That is, do you think that $\hat{\beta}_1 > \beta_1$ or $\hat{\beta}_1 < \beta_1$?)

3. A recent study found that the death rate for people who sleep six to seven hours per night is lower than the death rate for people who sleep eight or more hours, and higher than the death rate for people who sleep five or fewer hours. The 1.1 million observations used for this study came from a random survey of Americans aged 30 to 102. Each survey respondent was tracked for four years. The death rate for people sleeping seven hours was calculated as the ratio of the number of deaths over the span of the study among people sleeping seven hours to the total number of survey respondents who slept seven hours. This calculation was then repeated for people sleeping six hours, and so on. Based on this summary, would you recommend that Americans who sleep nine hours per night consider reducing their sleep to six or seven hours if they want to prolong their lives? Why or why not? Explain.
4. \((Y_i, X_{i1}, X_{i2})\) satisfy the assumptions in Key Concept 6.4 (Least Square Assumptions in the multiple regression model). You are interested in \(\beta_1\), the causal effect of \(X_1\) on \(Y\). Suppose that \(X_1\) and \(X_2\) are uncorrelated. You estimate \(\beta_1\) by regressing \(Y\) onto \(X_1\) (so that \(X_2\) is not included in the regression). Does this estimator suffer from omitted variable bias? Explain.

5. Using the data set Growth.dta (attached) described in Empirical Exercise 4.4, but excluding the data for Malta, carry out the following exercises.

a. Construct a table that shows the sample mean, standard deviation, minimum and maximum values for the series \(Growth, TradeShare, YearsSchool, Rev_Coup, Assasinations, RGDP60\).

b. Run a regression of \(Growth\) on \(TradeShare, YearsSchool, Rev_Coup, Assasinations, RGDP60\). What is the value of the coefficient on \(Rev_Coup\)? Interpret the value of this coefficient. Is it large or small in a real world sense?

c. Use the regression to predict the average annual growth rate for a country that has average values for all regressors.

d. Repeat (c) but now assume that the country’s value for \(TradeShare\) is one standard deviation above the mean.