1. Suppose a data set has a linear regression line of $y = 6 - 0.8x$. If the mean of the x's is 5, what is the mean of the y's?

> **A.** 2

> **B.** 5

> **C.** 10

> **D.** 6

> **E.** -5

2. If an observed *y*-value is below a line of best fit, then the residual is

   A. positive.

   B. negative.

   C. equal to the squared residual.

   D. greater than one.

   E. None of the above

3. You have the following regression equation for the effect of streetlights per block (x), on crimes per month (y): $y = 2.4 - 0.2x$. How many crimes a month are predicted when there are 7 street lights on a block?

    A. 3.8

    B. 1.7

    C. 16.6

    D. -11.6

    E. 1.0

4. You have the following regression equation for the effect of streetlights per block (x), on crimes per month (y): $y = 2.4 - 0.2x$. Calculate the residual for a block with 10 streetlights and 1 crime a month (10,1).

    A. -0.6

    B. 0.6

    C. -0.4

    D. 0.4

    E. -1.2

5. Describe the strength and direction of a relationship with the correlation
   coefficient $r = -0.8$.

   A. Moderate and negative

   B. Weak and negative

   C. Strong and negative

   D. Weak and positive

   E. There is no association

6. Which of the following determines the sign of r?

   A. The pattern of the residuals

   B. The strength of the relationship between variables

   C. Whether the y-intercept, b_0, is positive or negative

   D. Whether the value of y increases or decreases as the value of x increases

   E. Whether the sum of the squared residuals is positive or negative

7. A bivariate scatterplot has an r^2 of .85. This means:

    A. 15% of the variation in y is explained by the changes in x.

    B. 15% of the variation in x is explained by the changes in y.

    C. 15% of the variation in x isn't explained by y.

    D. 85% of the variation in y is explained by the changes in x.

    E. 85% of the variation in x is explained by the changes in y.

8. If the least-squares linear regression line explained the same amount of variation as the line $\hat{y} = \overline{y}$, what would be the value of r^2?

   A. 1.00

   B. 0.50

   C. 0

   D. -1.00

   E. Can't answer with only this information

9. A residual:

   A. is the amount of variation explained by the least-squares regression line of y on x.

   B. is how much an observed y-value differs from a predicted y-value.

   C. predicts how well x explains y.

   D. is the total variation of the data points.

   E. should be smaller than the mean of y.

10. A linear regression line indicates the amount of grams of the chemical $CuSO_4$ (the response variable, y) that dissolve in water at various temperatures, Celsius (the explanatory variable, x). The least-squares regression line is $y = 10.14 + 0.51x$. Give the meaning of the slope of the regression line in the context of the problem.

A. For each one-degree rise in the temperature, you can dissolve 10.14 more grams of $CuSO_4$.

B. If the temperature increases by 0.51, you can dissolve one more gram of $CuSO_4$.

C. When you dissolve one more gram of $CuSO_4$, then the temperature will rise by 0.51.

D. For each one-degree rise in the temperature, you can dissolve 0.51 fewer grams of $CuSO_4$.

E. For each one-degree rise in the temperature, you can dissolve 0.51 more grams of $CuSO_4$.

11. If you're attempting to predict a value of the response variable using a value of x that is outside the range of observed x-values in your data set, you're conducting a process of:

    A. predicting the slope of the regression line.

    B. interpolation.

    C. computing residuals.

    D. extrapolation.

    E. slope interpretation.

12. Which of the following statements about influential points is true?

    I. Removing an influential point from a data set can have a major effect on the regression line.

    II. If you calculate the residual between the influential point and a regression line based on the rest of the data, it will probably be large.

    III. You will typically find an influential point horizontally distant from the rest of the data along the x-axis.

A. I only

B. II only

C. III only

D. I and II only

E. I and III only

13. An outlier:

    A. usually does not have a strong effect on the correlation coefficient and regression line.

    B. can also be an influential point.

    C. may be an error.

    D. usually does not have a strong effect on the regression line, can also be an influential point, and may be an error.

    E. usually has a strong effect on the correlation coefficient and regression line and can also be an influential point.

14. Influential points and outliers:

A. sometimes have no effect on the regression line.

B. are useful when they cause a stronger correlation coefficient for a data set.

C. should be discarded if they cause a weaker correlation coefficient for a data set.

D. are rarely found in data sets.

E. should be examined carefully to determine if they're part of the data set.