

Pre-Calculus Multiple Choice Questions - Chapter S8

1 If every man married a women who was exactly 3 years younger than he, what would be the correlation between the ages of married men and women?

- a Somewhat negative
- b 0
- c Somewhat positive
- d Nearly 1
- e 1

	S9.1
--	------

2 Which of the following statements about the correlation of r is incorrect?

- a The correlation and the slope of the regression line always have the same sign
- b A correlation of -0.32 and a correlation of $+0.32$ show the same degree of clustering around the reg. line
- c Correlation r measures the strength and direction of linear association only
- d A correlation of 0.78 indicates a relationship that is 3 times as linear as one for which the correlation is 0.26
- e Outliers can greatly affect the value of r

	S9.1
--	------

3 Suppose a study finds that the correlation coefficient relating job satisfaction to salary is $r = +1$. Which of the following is a proper conclusion?

- a High salary causes high job satisfaction
- b Low salary causes low job satisfaction
- c There is a 100% cause-and-effect relationship between salary and job satisfaction
- d There is a very strong association between salary and job satisfaction
- e None of the above are proper conclusions

	S9.1
--	------

Pre-Calculus Multiple Choice Questions - Chapter S8

1 Suppose the correlation of a data set is negative. Given two points from the scatterplot, which of the following is possible?

- I. The first point has a larger x-value and a smaller y-value than the second point
- II. The first point has a larger x-value and a larger y-value than the second value
- III. The first point has a smaller x-value and a larger y-value than the second point.

- a I only
- b II only
- c III only
- d I and III only
- e I, II, and III

	S9.2
--	------

2 A study is conducted relating AP Statistics exam scores to the total number of study hours for the AP Statistics class put in by students during the academic year, and the correlation is found to be 0.6. Which of the following is a true statement?

- a On the average, a 40% increase in study time results in a 24% increase in average score
- b On the average, a 60% increase in study time results in a 100% increase in exam score
- c 60% of a student's exam score can be explained by the number of study hours
- d 60% of the variation in exam scores can be accounted for by this linear regression model
- e Higher exam scores tend to be associated with higher numbers of study hours

	S9.2
--	------

3 Which of the following is a true statement about the correlation coefficient r ?

- a A correlation of 0.3 means that 30% of the points are highly correlated
- b The square of the correlation measures the proportion of the y-variance that is predictable from a knowledge of x
- c Perfect correlation, that is, when the points lie exactly on a straight line, results in $r = 0$
- d Multiplying every y-value by -1 leaves the correlation unchanged
- e The unit of measurement for correlation is the y-unit per x-unit

	S9.2
--	------

Pre-Calculus Multiple Choice Questions - Chapter S8

2 Consider the three points (4, 33), (5, 27), and (6, 15). Given any straight line, we can calculate the sum of the squares of the three vertical distances from these points to the line. What is the smallest possible values this sum can be?

- a 2.45
- b 6
- c 8.66
- d 36
- e None of the above

	S9.3
--	-------------

3 Suppose the regression line for a set of data, $y = a + 4x$, passes through the point (1,3). If \bar{x} and \bar{y} are the sample means of the x- and y-values, respectively, then $y =$

- a \bar{x}
- b $4\bar{x}$
- c $3 + 4\bar{x}$
- d $2 + \bar{x}$
- e $-1 + 4\bar{x}$

	S9.3
--	-------------

1 Exercise physiologists are investigating the relationship between lean body mass (in kilograms) and the resting metabolic rate (in calories per day) in sedentary males

	S9.3
--	-------------

Predictor	Coef	StDev	T	P
Constant	264	276.9	0.95	0.363
Mass	22.563	6.36	3.55	0.005

S = 144.9 R-sq = 55.7% R-sq (adj) = 51.3%

Based on the computer output above, which of the following is the best interpretation of the value of the slope of the regression line?

- a For each additional kilogram of lean body mass, the resting metabolic rate increases on average by 22.563 calories per day
- b For each additional kilogram of lean body mass, the resting metabolic rate increases on average by 264.0 calories per day
- c For each additional kilogram of lean body mass, the resting metabolic rate increases on average by 144.9 calories per day
- d For each additional calorie per day for the resting metabolic rate, the lean body mass increases by 22.563 kilograms
- e For each additional calorie per day for the resting metabolic rate, the lean body mass increases on average by 264.0 kilograms

Pre-Calculus Multiple Choice Questions - Chapter S8

- 2 A simple random sample of 25 world-ranked tennis players provides the following statistics:
Number of hours of practice per day: $\bar{x} = 7.3$, $s_x = 1.2$.

Yearly winnings: $\bar{y} = \$1820000$, $s_y = \$310000$

Correlation = 0.23

Based on this data, what is the resulting linear regression equation

- a Winnings = $1390000 + 59400(\text{hours})$ b Winnings = $1300000 + 71300(\text{hours})$
c Winnings = $-63400 + 258000(\text{hours})$ d Winnings = $-443000 + 310000(\text{hours})$
e Winnings = $-10000000 + 1620000(\text{hours})$

	S9.3
--	------

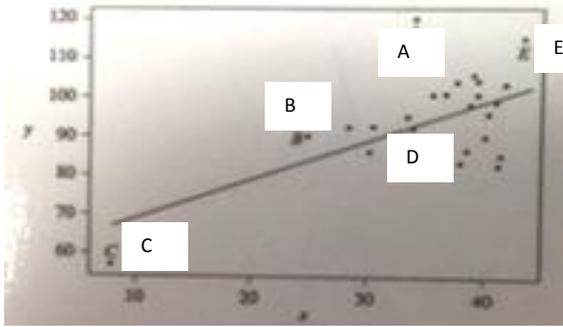
- 3 Consider n pairs of numbers. Suppose $\bar{x} = 4$, $s_x = 3$, $\bar{y} = 2$, and $s_y = 5$. Of the following, which could be the least squares line?

- a $y = 2 + x$ b $y = -6 + 2x$
c $y = -10 + 3x$ d $y = (5/3) - x$
e $y = 6 - x$

	S9.3
--	------

Pre-Calculus Multiple Choice Questions - Chapter S8

1



In the scatterplot of y versus x shown above, the least squares regression line is superimposed on the plot. Which of the following points has the largest residual?

- a A
- b B
- c C
- d D
- e E

	S9.4
--	-------------

2 Which of the following statements about residuals is incorrect?

- a The mean of the residuals is always zero
- b The sum of the residuals is always zero
- c The regression line for a residual plot is a horizontal line
- d A residual equals the predicted y minus the observed y
- e The standard deviation of the residuals gives a measure of how the points in the scatterplot are spread around the regression line

	S9.4
--	-------------

3 Which of the following statements about influential points is incorrect?

- a Looking at a residual plot is a helpful way of picking out influential points
- b Removal of an influential point sharply affects the regression line
- c Determine a regression model with and without a point is an excellent way of picking out inf. points
- d Points with high leverage may not be influential points
- e None of the above are incorrect statements

	S9.4
--	-------------

Pre-Calculus Multiple Choice Questions - Chapter S8

- 1 A linear regression analysis is performed on the data from two scatterplots, A and B, resulting in identical least squares regression lines with positive slopes. Which of the following statements is true?
- a The sum of the squares of the residuals in A equals the sum of the squares of the residuals in B
 - b The correlation in A equals the correlation in B
 - c If the sum of the squares of the residuals in A is greater than the sum of the squares of the residuals in B, then the correlation in A will be greater than the correlation in B.
 - d If the sum of the squares of the residuals in A is greater than the sum of the squares of the residuals in B, then the correlation in A will be less than the correlation in B.
 - e None of the above

	S9.5
--	-------------

- 2 The proportion of variability in Y that is explained by the independent variable X in the regression model is known as
- a Coefficient of Determination
 - b Coefficient of Non-Determination
 - c Coefficient of Correlation
 - d Standard Error of the Estimate

	S9.5
--	-------------

- 3 The coefficient of determination tells us
- a The proportion of total variation in x that cannot be explained by y
 - b The proportion of total variation in x that is explained by y
 - c The proportion of total variation in y that is explained by x
 - d The proportion of total variation in y that cannot be explained by x

	S9.5
--	-------------

- 1 Data are obtained from a random sample of adult women with regard to their ages and their monthly expenditures on health products. The resulting regression equation is:
Expenditure = 43 + 0.23(Age) $r = 0.27$
What percentage of the variation in expenditures can be explained by looking at ages?
- a 0.23%
 - b 23%
 - c 7.29%
 - d 27%
 - e 52%

	S9.5
--	-------------

