

Name: \_\_\_\_\_ Date: \_\_\_\_\_ Class: \_\_\_\_\_

Key

Statistical Reasoning - Collecting and Analyzing Bivariate Data Test Review

1. The regression line for the two variable data set given below is  $\hat{y} = 2.35 + 0.86x$ . What is the value of the residual for the point whose x value is 32?

x	y
23	19
15	18
26	22
24	20
22	27
29	25
32	32
40	38
41	35
46	45

$y = 2.35 + 0.86(32) = 29.87$   
Residual =  $32 - 29.87 = 2.13$

2. A least-squares regression line for predicting performance on a college entrance exam based on high school grade point average (GPA) is determined to be  $\text{Score} = 273.5 + 91.2(\text{GPA})$ . One student in the study had a high school GPA of 3.0 and an exam score of 510. What is the residual for this student?  
 $y = 273.5 + 91.2(3) = 547.1$   
Residual =  $510 - 547.1 = -37.1$

3. Suppose the (SRL) for predicting Weight (in pounds) from Height (in inches) is given by Weight =  $115 + 3.6(\text{Height})$ . Which of the following statements is correct?
- (a) I only
  - (b) II only
  - (c) III only
  - (d) II and III only
  - (e) I and III only
- II: For each additional inch of Height, Weight will increase on average by 3.6 pounds.  
III: There is a strong positive linear relationship between Height and Weight.



5. Classify the type of strength and direction of each of the correlation coefficients.
- |               |     |          |
|---------------|-----|----------|
| $r = 0.9994$  | Pos | Strong   |
| $r = 0.7342$  | Pos | Moderate |
| $r = -0.3332$ | Neg | Weak     |
| $r = -0.7992$ | Neg | Moderate |

Hybart 11

6. The following data has been collected regarding hours studied and grades on a final Exam:

Hours	Grade
1	61
2	68
3	77
4	82
5	85
6	90
7	94

- a. Make a graph of the data and evaluate the residual plot.  
b. Use your calculator to determine the regression formula and correlation coefficient. (Round to 3 decimals)  
Regression equation:  $y = 5.39x + 58$   $r = 0.99$   
c. Is a linear model a good fit for this data? Why? Discuss r value (correlation) and Residual Plot.  
 $y = 5.39x + 58$   $r = 0.99$   
d. Interpret the regression characteristics.

Slope: Linear / Non-linear  
Direction: Positive / Negative  
Strength: Strong / Moderate / Weak

- e. What grade would you expect someone to get if they studied for 9 hours?  
 $y = 5.39(9) + 58 = 106.51$   
Do you think this is a good prediction? Justify your response.  
 $y = 106.51$  gas, r is strong (only if grade can be over 100)  
f. Approximately how many hours would someone need to study if they would like to make a 100?  
 $100 = 5.39x + 58$   $x = 7.79$  hours  
Do you think this is a good prediction? Justify your response.  
Also

(Question 7 - 9) A study gathers data on the outside temperature during the winter, in degrees Fahrenheit, and the amount of natural gas a household consumes, in cubic feet per day. Call the temperature x and gas consumption y. The house is heated with gas, so x helps explain y. The least-squares regression line is  $y = 1344 - 19x$ .

7. On a day when the temperature is 45° F, the regression line predicts that gas used will be about  
 $y = 1344 - 19(45) = 409 \text{ ft}^3$
8. When the temperature goes up 1 degree, what happens to the gas usage predicted by the regression line?  
Gas use goes down 19 ft<sup>3</sup>
9. The correlation between temperature x and gas usage y is  $r = -0.7$ . What is the effect of changing Celsius to Fahrenheit on the correlation?  
None

Hybart 12

10. All 733 students in grades 1 through 6 in an elementary school are given a math test that was designed for third graders. The body weights of all 733 students are also recorded. What kind of association can we expect to see between the weight of a student and their test score?

NONE

11. The United Nations has data on the percent of adult males and females who are illiterate in each of these 142 countries. The correlation between male illiteracy rate and female illiteracy rate is  $r = 0.945$ . What does this  $r$  value tell us?

There is a strong, positive relationship b/w male and female illiteracy rates.

12. The regression line for the data given above is  $y = 2.35 + 0.86x$ . What is the residual for the point whose  $x$  value is 158?

x	y
23	15
19	18
26	22
24	20
22	27
29	32
40	38
41	35
46	45

$y = 2.35 + 0.86(158) = 15.25$        $18 - 15.25 = 2.75$

13. A study of how much heavy Mr. H's book bag weighs over the course of a semester is below.

Time (months)	1	2	3	4	5	6	7	8	9	10	11	12
Mean Wt. (kg)	4.3	4.9	5.7	6.6	7.5	8.8	10.1	13	15	16.3	19	21.1

- a. Make a graph of the data and evaluate the residual plot.

- b. Use your calculator to determine the regression formula and correlation coefficient. (Round to 3 decimals)

Regression equation:  $y = 1.56x + 0.90$        $r = 0.98$

- c. Is a linear model a good fit for this data? Why? Discuss  $r$  value (correlation) and Residual Plot.

Yes -  $r$  is strong

- d. Interpret the regression characteristics.

Shape: Linear / Non-linear  
 Direction: Positive / Negative  
 Strength: Strong / Moderate / Weak

- e. What would be the mean weight at 15 months? 15.25 weeks  
 Do you think this is a good prediction? Justify your response.

- f. Predict how long it would take for his bag to weigh 50 pounds.  
 $50 = 1.56x + 0.90$        $x = 31.47$  weeks  
 Do you think this is a good prediction? Justify your response.

14. At summer camp, one of Carol's counselors told her that you can determine air temperature from the number of cricket chirps.

a) What is the response variable?

Explanatory variable: air temp      Response variable: # chirps

15. A value of  $r = .40$  indicates there is a weak relationship between  $x$  and  $y$ .

16. The difference between an observed value of  $y$  and a predicted value of  $y$  for the same  $x$  value is called a residual.

17. The following data has been collected regarding time and the temperature of a cup of coffee as it sits on a desk.

- a. Make a graph of the data and evaluate the residual plot.

- b. Use your calculator to determine the regression formula and correlation coefficient. (Round to 3 decimals)

Regression equation:  $y = -1.51x + 168.86$        $r = -0.95$

- c. Is a linear model a good fit for this data? Why? Discuss  $r$  value (correlation) and Residual Plot.

- d. Interpret the regression characteristics.

Shape: Linear / Non-linear  
 Direction: Positive / Negative  
 Strength: Strong / Moderate / Weak

- e. How hot would you expect the coffee to be after 60 minutes?  
 $y = -1.51(60) + 168.86 = 78.86$   
 Do you think this is a good prediction? Justify your response.

Time (mins)	Temp (° F)
0	179.5
5	168.7
8	158.1
11	149.2
15	141.7
18	134.6
22	125.4
25	123.5
30	116.3
34	113.2
38	109.1
42	105.7
45	102.2
50	100.5

- f. If room temperature is currently 72°, approximately how long would the coffee need to sit on the desk to reach room temperature?  
 $72 = -1.51x + 168.86$        $x = 64.15$  mins  
 Do you think this is a good prediction? Justify your response.