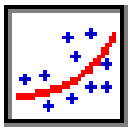
Solutions to Worktext Exercises



Chapter 18

Visualizing Regression Models

Basic Learning Exercises

1. a) The variables are GDP and Period. b) The scatter plot has 24 points that are upward sloping with a slight upward curve. c) The fitted line is nonlinear with an increasing positive slope. d) $\widehat{\text{Ln(GDP)}} = 1.48 + 0.0375$ Period e) Ln is the natural logarithm. f) ^ means the predicted value of Ln(GDP). g) GDP and Period.

2. When the axis labels are put in transformed units, Ln(GDP) in this case, the estimated line is linear.

3. a) On average, 95 out of 100 such 95% confidence intervals will contain the true model. b) On average, 95% of the observed data points will lie within the 95% prediction interval. c) They are closest to the average value of the independent variable, 12.5.
d) They are even more parabolic in shape but still closest to the fitted line at 12.5.

4. The estimated slope (0.037519) is the growth rate *after it is changed to a percent* (3.8%).

5. a) 1960–1997 b) 5.39% c) Inflation has been below the historic average.

6. a) The dependent variable is the firm's Output and Empl is the firm's employment. b) The scatter plot has 36 points that are upward sloping with a slight downward curve. c) The fitted line is nonlinear with a decreasing positive slope. d) $\widehat{\text{Quantity}} = 64{,}251 + 91{,}724$ Ln(Empl)
e) The dependent variable is in linear (not transformed) form and the independent variable is in log form.

7. When the axes labels are put in transformed units, Ln(Empl) in this case, the estimated line is linear.

8. Divide the estimated slope by 100. Since the function is nonlinear, you can't multiply the change due to a 1% increase by 100 to get the change due to a 100% change.

9. a) The independent variable is years in professional baseball and the dependent variable is the player's salary in thousands of dollars. b) A 1% increase in longevity increases a player's salary by \$3,040 (= 304/100 × 1000). It also means that the longer you stay a player the less you earn due to longevity.

10. a) The dependent variable is the Quantity demanded (in units) and the independent variable is the Price of the product in dollars. b) The scatter plot has 75 points. They are scattered in a downward pattern with a slight upward arc. c) The fitted line is nonlinear with a decreasingly negative slope. d) $\widehat{\text{Ln(Quantity)}} = 2.09 - 0.635$ Ln(Price)

11. When the axis labels are put in transformed units, Ln(Quantity) and Ln(Price) in this case, the estimated line is linear.

12. The estimated slope is –0.635. This coefficient estimates the *percentage change* in Y due to a 1% change in X. Because the function is nonlinear, what is true for a 1% change is not true for a 100% change. Therefore, you cannot simply multiply the coefficient by 100.

13. a) The independent variable is Gestation Period measured in weeks and the dependent variable is Birth Weight in ounces. b) A 1% increase in gestation results in a 2.57% change in birth weight. Therefore, increasing the length of time a baby is in the womb has a significant increase in birth weight. c) It would be a negatively-sloped line that gets increasingly steeper as X increases.

14. a) The independent variable is quantity of natural gas used in a month in hundreds of cubic feet and the dependent variable is energy cost for the month in dollars. b) A 1% increase in the quantity of gas used results in a 0.272% change in energy cost. Therefore, increasing the amount of natural gas used increases the cost of energy at a decreasing rate. c) It would be a negatively-sloped line that gets progressively flatter as X increases.

Intermediate Learning Exercises

15. a) The dependent variable is quarterly Profits in millions of dollars and the independent variable is Revenue in millions of dollars. b) The scatter plot has 60 points that are scattered in a rainbow-like arc. c) Quadratic model. d) The fitted line is a parabola with a maximum.
    e) $\hat{\text{Profit}} = 4.21 + 0.438\ \text{Revenue} - 0.0505\ \text{Revenue}^2$

16. a) The 3 slopes are 0.135, 0.034, and –0.067. b) Because the function is nonlinear.
    c) The parabola has a maximum because the squared term has a negative coefficient.

17. a) The independent variable is Time measured in months from 1997:01 to 1999:11. The dependent variable is the Gas Price in dollars. b) That gas prices rose and then decreased in the late 1990's. There is no straightforward interpretation to this model. c) It has a minimum because the squared term has a positive coefficient.

18. A cubic model has both a maximum (Time = 3) and a minimum (Time = 23). It has an inflection point halfway between these two extremes (Time = 13, i.e., 1998:1).

19. a) The dependent variable is Wage in dollars and the independent variable is Job Complexity measured as a scale between 0 and 40. b) The scatter plot has 36 points that are scattered in an upward-sloping line. c) The fitted line is a straight line with a positive slope.
    d) $\hat{\text{Std(Wage)}} = 0 + 0.899\ \text{Std(Cmplx)}$

20. a) The transformation is $(Y_i - \bar{y}) / s_y$, where s_y is the standard deviation of Y. The X transformation is similar. b) This is a linear transformation since Y_i is not raised to a power or transformed by a logarithmic function (or other nonlinear transformation). c) The estimated intercept is always zero because both the Std(Y) and Std(X) have a zero sample mean and the regression line goes through $(\bar{x}, \bar{y})$, i.e., through (0,0).

21. a) The coefficient is called a Beta Weight. b) A 1 standard deviation change in Job Complexity results in a 0.899 standard deviation change in Wages. c) The fitted line did not change since the variables were transformed with a linear transformation.

22. a) The independent variable is the student's grade on Exam 1 measured in tenths between 0 and 4. The dependent variable is the student's Final Course Grade measured in the same way. b) A 1-standard deviation change in exam 1 grade will result in a 0.701 standard deviation change in the final grade. This suggests that the grade on the first exam has a strong relation to the final course grade.

Advanced Learning Exercises

23. a) The dependent variable is Output in thousands of units and the independent variable is Capital in units. b) X or Capital is raised to the 0.2 power. c) The scatter plot has 70 points that are scattered upward with a slight downward arc. d) The fitted line is upward-sloping with a slight downward arc. e) $\hat{Output} = 1.86 + 3.07\ Capital^{0.2}$

24.

Exponent	ESS	Why Select the Next Exponent to Estimate
0.2	1.558	Bracket 0.2 with estimates using 0.3 and 0.1.
0.3	1.601	Since the ESS is larger, we will not be moving further in this direction.
0.1	1.579	Since ESS is smaller, we will decrease the exponent even more.
0.01	1.575	ESS is lower than for 0.1, but we must look between the two numbers.
0.05	1.576	ESS increased, so we must check between 0.05 and 0.01.
0.03	1.575	Same as 0.01, so check if 0.02 is smaller.
0.02	1.575	Function is flat between 0.01 and 0.03; select any of them.
Final Model – $\hat{Output} = -47.2 + 51.5\ Capital^{0.02}$		

25. a) The dependent variable is Quantity in units and the independent variable is Price in dollars. b) The scatter plot has 15 points that are scattered downward. c) The fitted line is downward sloping but with a decreasing slope. d) $\hat{Quantity} = 4.61 + 1{,}110 / Price$ e) The elasticity of this demand function is always –1.

26. a) The dependent variable is the value of a condominium in thousands of dollars and the independent variable is its Square Footage. b) The interaction variable is the number of people living in the condominium. c) The scatter plot has 60 points scattered upward. Some are red and some are blue corresponding to the number of people living in the condominium. d) $\hat{Value} = 56.1 + 0.0830\ SqFt + 0.0119\ (SqFt)(People)$ e) There are two lines: red when People = 1 and blue when People = 2. The lines are both upward sloping but the one for People = 2 is steeper.

27. a) <0.001 b) Two people living in a condominium will pay more for a *larger* condominium than will one person. c) Yes, since the interaction term is significantly different from zero, (SqFt)(People) is crucial in the interpretation.

28. a) The dependent variable is Income in thousands of dollars and the independent variable is years of Experience. b) The interaction variable is a categorical variable ranging from 1 to 4 describing the amount of education. c) $\hat{Income} = 27.9 + 0.393\ Exp + 0.433\ (Exp)(Ed)$. d) There are four lines, one for each category of Education: 1, high school diploma; 2, some college; 3, college degree; and 4, post baccalaureate degree. The fitted lines all have positive slope. The slope increases with the amount of education.

29. a) < 0.001 b) Employees with more education get *larger* salary increases per year than those with less education. c) Yes, since the interaction term is significantly different from zero, (Ed)(Exp), the years of Experience would be crucial in the interpretation.