

Directions: *Work on these sheets.*

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

- A study found correlation $r = 0.61$ between the sex of a worker and his or her income. You conclude that
 - Women earn more than men on the average.
 - Women earn less than men on average.
 - An arithmetic mistake was made; this is not a possible value of r .
 - This is nonsense because r makes no sense here.
 - There is a moderately strong, positive, linear relationship between sex and income.
- 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?
 - 0.86
 - 1.56
 - 0.48
 - None of these
 - Can't tell from the information given
- 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?
 - 86%
 - 93%
 - 74%
 - None of these
 - Can't tell from the information given
- 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
 - The points in A exhibit a stronger linear association than B.
 - The points in B exhibit a stronger linear association than A.
 - Neither A nor B has a stronger linear association.
 - You can't tell which data set has a stronger linear association without seeing the data or seeing the scatterplots.
 - A and B are actually the same data set.
- 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 give 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 in 5 chirps per minute?
 - 3.3°F
 - 16.5°F
 - 25.2°F
 - 28.5°F
 - 41.7°F

6. The equation of the least squares regression line for a certain set of data is $\hat{y} = 1.3 + 0.73x$. What is the residual for the point (4,7)?
- 2.78
 - 3.00
 - 4.00
 - 4.22
 - 7.00
7. Linear regression usually employs the method of least squares. Which of the following is the quantity that is minimized by the least squares process?
- \hat{y}_i
 - $x_i - \bar{x}_i$
 - $\sum (y_i - \bar{y}_i)^2$
 - (\bar{x}_i, \bar{y}_i)
 - $\sum (x_i - \bar{x}_i)^2$
8. A set of data relates the amount of annual salary raise and the performance rating. The least squares regression equation is $\hat{y} = 1,400 + 2,000x$ where y is the estimated raise and x is the performance rating. Which of the following statements is *not* correct?
- For each increase of one point in performance rating, the raise will increase on average by \$2,000.
 - This equation produces predicted raises with an average error of 0.
 - A rating of 0 will yield a predicted raise of \$1,400.
 - The correlation for the data is positive.
 - All of the above are true.
9. Which of the following would *not* be a correct interpretation of a correlation of $r = -.30$?
- The variables are inversely related.
 - The coefficient of determination is 0.09.
 - 30% of the variation between the variables is linear.
 - There exists a weak relationship between the variables.
 - All of the above statements are correct.
10. The following are resistant:
- Least squares regression line
 - Correlation coefficient
 - Both the least squares line and the correlation coefficient
 - Neither the least squares line nor the correlation coefficient
 - It depends

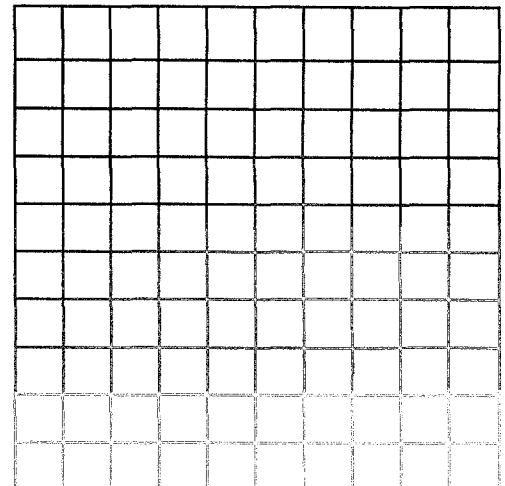
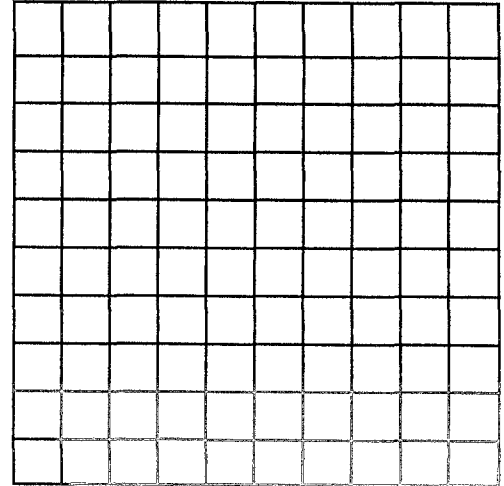
Part 2. Free Response

Answer completely, but be concise. Show your work.

11. Given the following observations of quantitative variables X and Y:

X	0	1	2	3	15
Y	3	4	6	10	0

- (a) Make a scatterplot of the data on the axes. Circle the most influential observation.
- (b) Determine the LSRL of Y on X. Draw this line carefully on your scatterplot. Indicate how you plotted the line.
- (c) What is a regression outlier?
- (d) Which data point is the biggest regression outlier?
- (e) What is the residual at the point identified in part (c)?
- (f) Construct a residual plot for these data.
- (g) Interpret your residual plot, i.e., what does the residual plot tell you?



12. 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 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) What is the slope of the LSRL of height on bone length?

(b) Calculate and interpret the r^2 -value.

(c) Based on your answer to (b), how would you describe the goodness of linear fit?

(d) Determine the equation of the LSRL of height on bone length.

Extra Credit: Suppose that x and y are both standardized variables (i.e., with means 0 and standard deviations 1). Determine the equation of the least squares regression line.

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