Test 3C AP Statistics Name:

Directions: *Work on these sheets.*

Part 1: Multiple Choice *Circle the letter corresponding the best answer.*

1. A study found correlation $r = 0.61$ between the sex of a worker and his or her income. You conclude that
(a) women earn more than men on the average.
(b) women earn less than men on average.
(c) an arithmetic mistake was made; this is not a possible value of r.
(d) this is nonsense because r makes no sense here.
(e) the correlation should have been $r = -0.61$.

2. A copy machine dealer has data on the number x of copy machines at each of 89 customer locations and the number y of service calls in a month at each location. Summary calculations give $\bar{x} = 8.4$, $s_x = 2.1$, $\bar{y} = 14.2$, $s_y = 3.8$, and $r = 0.86$. What is the slope of the least-squares regression line of number of service calls on number of copiers?
(a) 0.86
(b) 1.56
(c) 0.48
(d) None of these
(e) Can't tell from the information given

3. In the setting of the previous problem, about what percent of the variation in the number of service calls is explained by the linear relation between number of service calls and number of machines?
(a) 86%
(b) 93%
(c) 74%
(d) None of these
(e) Can't tell from the information given

4. If data set A of (x, y) data has correlation coefficient $r = 0.65$, and a second data set B has correlation $r = -0.65$, then
(a) the points in A exhibit a stronger linear association than B.
(b) the points in B exhibit a stronger linear association than A.
(c) neither A nor B has a stronger linear association.
(d) you can't tell which data set has a stronger linear association without seeing the data or seeing the scatterplots.
(e) a mistake has been made—r cannot be negative.

5. There is a linear relationship between the number of chirps made by the striped ground cricket and the air temperature. A least-squares fit of some data collected by a biologist gives the model $\hat{y} = 25.2 + 3.3x$, $9 < x < 25$, where x is the number of chirps per minute and $\hat{y}$ is the estimated temperature in degrees Fahrenheit. What is the estimated increase in temperature that corresponds to an increase of 5 chirps per minute?
(a) 3.3°F (b) 16.5°F (c) 25.2°F (d) 28.5°F (e) 41.7°F

6. Which of the following relationships is most likely to result in a strong negative correlation?
(a) The number of people showering in a college dorm and the water pressure in each shower.
(b) The outdoor temperature and the number of fans running in non-air-conditioned dorm rooms.
(c) The comfort rating of a mattress and the number of hours of uninterrupted sleep obtained.
(d) The price of a home and its square footage.
(e) The fuel efficiency of a car (miles per gallon) and its speed.

7. A set of data relates the amount of annual salary raise and the performance rating. The least squares regression equation is $\hat{y} = 1400 + 2000x$ where y is the raise amount and x is the performance rating. Which of statements (a) to (d) is *not* correct?
(a) For each increase of one point in performance rating, the raise will increase on average by $2000.
(b) This equation produces predicted raises with an average error of 0.
(c) A rating of 0 will yield a predicted raise of $1400.
(d) The correlation between salary raise and performance rating is positive.
(e) All of the above are true.

8. Leonardo da Vinci, the renowned painter, speculated that an ideal human would have an armspan (distance from outstretched fingertip of left hand to outstretched fingertip of right hand) that was equal to his height. The following computer regression printout shows the results of a least-squares regression on height and armspan, in inches, for a sample of 18 high school students.

Dependent variable is: **Height**
No Selector
R squared = 87.1% R squared (adjusted) = 86.3%
s = 1.613 with 18 - 2 = 16 degrees of freedom

Source	Sum of Squares	df	Mean Square	F-ratio
Regression	280.631	1	280.631	108
Residual	41.6185	16	2.60116	

Variable	Coefficient	s.e. of Coeff	t-ratio	prob
Constant	11.5474	5.6	2.06	0.0558
Armspan	0.840424	0.08091	10.4	≤ 0.0001

Which of the following statements is *false*?
(a) This least-squares regression model would make a prediction that is 1.63 inches higher than da Vinci projected for a 62-inch tall student.
(b) One of the students in the sample had a height of 70.5 inches and an armspan of 68 inches. The residual for this student is 1.83 inches.
(c) Da Vinci's projection is lower than the prediction that this least-squares line will make for any height.
(d) For every one-inch increase in armspan, the regression model predicts about a 0.84-inch increase in height.
(e) For a student 66 inches tall, our model would predict an armspan of about 67 inches.

Part 2. Free Response

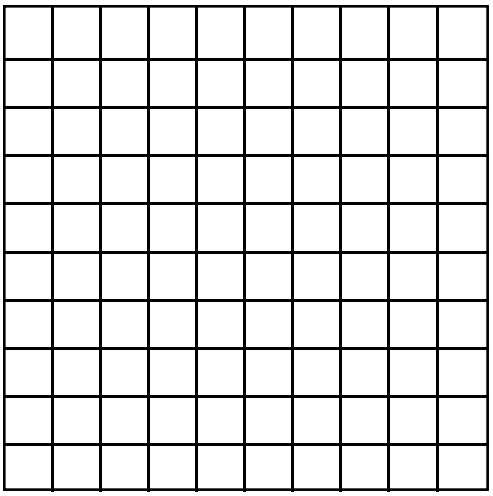
Answer completely, but be concise. Show your work.

Joey appears to be growing slowly as a toddler. His height between 18 and 30 months of age increases as follows:

Age (months)	Observed height (cm)	Predicted height	Residual
18	76.5		-0.08
21	78.7	79.09	
24	82.0	81.6	0.4
27	84.8	84.11	
30	86.0		-0.62

The least-squares regression line fitted to this data has equation

HEIGHT = 61.5 + 0.837 AGE



9. Finish filling in the table above.

10. Sketch a residual plot on the axes provided.

11. Based on your residual plot, would you describe Joey's growth pattern from 18 to 30 months as linear? Explain.

12. According to the least-squares principle, which of the lines below provides the best fit for the data shown in the scatterplot? Justify your answer.

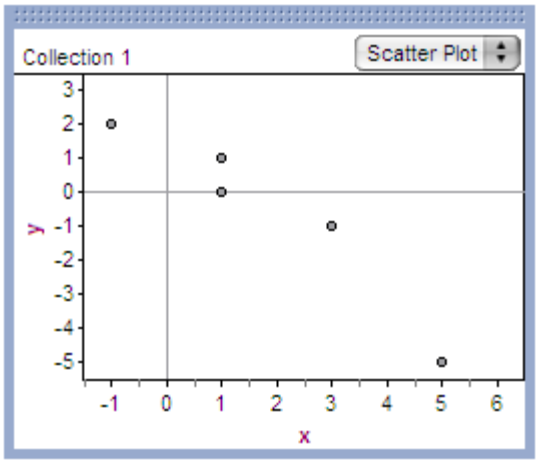
(a) $y = 2 - x$

(b) $y = 1.5 - x$

(c) $y = 1 - x$

(d) $y = 3 - 2x$

(e) $y = 3 - 1.5x$



13. Anthropologists must often estimate from human remains how tall the person was when alive. Carla is studying how overall height can be predicted from the length of a leg bone in a group of 36 living males. The data show that the bone lengths have mean 45.9 cm and standard deviation 4.2 cm, the overall heights have mean 172.7 cm and standard deviation 8.14 cm, and the correlation between bone length and height is 0.914.

(a) Determine the equation of the least-squares regression line of height on bone length. Show your work.

(b) Interpret the correlation in the context of this problem.

14. In general, is correlation a resistant measure of association? _____ Explain briefly or give a simple example to illustrate.

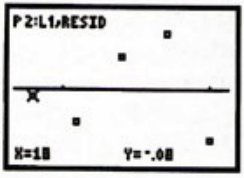
I pledge that I have neither given nor received aid on this test. _____

Test 3C

1. (d) 2. (b) 3. (c) 4. (c) 5. (b) 6. (a) 7. (e) 8. (c)

9.

Age (months)	Observed height (cm)	Predicted height	Residual
18	76.5	76.58	-0.08
21	78.7	79.09	-0.39
24	82.0	81.6	0.4
27	84.8	84.11	0.69
30	86.0	86.62	-0.62



10. The residual plot looks like this:

11. Joey's growth pattern is linear because there is no obvious pattern in the residual plot. 12. The sum of squared residuals for each of the proposed models is (a) 6, (b) 3.25, (c) 3, (d) 18, (e) 9.25. So choice (c), $y = 1 - x$, would provide the best fit.

13. (a) $b = r\frac{s_y}{s_x} = 0.914\frac{8.14}{4.2} = 1.771$; $172.7 = a + (1.771)(45.9) \Rightarrow a = 91.39$. The equation of the least-squares regression line is predicted height = 91.39 + 1.771(bone length). (b) There is a strong, positive, linear relationship between bone length and height. 14. No. Consider a scatterplot with five points in the lower-left corner that fall roughly along a positively sloped line. In this case, r is close to 1. Now add a point far to the right of the other five points, but at a height lower than the lowest of the 5 original points. The correlation will decrease noticeably and perhaps even become negative.