CHAPTER 10 FORM A

Name_____ Course Number:_____ Section Number:_____

**Directions: Answer the questions and solve the problems in the spaces provided, or attach paper. Circle the correct choice for each response set. Where allowed, use computer assistance.**

**Provide an appropriate response.**

1) A set of data consists of the number of years that applicants for foreign service jobs have studied German and the grades that they received on a proficiency test. The following regression equation is obtained: $\hat{y} = 31.6 + 10.9x$, where x represents the number of years of study and y represents the grade on the test. Identify the predictor and response variables.

**Given the linear correlation coefficient r and the sample size n, determine the critical values of r and use your finding to state whether or not the given r represents a significant linear correlation. Use a significance level of 0.05.**

2) $r = -0.844$, $n = 5$

A) Critical values: $r = \pm 0.950$, no significant linear correlation

B) Critical values: $r = \pm 0.878$, significant linear correlation

C) Critical values: $r = \pm 0.878$, no significant linear correlation

D) Critical values: $r = 0.950$, significant linear correlation

**Find the value of the linear correlation coefficient r.**

3) The paired data below consist of the test scores of 6 randomly selected students and the number of hours they studied for the test.

| Hours | 5 | 10 | 4 | 6 | 10 | 9 |
|---|---|---|---|---|---|---|
| Score | 64 | 86 | 69 | 86 | 59 | 87 |

A) 0.224 B) −0.224 C) −0.678 D) 0.678

**Describe the error in the stated conclusion.**

4) Given: There is no significant linear correlation between scores on a math test and scores on a verbal test. Conclusion: There is no relationship between scores on the math test and scores on the verbal test.

**Use the given data to find the best predicted value of the response variable.**

5) Eight pairs of data yield $r = 0.708$ and the regression equation $\hat{y} = 55.8 + 2.79x$. Also, $\bar{y} = 71.125$. What is the best predicted value of y for $x = 5.7$?

A) 71.13 B) 71.7 C) 320.85 D) 57.80

**Use the given data to find the equation of the regression line. Round the final values to three significant digits, if necessary.**

6)

| x | 0 | 3 | 4 | 5 | 12 |
|---|---|---|---|---|---|
| y | 8 | 2 | 6 | 9 | 12 |

A) $\hat{y} = 4.98 + 0.725x$

B) $\hat{y} = 4.88 + 0.525x$

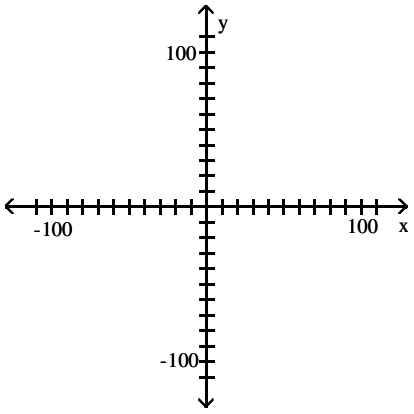
C) $\hat{y} = 4.88 + 0.625x$

D) $\hat{y} = 4.98 + 0.425x$

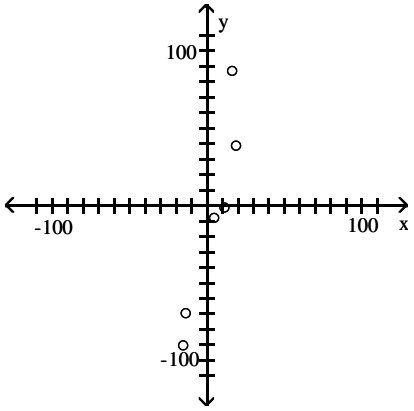
**Construct a scatterplot for the given data.**

7)

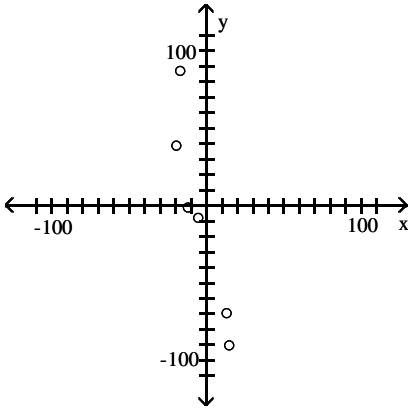
| x | 15 | −5 | −19 | 13 | −17 | −12 |
|---|---|---|---|---|---|---|
| y | −91 | −8 | 38 | −70 | 87 | −2 |



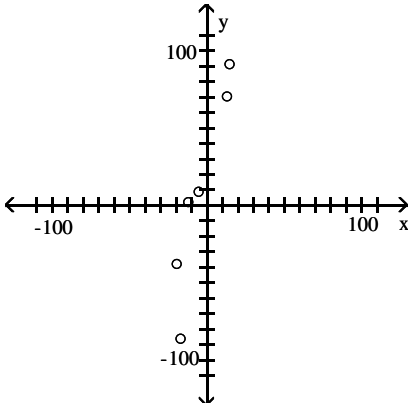

A)



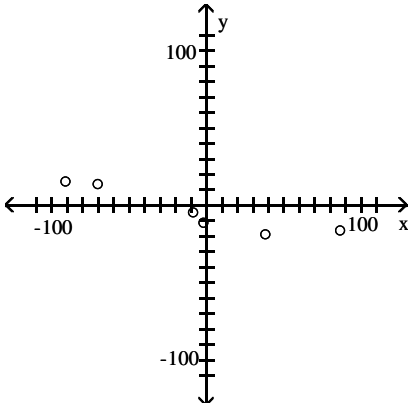

B)




C)




D)



**Is the data point, P, an outlier, an influential point, both, or neither?**

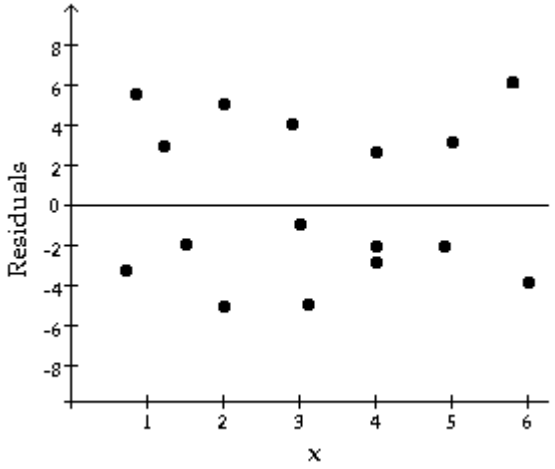
8) The regression equation for a set of paired data is $\hat{y} = 7 + 5x$. The correlation coefficient for the data is 0.88. A new data point, P(15, 94), is added to the set.

A) Neither B) Outlier

C) Both D) Influential point

**Provide an appropriate response.**

9) The following residual plot is obtained after a regression equation is determined for a set of data. Does the residual plot suggest that the regression equation is a bad model? Why or why not?




**Use the given information to find the coefficient of determination.**

10) A regression equation is obtained for a collection of paired data. It is found that the total variation is 29.045, the explained variation is 15.212, and the unexplained variation is 13.833. Find the coefficient of determination.

A) 0.476 B) 0.909 C) 0.524 D) 1.909

**Find the standard error of estimate for the given paired data.**

11) The equation of the regression line for the paired data below is $\hat{y} = 3x$. Find the standard error of estimate.

| x | 2 | 4 | 5 | 6 |
|---|---|---|---|---|
| y | 7 | 11 | 13 | 20 |

A) 5.00 B) 4.1892 C) 6.2750 D) 2.2361

**Use the computer display to answer the question.**

12) A collection of paired data consists of the number of years that students have studied Spanish and their scores on a Spanish language proficiency test. A computer program was used to obtain the least squares linear regression line and the computer output is shown below. Along with the paired sample data, the program was also given an x value of 2 (years of study) to be used for predicting test score.

The regression equation is

Score = 31.55 + 10.90 Years.

| Predictor | Coef | StDev | T | P |
|---|---|---|---|---|
| Constant | 31.55 | 6.360 | 4.96 | 0.000 |
| Years | 10.90 | 1.744 | 6.25 | 0.000 |

S = 5.651 R–Sq = 83.0% R–Sq (Adj) = 82.7%

Predicted values

| Fit | StDev Fit | 95.0% CI | 95.0% PI |
|---|---|---|---|
| 53.3 5 | 3.168 | (42.72, 63.98) | (31.61, 75.09) |

For a person who studies for 2 years, obtain the 95% prediction interval and write a statement interpreting the interval.

A) (31.61, 75.09); We can be 95% confident that the test score of an individual who studies 2 years will lie in the interval (31.61, 75.09)

B) (42.72, 63.98); We can be 95% confident that the test score of an individual who studies 2 years will lie in the interval (42.72, 63.98)

C) (31.61, 75.09); We can be 95% confident that the mean test score of all individuals who study 2 years will lie in the interval (31.61, 75.09)

D) (42.72, 63.98); We can be 95% confident that the mean test score of all individuals who study 2 years will lie in the interval (42.72, 63.98)

**Find the total variation for the paired data.**

13) The paired data below consists of heights and weights of 6 randomly selected adults. The equation of the regression line is $\hat{y} = -181.342 + 144.46x$. Find the total variation.

| x Height (meters) | 1.61 | 1.72 | 1.78 | 1.80 | 1.67 | 1.88 |
|---|---|---|---|---|---|---|
| y Weight (kg) | 54 | 62 | 70 | 84 | 61 | 92 |

A) 1,119.3 B) 100.06 C) 1,079.5 D) 979.44

**Construct the indicated prediction interval for an individual y.**

14) The paired data below consists of test scores and hours of preparation for 5 randomly selected students. The equation of the regression line is $\hat{y} = 44.845 + 3.524x$ and the standard error of estimate is $s_e = 5.40$. Find the 99% prediction interval for the test score of a person who spent 7 hours preparing for the test.

| x Hours of preparation | 5 | 2 | 9 | 6 | 10 |
|---|---|---|---|---|---|
| y Test score | 64 | 48 | 72 | 73 | 80 |

A) $35 < y < 104$ B) $32 < y < 107$ C) $58 < y < 82$ D) $62 < y < 78$

**Use computer software to find the multiple regression equation. Can the equation be used for prediction?**

15) FPEA, the Farm Production Enhancement Agency, regressed corn output against acreage, rainfall, and a trend line. The trend line is proxy for technological advancement in farming from improved pest control, fertilization, land management, and farming implements.

| CORNPROD | ACRES | RAIN | TREND |
|---|---|---|---|
| 456 | 9896 | 29.1 | 1 |
| 421 | 9680 | 42.3 | 2 |
| 653 | 10449 | 29.8 | 3 |
| 573 | 10811 | 26.0 | 4 |
| 546 | 10014 | 34.3 | 5 |
| 499 | 10293 | 22.7 | 6 |
| 504 | 9413 | 24.2 | 7 |
| 611 | 9860 | 31.6 | 8 |
| 646 | 9782 | 25.6 | 9 |
| 789 | 12139 | 37.9 | 10 |
| 773 | 12166 | 33.9 | 11 |
| 753 | 9976 | 37.4 | 12 |
| 852 | 10645 | 27.0 | 13 |
| 755 | 9738 | 31.5 | 14 |
| 815 | 9933 | 39.9 | 15 |
| 902 | 10132 | 25.3 | 16 |
| 986 | 11145 | 30.4 | 17 |
| 909 | 9775 | 32.7 | 18 |
| 945 | 9549 | 35.0 | 19 |
| 866 | 10077 | 33.8 | 20 |
| 1178 | 11550 | 29.4 | 21 |
| 1230 | 10600 | 37.1 | 22 |
| 1207 | 11280 | 42.9 | 23 |
| 968 | 12100 | 32.2 | 24 |
| 1118 | 12420 | 30.5 | 25 |

A) CORNPROD = −16.3 + 2.6ACRES + 3.9RAIN + 21.3TREND; Yes, because the the $R^2$ is high

B) CORNPROD = −21.1 + 0.036ACRES + 2.62RAIN + 27.6TREND; Yes, because the $R^2$ is high

C) CORNPROD = −0.9 + 1.68ACRES + 0.79RAIN + 10.2TREND; Yes, because the adjusted $R^2$ is high

D) CORNPROD = −21.1 + 0.036ACRES + 2.62RAIN + 27.6TREND; No, because the P−value is low

**Use computer software to obtain the multiple regression equation and identify $R^2$, adjusted $R^2$, and the P−value.**

16) A visitor to Yellowstone National Park sat down one day and observed Old Faithful, which faithfully spurts throughout the day, day in and day out. He surmised that the height of a given spurt was caused by the pressure build−up during the interval between spurts and by the momentum build−up during the duration of the spurt. He wrote down the data to test his hypothesis, but he didn't know what to do with his data. Can you help him out with his theory?

| HEIGHT | INTERVAL | DURATION |
|---|---|---|
| 150 | 86 | 240 |
| 154 | 86 | 237 |
| 140 | 62 | 122 |
| 140 | 104 | 267 |
| 160 | 62 | 113 |
| 140 | 95 | 258 |
| 150 | 79 | 232 |
| 150 | 62 | 105 |
| 160 | 94 | 276 |
| 155 | 79 | 248 |
| 125 | 86 | 243 |
| 136 | 85 | 241 |
| 140 | 86 | 214 |
| 155 | 58 | 114 |
| 130 | 89 | 272 |
| 125 | 79 | 227 |
| 125 | 83 | 237 |
| 139 | 82 | 238 |
| 125 | 84 | 203 |
| 140 | 82 | 270 |
| 140 | 82 | 270 |
| 140 | 78 | 218 |
| 135 | 87 | 270 |
| 140 | 70 | 241 |
| 100 | 56 | 102 |
| 105 | 81 | 271 |

A) 0.025, −0.060, 0.750

B) 0.213, 0.182, 0.213

C) 0.049, −0.021, 0.123

D) 0.089, 0.032, 0.634

**Use computer software to obtain the multiple regression equation. Use the estimated equation to find the predicted value.**

17) A wildlife analyst gathered the data in the table to develop an equation to predict the weights of bears. He used WEIGHT as the dependent variable and CHEST, LENGTH, and SEX as the independent variables. For SEX, he used male = 1 and female = 2. He took his equation "to the forest" and found a male bear whose chest measured 40.3 inches and who was 64.0 inches long.

| WEIGHT | CHEST | LENGTH | SEX |
|---|---|---|---|
| 344 | 45.0 | 67.5 | 1 |
| 416 | 54.0 | 72.0 | 1 |
| 220 | 41.0 | 70.0 | 2 |
| 360 | 49.0 | 68.5 | 1 |
| 332 | 44.0 | 73.0 | 1 |
| 140 | 32.0 | 63.0 | 2 |
| 436 | 48.0 | 72.0 | 1 |
| 132 | 33.0 | 61.0 | 2 |
| 356 | 48.0 | 64.0 | 2 |
| 150 | 35.0 | 59.0 | 1 |
| 202 | 40.0 | 63.0 | 2 |
| 365 | 50.0 | 70.5 | 1 |

A) 415 lb B) 293 lb C) 405 lb D) 252 lb

**Find the indicated multiple regression equation.**

18) Below are the results of two separate tests designed to measure a student's ability to solve problems.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Test A | 48 | 52 | 58 | 44 | 43 | 43 | 40 | 51 | 59 |
| Test B | 73 | 67 | 73 | 59 | 58 | 56 | 58 | 64 | 74 |

In addition to these results, a third test was designed to measure the same problem–solving ability, and the following results correspond to the same students.

Test C: 48 41 59 45 42 44 40 58 60

Find the multiple regression equation that expresses results from Test C in terms of Test A and Test B.

A) $\hat{C} = 8.16 + 1.12A + 3.49B$ B) $\hat{C} = 3.40 + 1.12A - 0.145B$

C) $\hat{C} = 1.60 + 2.02A - 1.93B$ D) $\hat{C} = 3.40 - 0.145A + 1.12B$

**Use computer software to find the best multiple regression equation to explain the variation in the dependent variable, Y, in terms of the independent variables, $X_1$, $X_2$, $X_3$.**

19)

| Y | $X_1$ | $X_2$ |
|---|---|---|
| 15 | 1.2 | 16 |
| 15 | 1.2 | 16 |
| 17 | 1.0 | 16 |
| 6 | 0.8 | 9 |
| 1 | 0.1 | 1 |
| 8 | 0.8 | 8 |
| 10 | 0.8 | 10 |
| 17 | 1.0 | 16 |
| 15 | 1.2 | 15 |
| 11 | 0.7 | 9 |
| 18 | 1.4 | 18 |
| 16 | 1.0 | 15 |
| 10 | 0.8 | 9 |
| 7 | 0.5 | 5 |
| 18 | 1.1 | 16 |

CORRELATION COEFFICIENT

$Y/X_1 = 0.886$

$Y/X_2 = 0.965$

COEFFICIENTS OF DETERMINATION

$Y/X_2 = 0.932$

$Y/X_2, X_1 = 0.943$

A) $\hat{Y} = 0.42 + 0.99 X_2$

B) $\hat{Y} = 1.25 - 1.55 X_1 + 5.79 X_2$

C) $\hat{Y} = -0.49 + 14.07 X_1$

D) $\hat{Y} = 1.38 - 5.53 X_1 + 1.33 X_2$

**Construct a scatterplot and identify the mathematical model that best fits the data. Assume that the model is to be used only for the scope of the given data and consider only linear, quadratic, logarithmic, exponential, and power models. Use a calculator or computer to obtain the regression equation of the model that best fits the data. You may need to fit several models and compare the values of $R^2$.**

20) A rock is dropped from a tall building and its distance (in feet) below the point of release is recorded as accurately as possible at various times after the moment of release. The results are shown in the table. Find the regression equation of the best model.

| x (seconds after release) | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| y (distance in feet) | 16 | 63 | 146 | 255 | 403 | 572 |

A) $y = -148.4 + 112 x$

B) $y = 13.0 e^{0.686 x}$

C) $y = -74.9 + 290 \ln x$

D) $y = 15.95 x^2$

1) The predictor variable is the number of years of study, and the response variable is the grade on the test.
2) C
3) A
4) Although there is no significant linear correlation between the two variables, we cannot conclude that no relationship exists as there could be a nonlinear relationship between the two variables.
5) B
6) B
7) B
8) C
9) No, the residual plot does not suggest that the regression equation is a bad model. The residual plot does not have an obvious pattern that is not a straight line. This confirms that a scatterplot of the sample data is a straight line. The residual plot does not become thicker or thinner when viewed from left to right. This confirms that for different fixed values of x, the distributions of the corresponding y–values all have the same standard deviation.
10) C
11) D
12) A
13) C
14) A
15) B
16) A
17) D
18) B
19) A
20) D