SOUTH-WESTERN
CENGAGE Learning

Chapter 12
Simple Linear Regression

Learning Objectives

1. Understand how regression analysis can be used to develop an equation that estimates mathematically how two variables are related.

2. Understand the differences between the regression model, the regression equation, and the estimated regression equation.

3. Know how to fit an estimated regression equation to a set of sample data based upon the least-squares method.

4. Be able to determine how good a fit is provided by the estimated regression equation and compute the sample correlation coefficient from the regression analysis output.

5. Understand the assumptions necessary for statistical inference and be able to test for a significant relationship.

6. Know how to develop confidence interval estimates of *y* given a specific value of *x* in both the case of a mean value of *y* and an individual value of *y*.

7. Learn how to use a residual plot to make a judgement as to the validity of the regression assumptions.

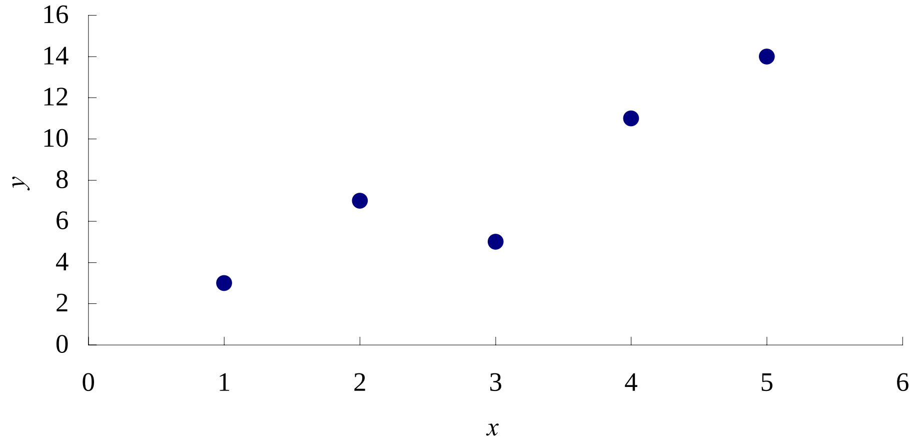
8. Know the definition of the following terms:

   independent and dependent variable
   simple linear regression
   regression model
   regression equation and estimated regression equation
   scatter diagram
   coefficient of determination
   standard error of the estimate
   confidence interval
   prediction interval
   residual plot

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.
STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

Chapter 12
Solutions:

1 a.



b. There appears to be a positive linear relationship between *x* and y.

c. Many different straight lines can be drawn to provide a linear approximation of the relationship between *x and* y; in part (d) we will determine the equation of a straight line that "best" represents the relationship according to the least squares criterion.

d. $\bar{x} = \dfrac{\Sigma x_i}{n} = \dfrac{15}{5} = 3 \quad \bar{y} = \dfrac{\Sigma y_i}{n} = \dfrac{40}{5} = 8$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 26 \quad \Sigma(x_i - \bar{x})^2 = 10$

$b_1 = \dfrac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \dfrac{26}{10} = 2.6$

$b_0 = \bar{y} - b_1\bar{x} = 8 - (2.6)(3) = 0.2$

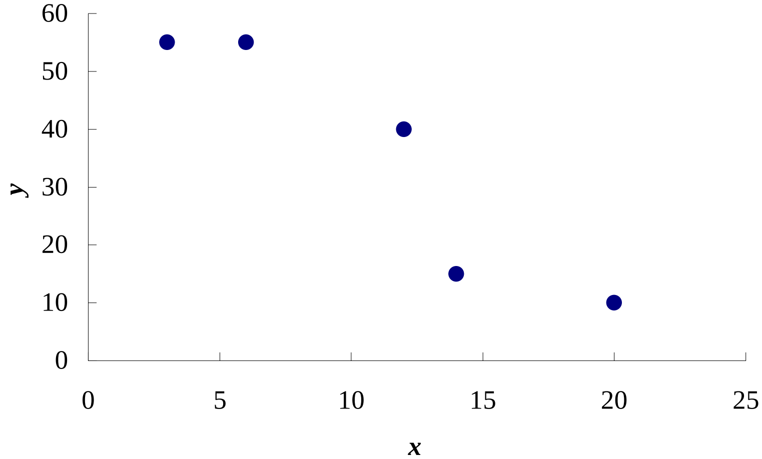
$\hat{y} = 0.2 + 2.6x$

e. $\hat{y} = 0.2 + 2.6(4) = 10.6$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

2. a.



b. There appears to be a negative linear relationship between *x* and *y*.

c. Many different straight lines can be drawn to provide a linear approximation of the relationship between *x and y*; in part (d) we will determine the equation of a straight line that "best" represents the relationship according to the least squares criterion.

d. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{55}{5} = 11 \quad \bar{y} = \frac{\Sigma y_i}{n} = \frac{175}{5} = 35$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = -540 \quad \Sigma(x_i - \bar{x})^2 = 180$

$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{-540}{180} = -3$

$b_0 = \bar{y} - b_1\bar{x} = 35 - (-3)(11) = 68$

$\hat{y} = 68 - 3x$

e. $\hat{y} = 68 - 3(10) = 38$

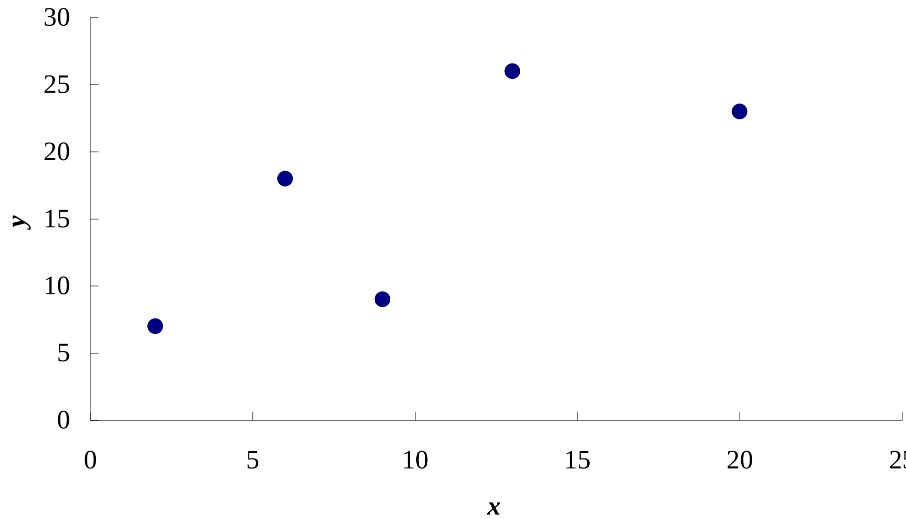
This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

SOUTH-WESTERN
CENGAGE Learning

Chapter 12

3. a.



b. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{50}{5} = 10 \quad \bar{y} = \frac{\Sigma y_i}{n} = \frac{83}{5} = 16.6$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 171 \quad \Sigma(x_i - \bar{x})^2 = 190$

$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{171}{190} = 0.9$

$b_0 = \bar{y} - b_1\bar{x} = 16.6 - (0.9)(10) = 7.6$

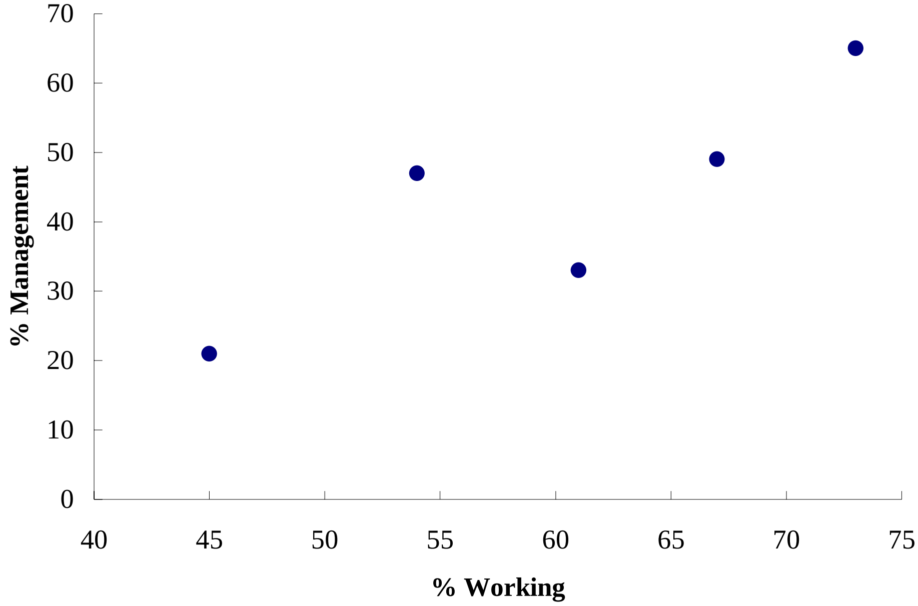
$\hat{y} = 7.6 + 0.9x$

c. $\hat{y} = 7.6 + 0.9(6) = 13$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

4. a.



b. There appears to be a positive linear relationship between the percentage of women working in the five companies (x)the percentage of management jobs held by women in that company (y)

c. Many different straight lines can be drawn to provide a linear approximation of the relationship between x and y; in part (d) we will determine the equation of a straight line that "best" represents the relationship according to the least squares criterion.

d. $\bar{x} = \dfrac{\Sigma x_i}{n} = \dfrac{300}{5} = 60 \qquad \bar{y} = \dfrac{\Sigma y_i}{n} = \dfrac{215}{5} = 43$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 624 \qquad \Sigma(x_i - \bar{x})^2 = 480$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{624}{480} = 1.3$$

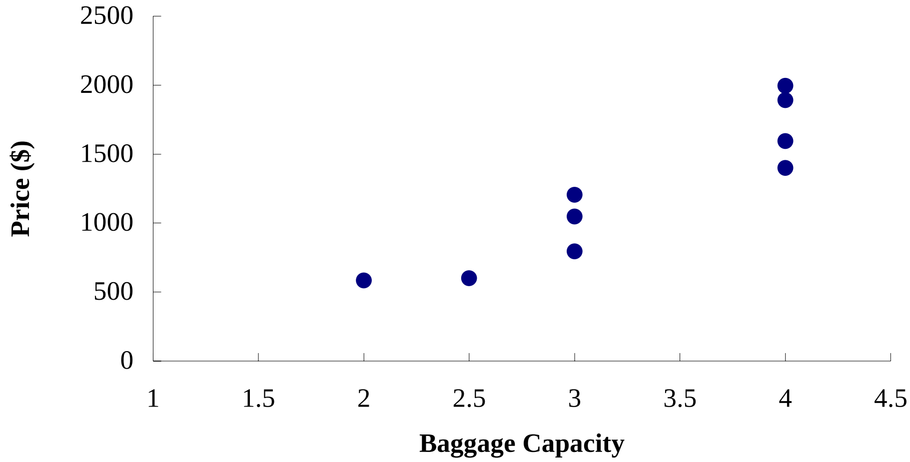
$\hat{y} = -35 + 1.3x$

e. $\hat{y} = -35 + 1.3x = -35 + 1.3(60) = 43\%$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

5. a.



b. Let x = baggage capacity and y = price ($).

There appears to be a positive linear relationship between x and y.

c. Many different straight lines can be drawn to provide a linear approximation of the relationship between x and y; in part (d) we will determine the equation of a straight line that "best" represents the relationship according to the least squares criterion.

d. $\bar{x} = \dfrac{\Sigma x_i}{n} = \dfrac{29.5}{9} = 3.277778$ $\quad \bar{y} = \dfrac{\Sigma y_i}{n} = \dfrac{11{,}110}{9} = 123.444444$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 2909.888891 \quad \Sigma(x_i - \bar{x})^2 = 4.555559$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{2909.888891}{4.555559} = 638.755615$$

$$b_0 = \bar{y} - b_1\bar{x} = 1234.4444 - (638.7561)(3.2778) = -859.254658$$

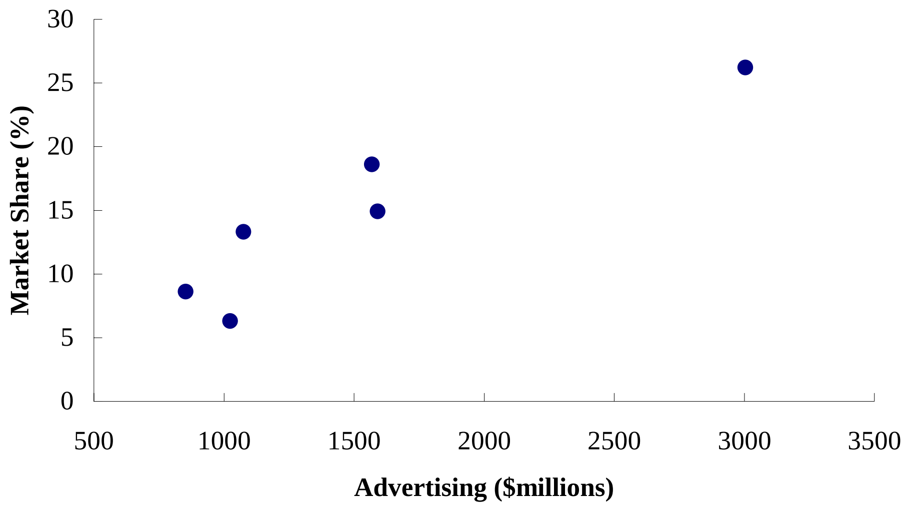
$$\hat{y} = -859.26 + 638.76x$$

e. A one point increase in the baggage capacity rating will increase the price by approximately $639.

f. $\hat{y} = -859.26 + 638.76x = -859.26 + 638.76(3) = \$1057$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.
STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

6. a.



b. There appears to be a positive linear relationship between advertising expenditure and market share.

c. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{9114}{6} = 1519 \quad \bar{y} = \frac{\Sigma y_i}{n} = \frac{87.9}{6} = 14.65$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 26{,}127.3 \quad \Sigma(x_i - \bar{x})^2 = 3{,}098{,}044$

$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{26{,}127.3}{3{,}098{,}044} = .00843348$

$b_0 = \bar{y} - b_1\bar{x} = 14.65 - (.00843348)(1519) = 1.8395$

$\hat{y} = 1.8395 + .0084x$

d. A one unit increase in advertising expenditure will increase the market share by .0084. Because advertising expenditure is measure in $million, an increase of $100 million would increase the market share by .84%.

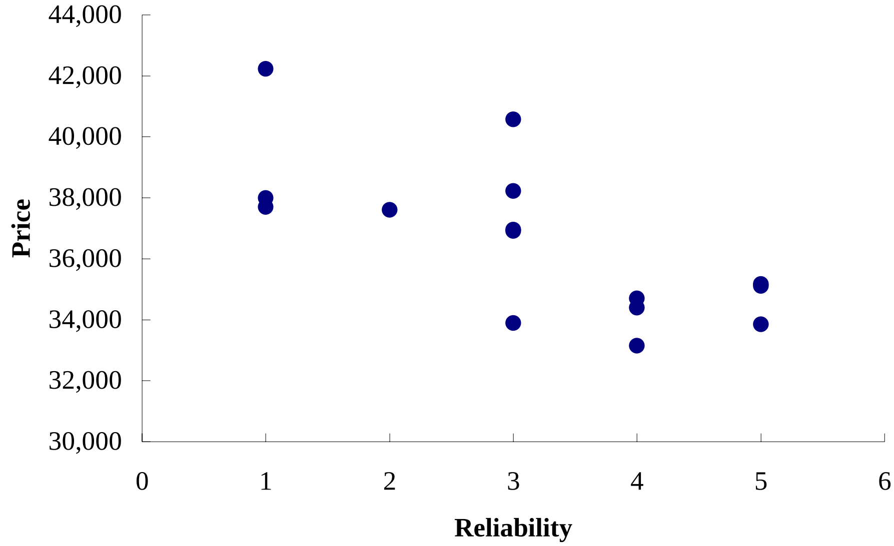
e. $\hat{y} = 1.8395 + .0084x = 1.8395 + .0084(1200) = 11.9$ or 11.9%

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

SOUTH-WESTERN
CENGAGE Learning

7. a.



b. $\bar{x} = \Sigma x_i / n = 47/15 = 3.1333$ $\quad \bar{y} = \Sigma y_i / n = 548,434/15 = 36,562.2667$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = -36,086.5333 \quad \Sigma(x_i - \bar{x})^2 = 27.7333$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{-36,086.5333}{27.7333} = -1301.1987$$

$$b_0 = \bar{y} - b_1\bar{x} = 36,562.2667 - (-1301.1987)(3.1333) = 40,639.3126$$

$$\hat{y} = 40,639 - 1301.2x$$

c. The scatter diagram and the slope of the estimated regression equation indicate a negative linear relationship between reliability and price. Thus, it appears that higher reliable cars actually cost less. Although this result may surprise you, it may be due to the fact that higher priced cars have more options that may increase the likelihood of problems.

d. A car with a good reliability rating corresponds to $x = 3$.

$$\hat{y} = 40,639 - 1301.2x = 40,639 - 1301.2(3) = 36,735.40$$

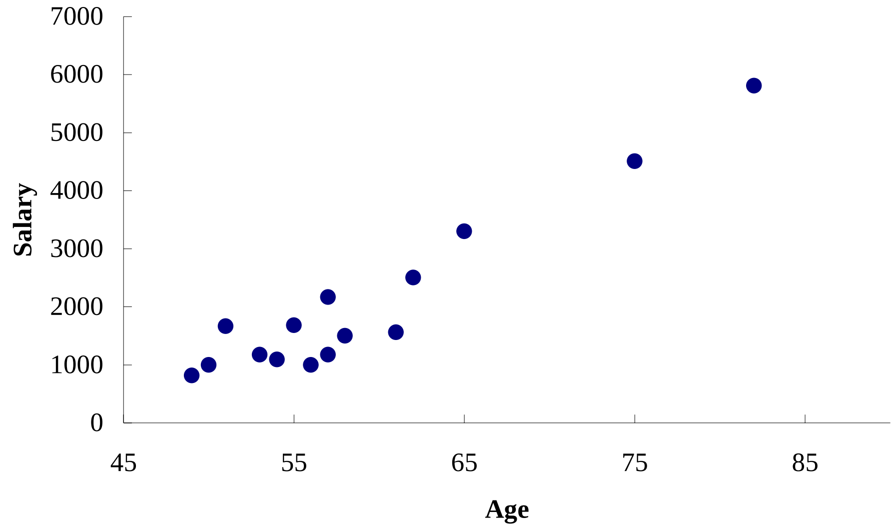
Thus, the estimate of the price of an upscale sedan with a good reliability rating is approximately $36,735.

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

SOUTH-WESTERN
CENGAGE Learning

8. a.



b. There appears to be a positive linear relationship between age and salary.

c. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{885}{15} = 59 \quad \bar{y} = \frac{\Sigma y_i}{n} = \frac{30,939}{15} = 2062.6$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 175,265 \quad \Sigma(x_i - \bar{x})^2 = 1174$

$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{175,265}{1174} = 149.2888$

$b_0 = \bar{y} - b_1\bar{x} = 2062.6 - (149.2888)(59) = -6745.44$

$\hat{y} = -6745.44 + 149.29x$

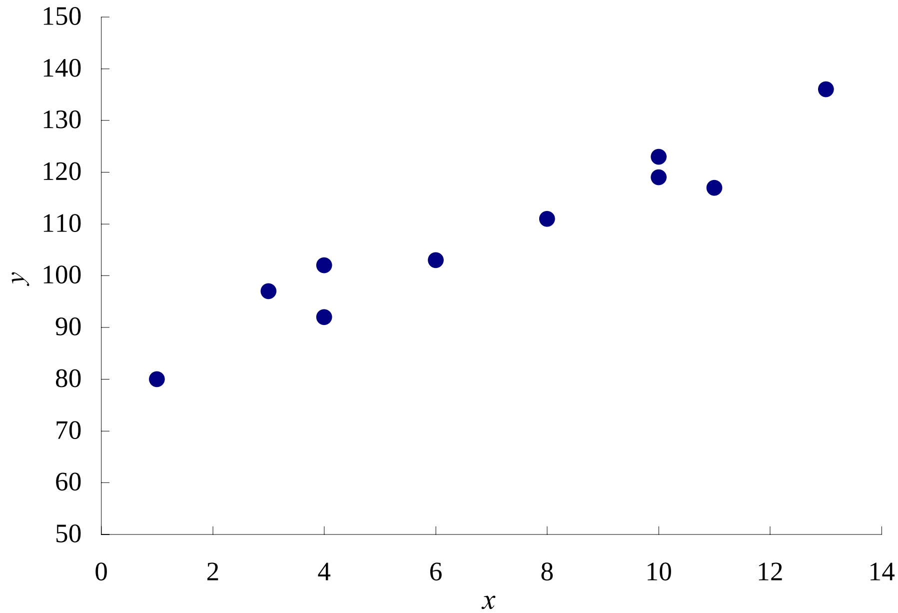
e. $\hat{y} = -6745.44 + 149.29x = -6745.44 + 149.29(72) = \$4003$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

SOUTH-WESTERN
CENGAGE Learning

9. a.



b. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{70}{10} = 7 \quad \bar{y} = \frac{\Sigma y_i}{n} = \frac{1080}{10} = 108$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 568 \quad \Sigma(x_i - \bar{x})^2 = 142$

$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{568}{142} = 4$

$b_0 = \bar{y} - b_1\bar{x} = 108 - (4)(7) = 80$

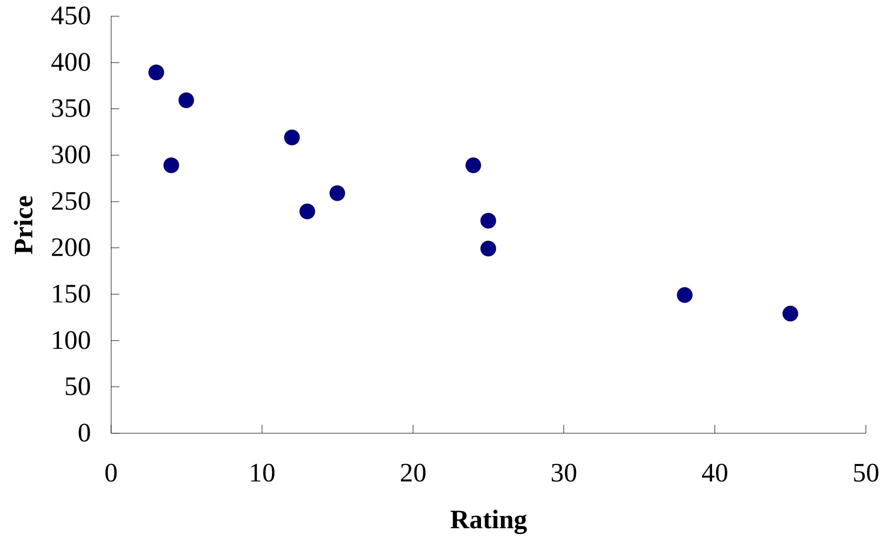
$\hat{y} = 80 + 4x$

c. $\hat{y} = 80 + 4x = 80 + 4(9) = 116$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

10. a.



b. The scatter diagram and the slope of the estimated regression equation indicate a negative linear relationship between rating and price. Thus, it appears that sleeping bags with a lower temperature rating cost more than sleeping bags with a higher temperature rating. In other words, it costs more to stay warmer.

c. $\bar{x} = \Sigma x_i / n = 209/11 = 19 \quad \bar{y} = \Sigma y_i / n = 2849/11 = 259$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = -10,090 \quad \Sigma(x_i - \bar{x})^2 = 1912$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{-10,090}{1912} = -5.2772$$

$b_0 = \bar{y} - b_1\bar{x} = 259 - (-5.2772)(19) = 359.2668$

$\hat{y} = 359.2668 - 5.2772x$

d. $\hat{y} = 359.2668 - 5.2772x = 359.2668 - 5.2772(20) = 253.72$

Thus, the estimate of the price of sleeping bag with a temperature rating of 20 is approximately \$254.

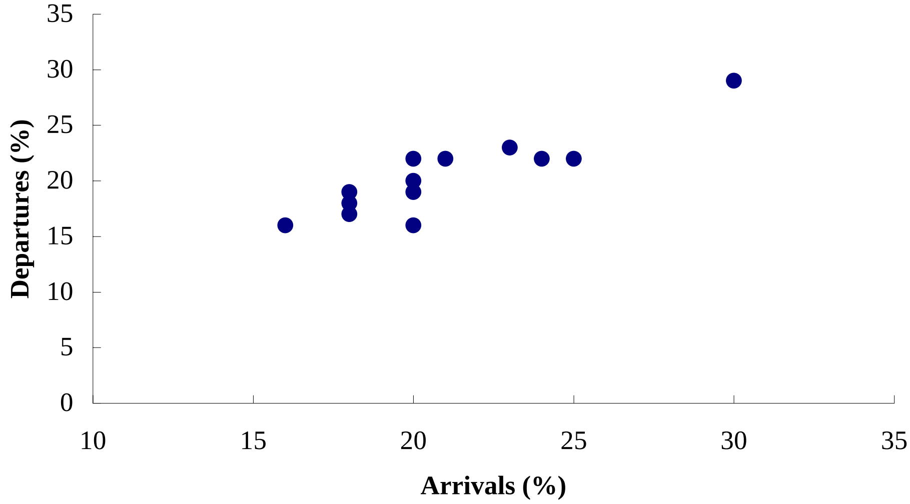
This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

SOUTH-WESTERN
CENGAGE Learning

Chapter 12

11. a.



b. There appears to be a positive linear relationship between the variables.

c. Let x = percentage of late arrivals and y = percentage of late departures.

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{273}{13} = 21 \quad \bar{y} = \frac{\Sigma y_i}{n} = \frac{265}{13} = 20.3846$$

$$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 142 \quad \Sigma(x_i - \bar{x})^2 = 166$$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{142}{166} = .8554$$

$$b_0 = \bar{y} - b_1\bar{x} = 20.3846 - (.8554)(21) = 2.4212$$

$$\hat{y} = 2.42 + .86x$$

d. A one percent increase in the percentage of late arrivals will increase the percentage of late arrivals by .86 or slightly less than one percent.

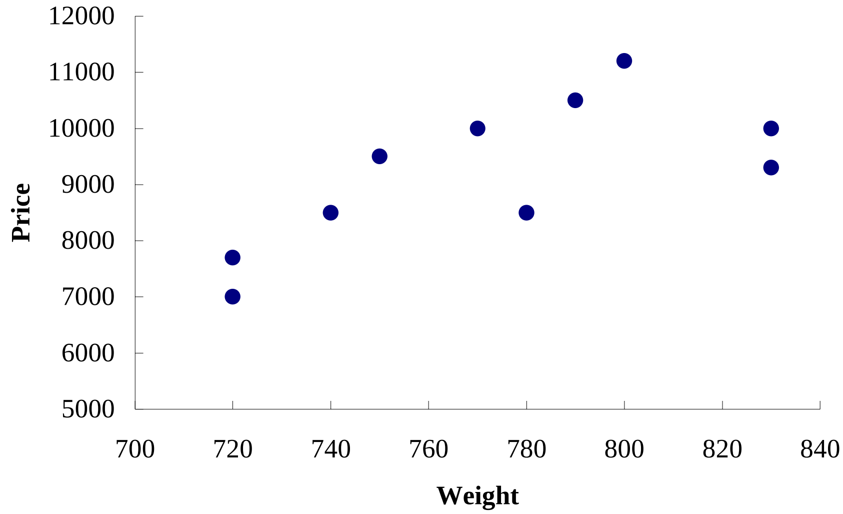
e. $\hat{y} = 2.42 + .86x = 2.42 + .86(22) = 21.34\%$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

SOUTH-WESTERN
CENGAGE Learning

12. a.



b. The scatter diagram indicates a positive linear relationship between weight and price. Thus, it appears that PWC's that weigh more have a higher price.

c. $\bar{x} = \Sigma x_i / n = 7730/10 = 773 \quad \bar{y} = \Sigma y_i / n = 92,200/10 = 9220$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 332,400 \quad \Sigma(x_i - \bar{x})^2 = 14,810$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{332,400}{14,810} = 22.4443$$

$b_0 = \bar{y} - b_1\bar{x} = 9220 - (22.4443)(773) = -8129.4439$

$\hat{y} = -8129.4439 + 22.4443x$

d. $\hat{y} = -8129.4439 + 22.4443x = -8129.4439 + 22.4443(750) = 8703.78$

Thus, the estimate of the price of Jet Ski with a weight of 750 pounds is approximately $8704.

e. No. The relationship between weight and price is not deterministic.

f. The weight of the Kawasaki SX-R 800 is so far below the lowest weight for the data used to develop the estimated regression equation that we would not recommend using the estimated regression equation to predict the price for this model.

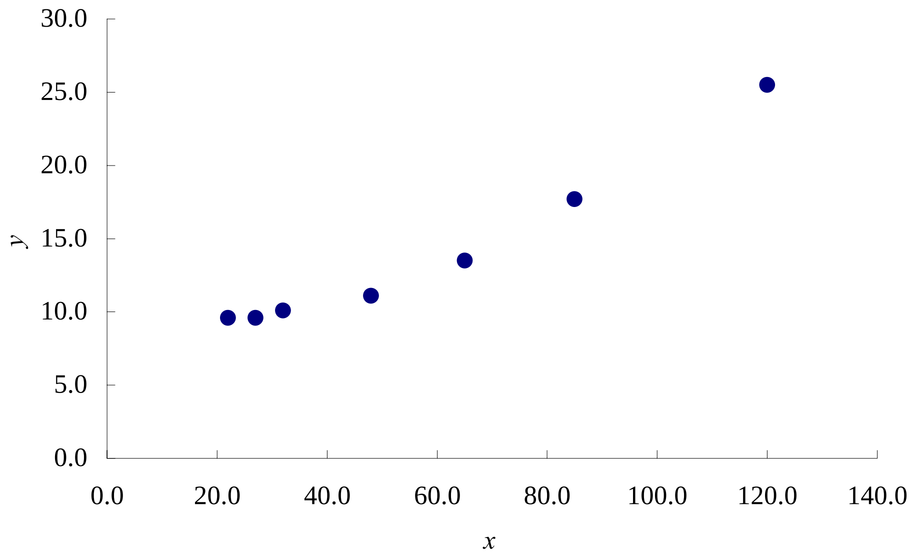
This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

SOUTH-WESTERN
CENGAGE Learning

Chapter 12

13. a.



b. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{399}{7} = 57 \quad \bar{y} = \frac{\Sigma y_i}{n} = \frac{97.1}{7} = 13.8714$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 1233.7 \quad \Sigma(x_i - \bar{x})^2 = 7648$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{1233.7}{7648} = 0.1613$$

$b_0 = \bar{y} - b_1\bar{x} = 13.8714 - (0.1613)(57) = 4.6773$

$\hat{y} = 4.68 + 0.16x$

c. $\hat{y} = 4.68 + 0.16x = 4.68 + 0.16(52.5) = 13.08$ or approximately \$13,080.

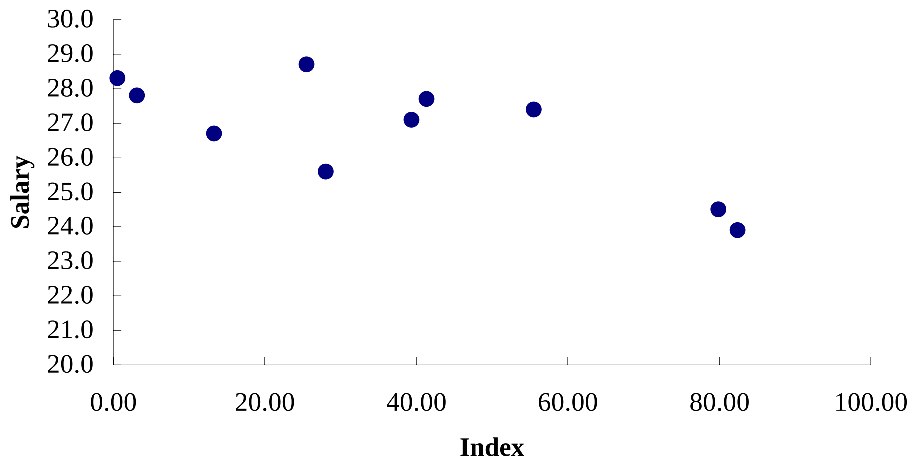
The agent's request for an audit appears to be justified.

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

SOUTH-WESTERN
CENGAGE Learning

14. a.



b. Let x = cost of living index and y = starting salary ($1000s)

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{369.16}{10} = 36.916 \qquad \bar{y} = \frac{\Sigma y_i}{n} = \frac{267.7}{10} = 26.77$$

$$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = -311.9592 \qquad \Sigma(x_i - \bar{x})^2 = 7520.4042$$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{-311.9592}{7520.4042} = -.0415$$

$$b_0 = \bar{y} - b_1\bar{x} = 26.77 - (-.0415)(36.916) = 28.30$$

$$\hat{y} = 28.30 - .0415x$$

c. $\hat{y} = 28.30 - .0415x = 28.30 - .0415(50) = 26.2$

15. a. The estimated regression equation and the mean for the dependent variable are:

$$\hat{y}_i = 0.2 + 2.6x_i \qquad \bar{y} = 8$$

The sum of squares due to error and the total sum of squares are

$$\text{SSE} = \sum(y_i - \hat{y}_i)^2 = 12.40 \qquad \text{SST} = \sum(y_i - \bar{y})^2 = 80$$

Thus, SSR = SST - SSE = 80 - 12.4 = 67.6

b. r^2 = SSR/SST = 67.6/80 = .845

The least squares line provided a very good fit; 84.5% of the variability in y has been explained by the least squares line.

c. $r_{xy} = \sqrt{.845} = +.9192$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

16. a. The estimated regression equation and the mean for the dependent variable are:

$\hat{y}_i = 68 - 3x \qquad \bar{y} = 35$

The sum of squares due to error and the total sum of squares are

$\text{SSE} = \sum(y_i - \hat{y}_i)^2 = 230 \qquad \text{SST} = \sum(y_i - \bar{y})^2 = 1850$

Thus, SSR = SST - SSE = 1850 - 230 = 1620

b. r^2 = SSR/SST = 1620/1850 = .876

The least squares line provided an excellent fit; 87.6% of the variability in *y* has been explained by the estimated regression equation.

c. $r_{xy} = \sqrt{.876} = -.936$

Note: the sign for *r* is negative because the slope of the estimated regression equation is negative. (b_1 = -3)

17. The estimated regression equation and the mean for the dependent variable are:

$\hat{y}_i = 7.6 + .9x \qquad \bar{y} = 16.6$

The sum of squares due to error and the total sum of squares are

$\text{SSE} = \sum(y_i - \hat{y}_i)^2 = 127.3 \qquad \text{SST} = \sum(y_i - \bar{y})^2 = 281.2$

Thus, SSR = SST - SSE = 281.2 – 127.3 = 153.9

r^2 = SSR/SST = 153.9/281.2 = .547

We see that 54.7% of the variability in *y* has been explained by the least squares line.

$r_{xy} = \sqrt{.547} = +.740$

18. a. The estimated regression equation and the mean for the dependent variable are:

$\hat{y} = 1790.5 + 581.1x \qquad \bar{y} = 3650$

The sum of squares due to error and the total sum of squares are

$\text{SSE} = \sum(y_i - \hat{y}_i)^2 = 85{,}135.14 \qquad \text{SST} = \sum(y_i - \bar{y})^2 = 335{,}000$

Thus, SSR = SST - SSE = 335,000 - 85,135.14 = 249,864.86

b. r^2 = SSR/SST = 249,864.86/335,000 = .746

We see that 74.6% of the variability in *y* has been explained by the least squares line.

c. $r_{xy} = \sqrt{.746} = +.8637$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

19. The estimated regression equation and the mean for the dependent variable are:

$\hat{y}_i = 40,639 - 1301.2x \qquad \bar{y} = 36,562.27$

The sum of squares due to error and the total sum of squares are

$\text{SSE} = \Sigma(y_i - \hat{y}_i)^2 = 47,116,828 \qquad \text{SST} = \Sigma(y_i - \bar{y})^2 = 94,072,519$

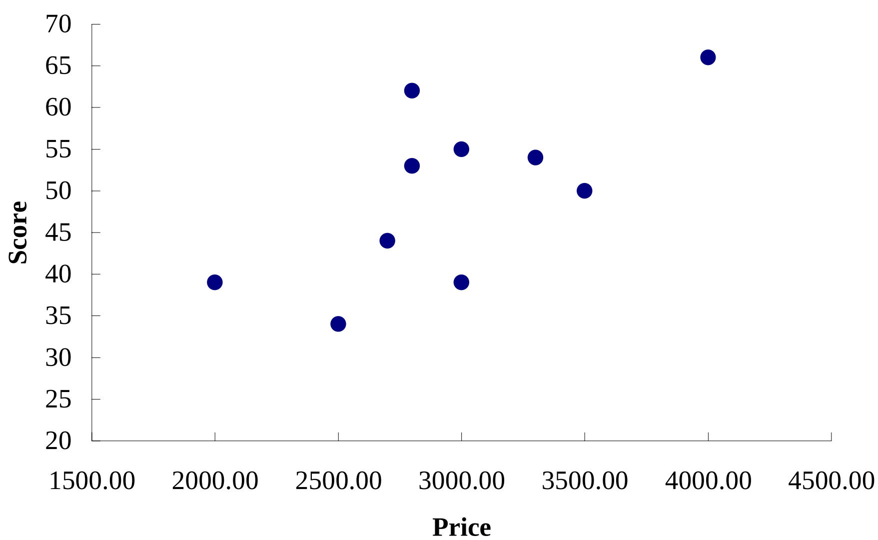
Thus, SSR = SST - SSE = 94,072,519 – 47,116,828 = 46,955,691

r^2 = SSR/SST = 46,955,691/94,072,519 = .4991

We see that 49.91% of the variability in y has been explained by the least squares line.

$r_{xy} = \sqrt{.4991} = -.71$

20. a.



The scatter diagram indicates a positive linear relationship between price and score.

$\bar{x} = \Sigma x_i / n = 29,600/10 = 2960 \qquad \bar{y} = \Sigma y_i / n = 496/10 = 49.6$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 34,840 \qquad \Sigma(x_i - \bar{x})^2 = 2,744,000$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{34,840}{2,744,000} = .012697$$

$b_0 = \bar{y} - b_1\bar{x} = 49.6 - (.012697)(2960) = 12.0169$

$\hat{y} = 12.0169 + .0127x$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.
STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

b. The sum of squares due to error and the total sum of squares are

$$\text{SSE} = \sum(y_i - \hat{y}_i)^2 = 540.0446 \qquad \text{SST} = \sum(y_i - \bar{y})^2 = 982.4$$

Thus, SSR = SST - SSE = 982.4 – 540.0446 = 442.3554

r^2 = SSR/SST = 442.3554/982.4 = .4503

The fit provided by the estimated regression equation is not that good; only 45.03% of the variability in y has been explained by the least squares line.

c. $\hat{y} = 12.0169 + .0127x = 12.0169 + .0127(3200) = 52.66$

The estimate of the overall score for a 42-inch plasma television is approximately 53.

21. a. $\bar{x} = \dfrac{\Sigma x_i}{n} = \dfrac{3450}{6} = 575 \qquad \bar{y} = \dfrac{\Sigma y_i}{n} = \dfrac{33,700}{6} = 5616.67$

$$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 712,500 \qquad \Sigma(x_i - \bar{x})^2 = 93,750$$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{712,500}{93,750} = 7.6$$

$$b_0 = \bar{y} - b_1\bar{x} = 5616.67 - (7.6)(575) = 1246.67$$

$$\hat{y} = 1246.67 + 7.6x$$

b. $7.60

c. The sum of squares due to error and the total sum of squares are:

$$\text{SSE} = \sum(y_i - \hat{y}_i)^2 = 233,333.33 \qquad \text{SST} = \sum(y_i - \bar{y})^2 = 5,648,333.33$$

Thus, SSR = SST - SSE = 5,648,333.33 - 233,333.33 = 5,415,000

r^2 = SSR/SST = 5,415,000/5,648,333.33 = .9587

We see that 95.87% of the variability in y has been explained by the estimated regression equation.

d. $\hat{y} = 1246.67 + 7.6x = 1246.67 + 7.6(500) = \$5046.67$

22. a. Let x = speed (ppm) and y = price ($)

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{149.9}{10} = 14.99 \qquad \bar{y} = \frac{\Sigma y_i}{n} = \frac{10,221}{10} = 1022.1$$

$$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 34,359.810 \qquad \Sigma(x_i - \bar{x})^2 = 291.3890$$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{34,359.810}{291.3890} = 117.917320$$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

$b_0 = \overline{y} - b_1\overline{x} = 1022.1 - (117.917320)(14.99) = -745.480627$

$\hat{y} = -745.480627 + 117.917320x$

b. The sum of squares due to error and the total sum of squares are:

$\text{SSE} = \sum(y_i - \hat{y}_i)^2 = 1{,}678{,}294 \qquad \text{SST} = \sum(y_i - \overline{y})^2 = 5{,}729{,}911$

Thus, SSR = SST - SSE = 5,729,911 - 1,678,294 = 4,051,617

r^2 = SSR/SST = 4,051,617/5,729,911 = 0.7071

Approximately 71% of the variability in price is explained by the speed.

c. $r_{xy} = \sqrt{.7071} = +.84$

It reflects a linear relationship that is between weak and strong.

23. a. s^2 = MSE = SSE / (n - 2) = 12.4 / 3 = 4.133

b. $s = \sqrt{\text{MSE}} = \sqrt{4.133} = 2.033$

c. $\Sigma(x_i - \overline{x})^2 = 10$

$$s_{b_1} = \frac{s}{\sqrt{\Sigma(x_i - \overline{x})^2}} = \frac{2.033}{\sqrt{10}} = 0.643$$

d. $t = \dfrac{b_1}{s_{b_1}} = \dfrac{2.6}{.643} = 4.044$

Using t table (3 degrees of freedom), area in tail is between .01 and .025

p-value is between .02 and .05

Using Excel or Minitab, the p-value corresponding to t = 4.04 is .0272.

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$

e. MSR = SSR / 1 = 67.6

F = MSR / MSE = 67.6 / 4.133 = 16.36

Using F table (1 degree of freedom numerator and 3 denominator), p-value is between .025 and .05

Using Excel or Minitab, the p-value corresponding to F = 16.36 is .0272.

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Regression	67.6	1	67.6	16.36	.0272
Error	12.4	3	4.133		
Total	80.0	4			

24. a. s^2 = MSE = SSE/(n - 2) = 230/3 = 76.6667

b. $s = \sqrt{\text{MSE}} = \sqrt{76.6667} = 8.7560$

c. $\Sigma(x_i - \bar{x})^2 = 180$

$$s_{b_1} = \frac{s}{\sqrt{\Sigma(x_i - \bar{x})^2}} = \frac{8.7560}{\sqrt{180}} = 0.6526$$

d. $t = \dfrac{b_1}{s_{b_1}} = \dfrac{-3}{.653} = -4.59$

Using t table (3 degrees of freedom), area in tail is between .005 and .01; p-value is between .01 and .02

Using Excel or Minitab, the p-value corresponding to t = -4.59 is .0193.

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$

e. MSR = SSR/1 = 1620

F = MSR/MSE = 1620/76.6667 = 21.13

Using F table (1 degree of freedom numerator and 3 denominator), p-value is between .01 and .025

Using Excel or Minitab, the p-value corresponding to F = 21.13 is .0193.

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Regression	230	1	230	21.13	.0193
Error	1620	3	76.6667		
Total	1850	4			

25. a. s^2 = MSE = SSE/(n - 2) = 127.3/3 = 42.4333

$s = \sqrt{\text{MSE}} = \sqrt{42.4333} = 6.5141$

b. $\Sigma(x_i - \bar{x})^2 = 190$

$$s_{b_1} = \frac{s}{\sqrt{\Sigma(x_i - \bar{x})^2}} = \frac{6.5141}{\sqrt{190}} = 0.4726$$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

$$t = \frac{b_1}{s_{b_1}} = \frac{.9}{.4726} = 1.90$$

Using t table (3 degrees of freedom), area in tail is between .05 and .10

p-value is between .10 and .20

Using Excel or Minitab, the p-value corresponding to t = 1.90 is .1530.

Because $p\text{-value} > \alpha$, we cannot reject H_0: $\beta_1 = 0$; x and y do not appear to be related.

c. MSR = SSR/1 = 153.9 /1 = 153.9

F = MSR/MSE = 153.9/42.4333 = 3.63

Using F table (1 degree of freedom numerator and 3 denominator), p-value is greater than .10

Using Excel or Minitab, the p-value corresponding to F = 3.63 is .1530.

Because $p\text{-value} > \alpha$, we cannot reject H_0: $\beta_1 = 0$; x and y do not appear to be related.

26. a. In solving exercise 18, we found SSE = 85,135.14

s^2 = MSE = SSE/(n - 2) = 85,135.14/4 = 21,283.79

$$s = \sqrt{\text{MSE}} = \sqrt{21{,}283.79} = 145.89$$

$$\Sigma(x_i - \bar{x})^2 = 0.74$$

$$s_{b_1} = \frac{s}{\sqrt{\Sigma(x_i - \bar{x})^2}} = \frac{145.89}{\sqrt{0.74}} = 169.59$$

$$t = \frac{b_1}{s_{b_1}} = \frac{581.1}{169.59} = 3.43$$

Using t table (4 degrees of freedom), area in tail is between .01 and .025

p-value is between .02 and .05

Using Excel or Minitab, the p-value corresponding to t = 3.43 is .0266.

Because $p\text{-value} \le \alpha$, we reject H_0: $\beta_1 = 0$

b. MSR = SSR/1 = 249,864.86/1 = 249.864.86

F = MSR/MSE = 249,864.86/21,283.79 = 11.74

Using F table (1 degree of freedom numerator and 4 denominator), p-value is between .025 and .05

Using Excel or Minitab, the p-value corresponding to F = 11.74 is .0266.

Because $p\text{-value} \le \alpha$, we reject H_0: $\beta_1 = 0$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

c.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Regression	249864.86	1	249864.86	11.74	.0266
Error	85135.14	4	21283.79		
Total	335000	5			

27. a. $\bar{x} = \dfrac{\Sigma x_i}{n} = \dfrac{37}{10} = 3.7 \quad \bar{y} = \dfrac{\Sigma y_i}{n} = \dfrac{1654}{10} = 165.4$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 315.2 \quad \Sigma(x_i - \bar{x})^2 = 10.1$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{315.2}{10.1} = 31.2079$$

$b_0 = \bar{y} - b_1\bar{x} = 165.4 - (31.2079)(3.7) = 49.9308$

$\hat{y} = 49.9308 + 31.2079x$

b. SSE = $\Sigma(y_i - \hat{y}_i)^2 = 2487.66$ SST = $\Sigma(y_i - \bar{y})^2 = 12{,}324.4$

Thus, SSR = SST - SSE = 12,324.4 - 2487.66 = 9836.74

MSR = SSR/1 = 9836.74

MSE = SSE/(n - 2) = 2487.66/8 = 310.96

F = MSR / MSE = 9836.74/310.96 = 31.63

Using F table (1 degree of freedom numerator and 8 denominator), p-value is less than .01

Using Excel or Minitab, the p-value corresponding to F = 31.63 is .001.

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$

Upper support and price are related.

c. r^2 = SSR/SST = 9,836.74/12,324.4 = .80

The estimated regression equation provided a good fit; we should feel comfortable using the estimated regression equation to estimate the price given the upper support rating.

d. $\hat{y}$ = 49.93 + 31.21(4) = 174.77

28. The sum of squares due to error and the total sum of squares are

SSE = $\sum(y_i - \hat{y}_i)^2 = 12{,}953.09$ SST = $\sum(y_i - \bar{y})^2 = 66{,}200$

Thus, SSR = SST - SSE = 66,200 – 12,953.09 = 53,246.91

s^2 = MSE = SSE / (n - 2) = 12,953.09 / 9 = 1439.2322

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

$s = \sqrt{\text{MSE}} = \sqrt{1439.2322} = 37.9372$

We can use either the t test or F test to determine whether temperature rating and price are related.

We will first illustrate the use of the t test.

Note: from the solution to exercise 10 $\Sigma(x_i - \bar{x})^2 = 1912$

$$s_{b_1} = \frac{s}{\sqrt{\Sigma(x_i - \bar{x})^2}} = \frac{37.9372}{\sqrt{1912}} = .8676$$

$$t = \frac{b_1}{s_{b_1}} = \frac{-5.2772}{.8676} = -6.0825$$

Using t table (9 degrees of freedom), area in tail is less than .005; p-value is less than .01

Using Excel or Minitab, the p-value corresponding to t = -6.0825 is .000.

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$

Because we can reject H_0: $\beta_1 = 0$ we conclude that temperature rating and price are related.

Next we illustrate the use of the F test.

MSR = SSR / 1 = 53,246.91

F = MSR / MSE = 53,246.91 / 1439.2322 = 37.00

Using F table (1 degree of freedom numerator and 9 denominator), p-value is less than .01

Using Excel or Minitab, the p-value corresponding to F = 37.00 is .000.

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$

Because we can reject H_0: $\beta_1 = 0$ we conclude that temperature rating and price are related.

The ANOVA table is shown below.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Regression	53,246.91	1	53,246.91	37.00	.000
Error	12,953.09	9	1439.2322		
Total	66,200	10			

29. $\text{SSE} = \Sigma(y_i - \hat{y}_i)^2 = 233{,}333.33$ $\text{SST} = \Sigma(y_i - \bar{y})^2 = 5{,}648{,}333.33$

Thus, SSR = SST – SSE= 5,648,333.33 –233,333.33 = 5,415,000

MSE = SSE/(n - 2) = 233,333.33/(6 - 2) = 58,333.33

MSR = SSR/1 = 5,415,000

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

F = MSR / MSE = 5,415,000 / 58,333.25 = 92.83

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Regression	5,415,000.00	1	5,415,000	92.83	.0006
Error	233,333.33	4	58,333.33		
Total	5,648,333.33	5			

Using F table (1 degree of freedom numerator and 4 denominator), p-value is less than .01

Using Excel or Minitab, the p-value corresponding to F = 92.83 is .0006.

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$. Production volume and total cost are related.

30. $\text{SSE} = \Sigma(y_i - \hat{y}_i)^2 = 1{,}678{,}294$ $\quad \text{SST} = \Sigma(y_i - \overline{y})^2 = 5{,}729{,}911$

Thus, SSR = SST – SSE = 5,729,911 – 1,678,294 = 4,051,617

s^2 = MSE = SSE/(n-2) = 1,678,294/8 = 209,786.8

$$s = \sqrt{209{,}786.8} = 458.02$$

$$\sum(x_i - \overline{x})^2 = 291.3890$$

$$s_{b_1} = \frac{s}{\sqrt{\Sigma(x_i - \overline{x})^2}} = \frac{458.02}{\sqrt{291.3890}} = 26.83$$

$$t = \frac{b_1}{s_{b_1}} = \frac{117.9173}{26.83} = 4.39$$

Using t table (1 degree of freedom numerator and 8 denominator), area in tail is less than .005

p-value is less than .01

Using Excel or Minitab, the p-value corresponding to t = 4.39 is .002.

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$

There is a significant relationship between x and y.

31. SSE = 540.04 and SST = 982.40

Thus, SSR = SST - SSE = 982.40 – 540.04 = 442.36

s^2 = MSE = SSE / (n - 2) = 540.04 / 8 = 67.5050

$$s = \sqrt{\text{MSE}} = \sqrt{67.5050} = 8.216$$

MSR = SSR / 1 = 442.36

F = MSR / MSE = 442.36 / 67.5050 = 6.55

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

Using F table (1 degree of freedom numerator and 8 denominator), p-value is between .025 and .05

Using Excel or Minitab, the p-value corresponding to $F = 6.55$ is .034.

Because p-value $\leq \alpha$, we reject H_0: $\beta_1 = 0$

Conclusion: price and overall score are related

32. a. $s = 2.033$

$\bar{x} = 3 \quad \Sigma(x_i - \bar{x})^2 = 10$

$$s_{\hat{y}_p} = s\sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}} = 2.033\sqrt{\frac{1}{5} + \frac{(4-3)^2}{10}} = 1.11$$

b. $\hat{y} = 0.2 + 2.6x = 0.2 + 2.6(4) = 10.6$

$\hat{y}_p \pm t_{\alpha/2} s_{\hat{y}_p}$

$10.6 \pm 3.182\ (1.11) = 10.6 \pm 3.53$

or 7.07 to 14.13

c. $$s_{\text{ind}} = s\sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}} = 2.033\sqrt{1 + \frac{1}{5} + \frac{(4-3)^2}{10}} = 2.32$$

d. $\hat{y}_p \pm t_{\alpha/2} s_{\text{ind}}$

$10.6 \pm 3.182\ (2.32) = 10.6 \pm 7.38$

or 3.22 to 17.98

33. a. $s = 8.7560$

b. $\bar{x} = 11 \quad \Sigma(x_i - \bar{x})^2 = 180$

$$s_{\hat{y}_p} = s\sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}} = 8.7560\sqrt{\frac{1}{5} + \frac{(8-11)^2}{180}} = 4.3780$$

$\hat{y} = 68 - 3x = 68 - 3(8) = 44$

$\hat{y}_p \pm t_{\alpha/2} s_{\hat{y}_p}$

$44 \pm 3.182\ (4.3780) = 44 \pm 13.93$

or 30.07 to 57.93

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

c. $s_{\text{ind}} = s\sqrt{1+\dfrac{1}{n}+\dfrac{(x_{\text{p}}-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}} = 8.7560\sqrt{1+\dfrac{1}{5}+\dfrac{(8-11)^2}{180}} = 9.7895$

d. $\hat{y}_{\text{p}} \pm t_{\alpha/2} s_{\text{ind}}$

44 ± 3.182(9.7895) = 44 ± 31.15

or 12.85 to 75.15

34. $s = 6.5141$

$\bar{x} = 10 \quad \Sigma(x_i-\bar{x})^2 = 190$

$$s_{\hat{y}_{\text{p}}} = s\sqrt{\frac{1}{n}+\frac{(x_{\text{p}}-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}} = 6.5141\sqrt{\frac{1}{5}+\frac{(12-10)^2}{190}} = 3.0627$$

$\hat{y} = 7.6 + .9x = 7.6 + .9(12) = 18.40$

$\hat{y}_{\text{p}} \pm t_{\alpha/2} s_{\hat{y}_{\text{p}}}$

18.40 ± 3.182(3.0627) = 18.40 ± 9.75

or 8.65 to 28.15

$$s_{\text{ind}} = s\sqrt{1+\frac{1}{n}+\frac{(x_{\text{p}}-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}} = 6.5141\sqrt{1+\frac{1}{5}+\frac{(12-10)^2}{190}} = 7.1982$$

$\hat{y}_{\text{p}} \pm t_{\alpha/2} s_{\text{ind}}$

18.40 ± 3.182(7.1982) = 18.40 ± 22.90

or -4.50 to 41.30

35. a. Note: some of the values shown were computed in exercises 18 and 26.

$s = 145.89$

$\bar{x} = 3.2 \quad \Sigma(x_i-\bar{x})^2 = 0.74$

$$s_{\hat{y}_{\text{p}}} = s\sqrt{\frac{1}{n}+\frac{(x_{\text{p}}-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}} = 145.89\sqrt{\frac{1}{6}+\frac{(3-3.2)^2}{0.74}} = 68.54$$

$\hat{y} = 1790.5 + 581.1x = 1790.5 + 581.1(3) = 3533.8$

$\hat{y}_{\text{p}} \pm t_{\alpha/2} s_{\hat{y}_{\text{p}}}$

3533.8 ± 2.776 (68.54) = 3533.8 ± 190.27
or \$3343.53 to \$3724.07

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

b. $s_{\text{ind}} = s\sqrt{1+\frac{1}{n}+\frac{(x_{\text{p}}-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}} = 145.89\sqrt{1+\frac{1}{6}+\frac{(3-3.2)^2}{0.74}} = 161.19$

$\hat{y}_{\text{p}} \pm t_{\alpha/2} s_{\text{ind}}$

3533.8 ± 2.776 (161.19) = 3533.8 ± 447.46

or \$3086.34 to \$3981.26

36. a. $\hat{y} = 359.2668 - 5.2772x$ =359.2668 – 5.2772(30) = 200.95 or approximately \$201.

b. s = 37.9372

$\bar{x} = 19 \quad \Sigma(x_i-\bar{x})^2 = 1912$

$s_{\hat{y}_{\text{p}}} = s\sqrt{\frac{1}{n}+\frac{(x_{\text{p}}-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}} = 37.9372\sqrt{\frac{1}{11}+\frac{(30-19)^2}{1912}} = 14.90$

$\hat{y}_{\text{p}} \pm t_{\alpha/2} s_{\hat{y}_{\text{p}}}$

200.95 ± 2.262 (14.90) = 200.95 ± 33.70

or 167.25 to 234.65

c. $s_{\text{ind}} = s\sqrt{1+\frac{1}{n}+\frac{(x_{\text{p}}-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}} = 37.9372\sqrt{1+\frac{1}{11}+\frac{(30-19)^2}{1912}} = 40.76$

$\hat{y}_{\text{p}} \pm t_{\alpha/2} s_{\text{ind}}$

200.95 ± 2.262 (40.76) = 200.95 ± 92.20

or 108.75 to 293.15

d. As expected, the prediction interval is much wider than the confidence interval. This is due to the fact that it is more difficult to predict the results for one new model with a temperature rating of 30 F^{o} than it is to estimate the mean for all models with a temperature rating of 30 F^{o} .

37. a. $\bar{x} = 57 \quad \Sigma(x_i-\bar{x})^2 = 7648$

$s^2 = 1.88 \quad s = 1.37$

$s_{\hat{y}_{\text{p}}} = s\sqrt{\frac{1}{n}+\frac{(x_{\text{p}}-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}} = 1.37\sqrt{\frac{1}{7}+\frac{(52.5-57)^2}{7648}} = 0.52$

$\hat{y}_{\text{p}} \pm t_{\alpha/2} s_{\hat{y}_{\text{p}}}$

$\hat{y} = 4.68 + 0.16x = 4.68 + 0.16(52.5) = 13.08$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

$13.08 \pm 2.571\,(.52) = 13.08 \pm 1.34$

or 11.74 to 14.42 or $11,740 to $14,420

b. $s_{\text{ind}} = 1.47$

$13.08 \pm 2.571\,(1.47) = 13.08 \pm 3.78$

or 9.30 to 16.86 or $9,300 to $16,860

c. Yes, $20,400 is much larger than anticipated.

d. Any deductions exceeding the $16,860 upper limit could suggest an audit.

38. a. $\hat{y} = 1246.67 + 7.6(500) = \$5046.67$

b. $\bar{x} = 575 \quad \Sigma(x_i - \bar{x})^2 = 93,750$

$s^2 = \text{MSE} = 58,333.33 \quad s = 241.52$

$$s_{\text{ind}} = s\sqrt{1 + \frac{1}{n} + \frac{(x_{\text{p}} - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}} = 241.52\sqrt{1 + \frac{1}{6} + \frac{(500 - 575)^2}{93,750}} = 267.50$$

$\hat{y}_{\text{p}} \pm t_{\alpha/2} s_{\text{ind}}$

$5046.67 \pm 4.604\,(267.50) = 5046.67 \pm 1231.57$

or $3815.10 to $6278.24

c. Based on one month, $6000 is not out of line since $3815.10 to $6278.24 is the prediction interval. However, a sequence of five to seven months with consistently high costs should cause concern.

39. a. Let x = miles of track and y = weekday ridership in thousands.

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{203}{7} = 29 \quad \bar{y} = \frac{\Sigma y_i}{n} = \frac{309}{7} = 44.1429$$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 1471 \quad \Sigma(x_i - \bar{x})^2 = 838$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{1471}{838} = 1.7554$$

$b_0 = \bar{y} - b_1\bar{x} = 44.1429 - (1.7554)(29) = -6.76$

$\hat{y} = -6.76 + 1.755x$

b. SST =3620.9 SSE = 1038.7 SSR = 2582.1

r^2 = SSR/SST = 2582.1/3620.9 = .713

The estimated regression equation explained 71.3% of the variability in y; a good fit.

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

c. s^2 = MSE = 1038.7/5 = 207.7

$s = \sqrt{207.7} = 14.41$

$s_{\hat{y}_p} = s\sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}} = 14.41\sqrt{\frac{1}{7} + \frac{(30-29)^2}{838}} = 5.47$

$\hat{y} = -6.76 + 1.755x = -6.76 + 1.755(30) = 45.9$

$\hat{y}_p \pm t_{\alpha/2} s_{\hat{y}_p}$

45.9 ± 2.571(5.47) = 45.9 ± 14.1

or 31.8 to 60

d. $s_{ind} = s\sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\Sigma(x_i - \bar{x})^2}} = 14.41\sqrt{1 + \frac{1}{7} + \frac{(30-29)^2}{838}} = 15.41$

$\hat{y}_p \pm t_{\alpha/2} s_{ind}$

45.9 ± 2.571(15.41) = 45.9 ± 39.6

or 6.3 to 85.5

The prediction interval is so wide that it would not be of much value in the planning process. A larger data set would be beneficial.

40. a. 9

b. $\hat{y} = 20.0 + 7.21x$

c. 1.3626

d. SSE = SST - SSR = 51,984.1 - 41,587.3 = 10,396.8

MSE = 10,396.8/7 = 1,485.3

F = MSR / MSE = 41,587.3 /1,485.3 = 28.00

Using F table (1 degree of freedom numerator and 7 denominator), p-value is less than .01

Using Excel or Minitab, the p-value corresponding to F = 28.00 is .0011.

Because p-value $\leq \alpha$, we reject H_0: $B_1 = 0$.

e. $\hat{y} = 20.0 + 7.21(50) = 380.5$ or $380,500

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

41. a. $\hat{y} = 6.1092 + .8951x$

b. $t = \dfrac{b_1 - B_1}{s_{b_1}} = \dfrac{.8951 - 0}{.149} = 6.01$

Using the t table (8 degrees of freedom), area in tail is less than .005
p-value is less than .01

Using Excel or Minitab, the p-value corresponding to $t = 6.01$ is .0003.

Because $p\text{-value} \le \alpha$, we reject H_0: $B_1 = 0$

c. $\hat{y} = 6.1092 + .8951(25) = 28.49$ or $28.49 per month

42 a. $\hat{y} = 80.0 + 50.0x$

b. 30

c. F = MSR / MSE = 6828.6/82.1 = 83.17

Using F table (1 degree of freedom numerator and 28 denominator), p-value is less than .01

Using Excel or Minitab, the p-value corresponding to $F = 83.17$ is .000.

Because $p\text{-value} < \alpha = .05$, we reject H_0: $B_1 = 0$.

Branch office sales are related to the salespersons.

d. $\hat{y} = 80 + 50\,(12) = 680$ or $680,000

43. a. The Minitab output is shown below:

```
The regression equation is
Price = 4.98 + 2.94 Weight

Predictor        Coef     SE Coef          T        P
Constant        4.979       3.380       1.47    0.154
Weight         2.9370      0.2934      10.01    0.000

S = 8.457       R-Sq = 80.7%     R-Sq(adj) = 79.9%

Analysis of Variance

Source            DF          SS          MS        F        P
Regression         1      7167.9      7167.9   100.22    0.000
Residual Error    24      1716.6        71.5
Total             25      8884.5

Predicted Values for New Observations

New Obs     Fit     SE Fit         95.0% CI              95.0% PI
     1        34.35       1.66   (   30.93,   37.77)  (   16.56,   52.14)
```

b. The p-value = .000 < α = .05 (t or F); significant relationship

c. $r^2 = .807$. The least squares line provided a very good fit.

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

d. The 95% confidence interval is 30.93 to 37.77.

e. The 95% prediction interval is 16.56 to 52.14.

44. a/b. The scatter diagram shows a linear relationship between the two variables.

c. The Minitab output is shown below:

```
The regression equation is
Rental$ = 37.1 - 0.779 Vacancy%

Predictor        Coef     SE Coef          T        P
Constant       37.066       3.530      10.50    0.000
Vacancy%      -0.7791      0.2226      -3.50    0.003

S = 4.889       R-Sq = 43.4%     R-Sq(adj) = 39.8%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression         1      292.89      292.89     12.26    0.003
Residual Error    16      382.37       23.90
Total             17      675.26

Predicted Values for New Observations

New Obs     Fit     SE Fit       95.0% CI              95.0% PI
1         17.59       2.51   ( 12.27,   22.90)  (     5.94,   29.23)
2         28.26       1.42   ( 25.26,   31.26)  (    17.47,   39.05)

Values of Predictors for New Observations

New Obs  Vacancy%
1            25.0
2            11.3
```

d. Since the p-value = 0.003 is less than α = .05, the relationship is significant.

e. r^2 = .434. The least squares line does not provide a very good fit.

f. The 95% confidence interval is 12.27 to 22.90 or $12.27 to $22.90.

g. The 95% prediction interval is 17.47 to 39.05 or $17.47 to $39.05.

45. a. $\bar{x} = \frac{\Sigma x_i}{n} = \frac{70}{5} = 14 \quad \bar{y} = \frac{\Sigma y_i}{n} = \frac{76}{5} = 15.2$

$\Sigma(x_i - \bar{x})(y_i - \bar{y}) = 200 \quad \Sigma(x_i - \bar{x})^2 = 126$

$$b_1 = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{200}{126} = 1.5873$$

$$b_0 = \bar{y} - b_1\bar{x} = 15.2 - (1.5873)(14) = -7.0222$$

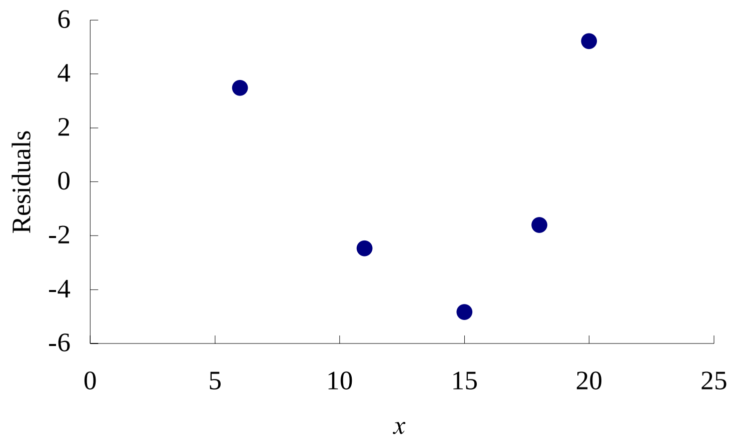
$$\hat{y} = -7.02 + 1.59x$$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

b. The residuals are 3.48, -2.47, -4.83, -1.6, and 5.22

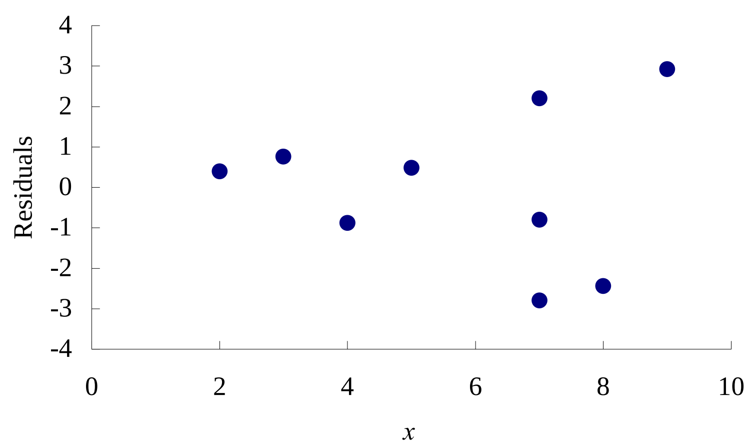
c.



With only 5 observations it is difficult to determine if the assumptions are satisfied. However, the plot does suggest curvature in the residuals that would indicate that the error term assumptions are not satisfied. The scatter diagram for these data also indicates that the underlying relationship between *x* and *y* may be curvilinear.

46. a. $\hat{y} = 2.32 + .64x$

b.



The assumption that the variance is the same for all values of *x* is questionable. The variance appears to increase for larger values of *x*.

47. a. Let *x* = advertising expenditures and *y* = revenue

$\hat{y} = 29.4 + 1.55x$

b. SST = 1002   SSE = 310.28   SSR = 691.72

MSR = SSR / 1 = 691.72

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com   Uploaded By: 1210711@student.birzeit.edu

SOUTH-WESTERN
CENGAGE Learning

MSE = SSE / (n - 2) = 310.28/ 5 = 62.0554

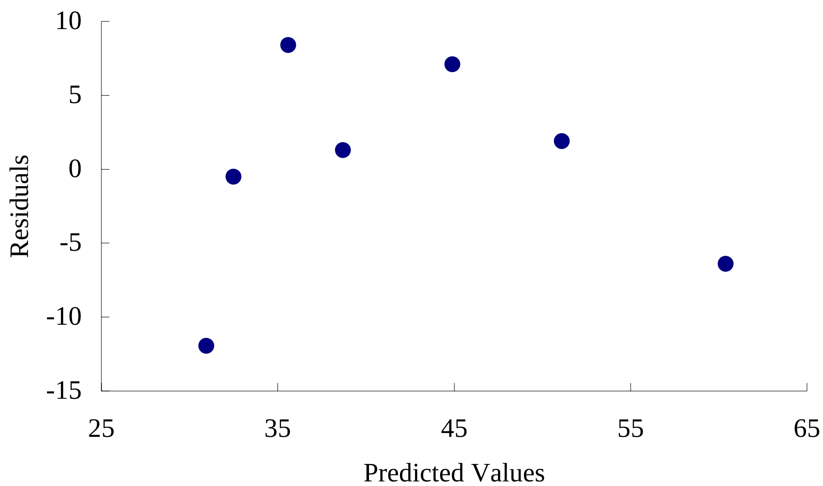
F = MSR / MSE = 691.72/ 62.0554= 11.15

Using F table (1 degree of freedom numerator and 5 denominator), p-value is between .01 and .025

Using Excel or Minitab, the p-value corresponding to F = 11.15 is .0206.

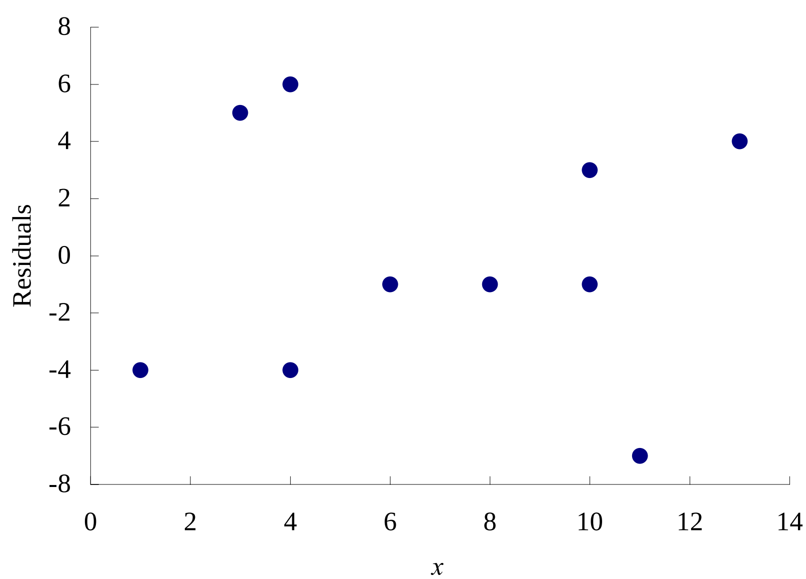
Because p-value $\leq \alpha$ = .05, we conclude that the two variables are related.

c.



d. The residual plot leads us to question the assumption of a linear relationship between x and y. Even though the relationship is significant at the .05 level of significance, it would be extremely dangerous to extrapolate beyond the range of the data.

48. a. $\hat{y} = 80 + 4x$



This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

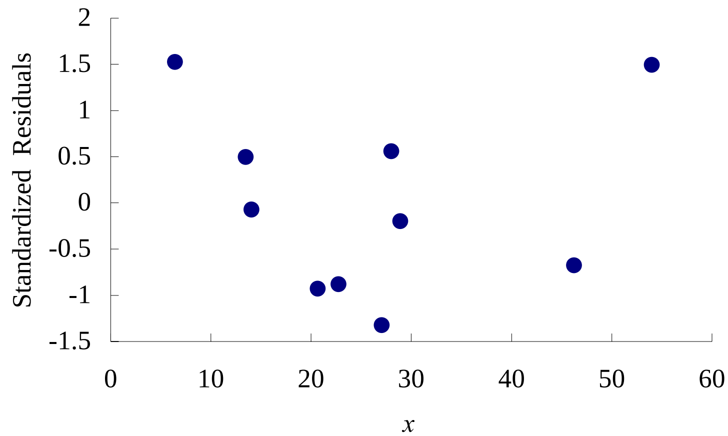
STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

b. The assumptions concerning the error term appear reasonable.

49. a. Let x = return on investment (ROE) and y = price/earnings (P/E) ratio.

$$\hat{y} = -32.13 + 3.22x$$

b.



c. There is an unusual trend in the residuals. The assumptions concerning the error term appear questionable.

50. a. The Minitab output is shown below:

```
The regression equation is
Price = 9.26 + 0.711 Shares

Predictor        Coef     SE Coef          T        P
Constant        9.265       1.099       8.43    0.000
Shares         0.7105      0.1474       4.82    0.001

S = 1.419       R-Sq = 74.4%     R-Sq(adj) = 71.2%

Analysis of Variance

Source           DF          SS          MS         F        P
Regression        1      46.784      46.784     23.22    0.001
Residual Error    8      16.116       2.015
Total             9      62.900
```

b. Since the p-value corresponding to F = 23.22 = .001 < α = .05, the relationship is significant.

c. r^2 = .744; a good fit. The least squares line explained 74.4% of the variability in Price.

d. $\hat{y} = 9.26 + .711(6) = 13.53$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

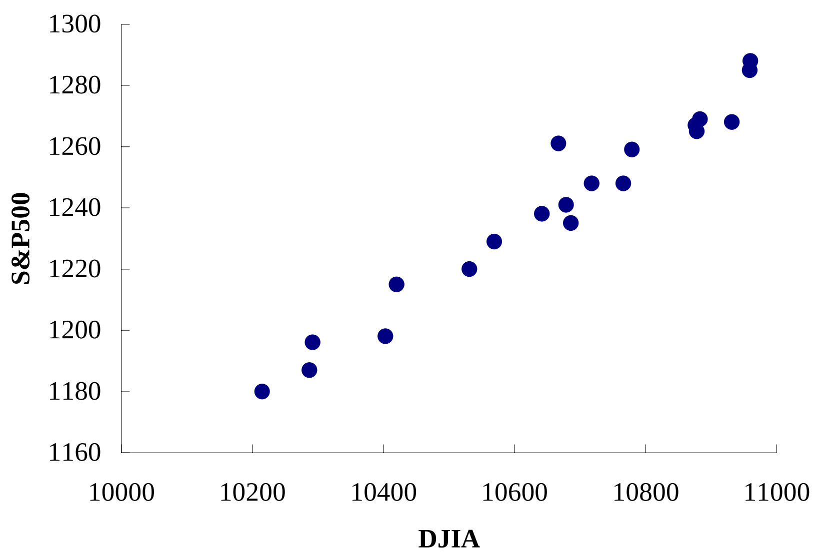
51. a. The Minitab output is shown below:

```
The regression equation is
Options = - 3.83 + 0.296 Common

Predictor          Coef      SE Coef          T        P
Constant         -3.834        5.903      -0.65    0.529
Common          0.29567      0.02648      11.17    0.000

S = 11.04       R-Sq = 91.9%     R-Sq(adj) = 91.2%
Analysis of Variance

Source            DF          SS          MS         F         P
Regression         1       15208       15208    124.72     0.000
Residual Error    11        1341         122
Total             12       16550
```

b. $\hat{y} = -3.83 + .296(150) = 40.57$; approximately 40.6 million shares of options grants outstanding.

c. r^2 = .919; a very good fit. The least squares line explained 91.9% of the variability in Options.

52. a.



b. The Minitab output is shown below:

```
The regression equation is
S&P500 = - 182 + 0.133 DJIA

Predictor       Coef    SE Coef      T      P
Constant     -182.11      71.83  -2.54  0.021
DJIA        0.133428   0.006739  19.80  0.000


S = 6.89993   R-Sq = 95.6%   R-Sq(adj) = 95.4%

Analysis of Variance

Source          DF     SS     MS       F      P
Regression       1  18666  18666  392.06  0.000
Residual Error  18    857     48
Total           19  19523
```


This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

c. Using the F test, the p-value corresponding to F = 392.06 is .000. Because the p-value $\leq \alpha$ =.05, we reject $H_0: \beta_1 = 0$; there is a significant relationship.

d. With R-Sq = 95.6%, the estimated regression equation provided an excellent fit.

e. $\hat{y} = -182.11 + .133428\text{DJIA} = -182.11 + .133428(11,000) = 1285.60$ or 1286.

f. The DJIA is not that far beyond the range of the data. With the excellent fit provided by the estimated regression equation, we should not be too concerned about using the estimated regression equation to predict the S&P500.

53. The Minitab output is shown below:

```
The regression equation is
Expense = 10.5 + 0.953 Usage

Predictor       Coef       SE Coef        T        p
Constant      10.528        3.745       2.81    0.023
X             0.9534       0.1382       6.90    0.000

S = 4.250       R-sq = 85.6%     R-sq(adj) = 83.8%


Analysis of Variance

SOURCE         DF          SS          MS         F        p
Regression      1      860.05      860.05     47.62    0.000
Residual Error 8      144.47       18.06
Total           9     1004.53

  Fit  Stdev.Fit         95% C.I.         95% P.I.
39.13       1.49   (  35.69,  42.57)  (  28.74,  49.52)
```

a. $\hat{y} =$ 10.5 + .953 Usage

b. Since the p-value corresponding to $F = 47.62 = .000 < \alpha = .05$, we reject H_0: $\beta_1 = 0$.

c. The 95% prediction interval is 28.74 to 49.52 or $2874 to $4952

d. Yes, since the expected expense is $3913.

54. a. The Minitab output is shown below:

```
The regression equation is
Defects = 22.2 - 0.148 Speed

Predictor        Coef     SE Coef          T        P
Constant       22.174       1.653      13.42    0.000
Speed        -0.14783     0.04391      -3.37    0.028

S = 1.489       R-Sq = 73.9%     R-Sq(adj) = 67.4%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression         1      25.130      25.130     11.33    0.028
Residual Error     4       8.870       2.217
Total              5      34.000
```


This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

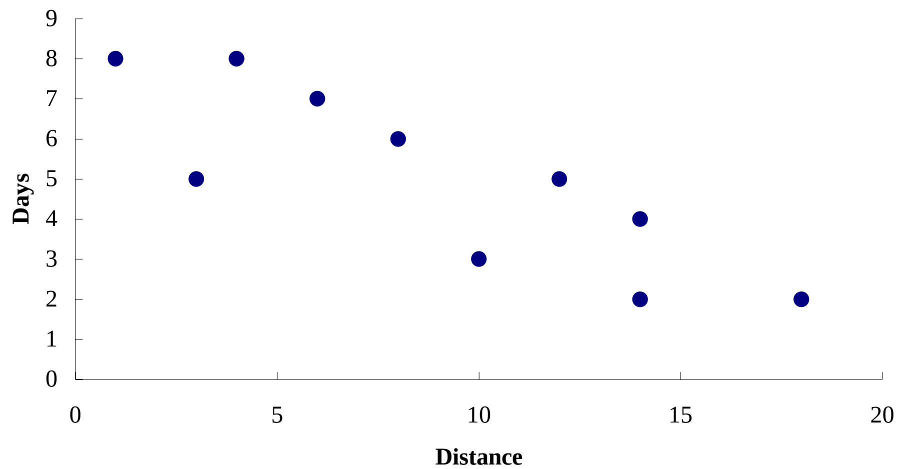
```
Predicted Values for New Observations

New Obs     Fit     SE Fit          95.0% CI              95.0% PI
1        14.783      0.896    (  12.294,  17.271)  (   9.957,  19.608)
```

b. Since the p-value corresponding to F = 11.33 = .028 < α = .05, the relationship is significant.

c. r^2 = .739; a good fit. The least squares line explained 73.9% of the variability in the number of defects.

d. Using the Minitab output in part (a), the 95% confidence interval is 12.294 to 17.271.

55. a.



There appears to be a negative linear relationship between distance to work and number of days absent.

b. The MINITAB output is shown below:

```
The regression equation is
Days = 8.10 - 0.344 Distance

Predictor        Coef      SE Coef          T        p
Constant       8.0978       0.8088      10.01    0.000
X            -0.34420      0.07761      -4.43    0.002

S = 1.289       R-sq = 71.1%     R-sq(adj) = 67.5%

Analysis of Variance

SOURCE          DF          SS          MS         F        p
Regression       1      32.699      32.699     19.67    0.002
Residual Error 8      13.301       1.663
Total            9      46.000

      Fit   Stdev.Fit         95% C.I.          95% P.I.
   6.377       0.512    (  5.195,   7.559)  (  3.176,   9.577)
```

c. Since the p-value corresponding to F = 419.67 is .002 < α = .05. We reject $H_0 : \beta_1 = 0$.

d. r^2 = .711. The estimated regression equation explained 71.1% of the variability in y; this is a reasonably good fit.

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

e. The 95% confidence interval is 5.195 to 7.559 or approximately 5.2 to 7.6 days.

56. a. The Minitab output is shown below:

```
The regression equation is
Cost = 220 + 132 Age

Predictor       Coef     SE Coef          T        p
Constant      220.00       58.48       3.76    0.006
X             131.67       17.80       7.40    0.000

S = 75.50       R-sq = 87.3%     R-sq(adj) = 85.7%

Analysis of Variance

SOURCE         DF          SS          MS         F        p
Regression      1      312050      312050     54.75    0.000
Residual Error 8        45600        5700
Total           9      357650

     Fit  Stdev.Fit         95% C.I.         95% P.I.
   746.7       29.8   (  678.0,  815.4)  (  559.5,  933.9)
```

b. Since the p-value corresponding to $F = 54.75$ is $.000 < \alpha = .05$, we reject H_0: $\beta_1 = 0$.

c. $r^2 = .873$. The least squares line provided a very good fit.

d. The 95% prediction interval is 559.5 to 933.9 or \$559.50 to \$933.90

57. a. The Minitab output is shown below:

```
The regression equation is
Points = 5.85 + 0.830 Hours

Predictor       Coef     SE Coef          T        p
Constant       5.847       7.972       0.73    0.484
X             0.8295      0.1095       7.58    0.000

S = 7.523       R-sq = 87.8%     R-sq(adj) = 86.2%

Analysis of Variance

SOURCE         DF          SS          MS         F        p
Regression      1      3249.7      3249.7     57.42    0.000
Residual Error 8        452.8        56.6
Total           9      3702.5

     Fit  Stdev.Fit         95% C.I.         95% P.I.
   84.65       3.67   (  76.19,  93.11)  (  65.35, 103.96)
```

b. Since the p-value corresponding to $F = 57.42$ is $.000 < \alpha = .05$, we reject H_0: $\beta_1 = 0$.

c. 84.65 points

d. The 95% prediction interval is 65.35 to 103.96

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

58. a. The Minitab output is shown below:

```
The regression equation is
Horizon = 0.275 + 0.950 S&P 500

Predictor       Coef     SE Coef          T        P
Constant      0.2747      0.9004       0.31    0.768
S&P 500       0.9498      0.3569       2.66    0.029

S = 2.664       R-Sq = 47.0%     R-Sq(adj) = 40.3%

Analysis of Variance

Source         DF          SS          MS         F       P
Regression      1      50.255      50.255      7.08    0.029
Residual Error 8       56.781       7.098
Total           9     107.036
```

The market beta for Horizon is $b_1 = .95$

b. Since the p-value = 0.029 is less than $\alpha = .05$, the relationship is significant.

c. $r^2 = .470$. The least squares line does not provide a very good fit.

d. Texas Instruments has higher risk with a market beta of 1.46.

59. a. The Minitab output is shown below:

```
The regression equation is
Audit% = - 0.471 +0.000039 Income

Predictor        Coef     SE Coef          T        P
Constant      -0.4710      0.5842      -0.81    0.431
Income     0.00003868  0.00001731       2.23    0.038

S = 0.2088      R-Sq = 21.7%     R-Sq(adj) = 17.4%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression         1     0.21749     0.21749      4.99    0.038
Residual Error    18     0.78451     0.04358
Total             19     1.00200

Predicted Values for New Observations

New Obs     Fit    SE Fit        95.0% CI              95.0% PI
1        0.8828    0.0523   ( 0.7729,  0.9927)  ( 0.4306,  1.3349)
```

b. Since the p-value = 0.038 is less than $\alpha = .05$, the relationship is significant.

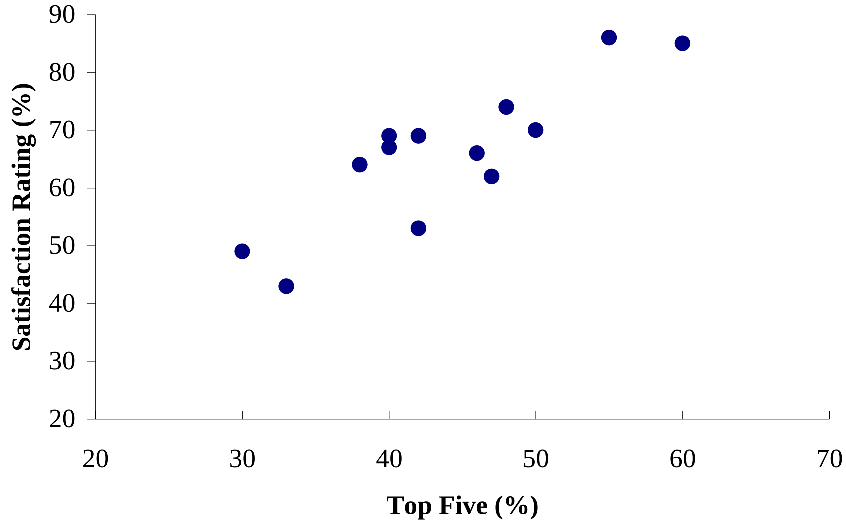
c. $r^2 = .217$. The least squares line does not provide a very good fit.

d. The 95% confidence interval is .7729 to .9927.

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu

60. a.



b. There appears to be a positive linear relationship between the two variables.

A portion of the Minitab output for this problem follows.

```
The regression equation is
Rating (%) = 9.4 + 1.29 Top Five (%)

Predictor        Coef  SE Coef     T      P
Constant         9.37    10.24  0.92  0.380
Top Five (%)   1.2875   0.2293  5.61  0.000

S = 6.62647   R-Sq = 74.1%   R-Sq(adj) = 71.8%

Analysis of Variance

Source          DF      SS      MS      F      P
Regression       1  1383.9  1383.9  31.52  0.000
Residual Error  11   483.0    43.9
Total           12  1866.9
```

c. $\hat{y}$ = 9.37 + 1.2875 Top Five (%)

d. Since the p-value = .000 < α = .05, the relationship is significant.

e. r^2 = .741; a good fit. The least squares line explained 74.1% of the variability in the satisfaction rating.

f. $r_{xy} = \sqrt{r^2} = \sqrt{.741} = .86$

This edition is intended for use outside of the U.S. only, with content that may be different from the U.S. Edition. This may not be resold, copied, or distributed without the prior consent of the publisher.

STUDENTS-HUB.com Uploaded By: 1210711@student.birzeit.edu