Name KEY

I. Identify which variable is the response, guess whether the relationship is positive (+) or negative (–).

1. A study has shown then number of fatal highway accidents (A) that occur on highways is related to the mean speed of traffic (S) on that highway.

response: A

+ or –? +

2. An psychology experiment showed that the time it takes elementary students to complete a puzzle (T) is related to the student's score on a mathematics test (M).

response: T

+ or –? –

3. A study found that counties in the US with higher median income levels (I) have higher rates of breast cancer diagnoses (C).

response: C

+ or –? +

II. For bivariate data, match one of the following with the description below.

a. scatterplot b. correlation coefficient c. regression equation

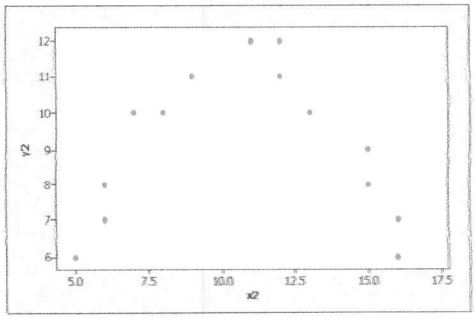
b 1. A measure of linear association.

a 2. What we use to try to determine if the two variables are related and how.

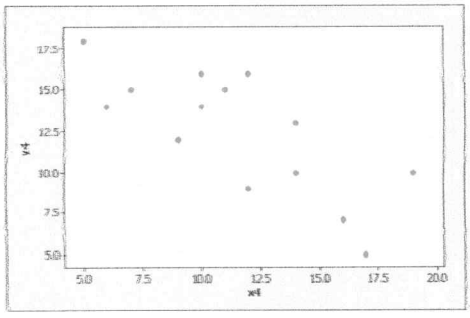
c 3. The mathematical model we use to predict a value of the response.

III. Match the scatterplot to the correlation coefficient, the calculated correlations are approximately a) 0.76, b) 0.04, c) –0.08, d) –0.58

1. b



2. d



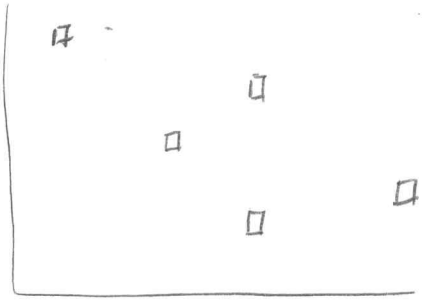
Name KEY

IV Given the following set of bivariate data: (3, 22), (5, 17), (6, 18), (6, 12), (8, 13)

1. Enter the data into your calculator.

2. Create the scatterplot on your calculator.

more or less



3. Find r, the correlation coefficient, and the least-squares regression equation and store the regression equation in Y_1.

OUTPUT:

$a = 26.6666\ldots$

$b = -1.8333\ldots$

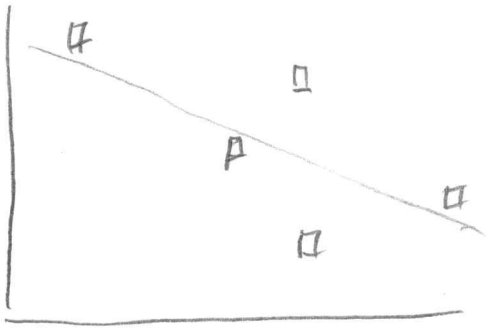
$r^2 = .6804703476$

$r = -.8249406266$

so $y = 26.67 - 1.83x$

$r = -.82$

4. Display the graph with the regression line. [GRAPH]



5. Compute the predicted value of y for $x = 5$. [Y1(5)]

$Y1(5) = 17.5$

residual for (5, 17) is $17 - 17.5 = -.5$

Name KEY

V. A large study of statistics students in U.S. colleges and universities found a strong negative association between the number of absences a student had during the term and their final score. The following regression equation relates the number of absences (A) to their final score (S, in percent of total points possible), $\hat{S} = 84.6 - 1.9A$ and $R^2 = 82\%$. For this regression:

1. Identify the

   response variable S (score)

   explanatory variable A (absences)

2. give the slope (*include units*) −1.9 %/absence

3. interpret the slope *with respect to this situation and this regression* (your response should have to do the number of absences and the student's grade)

   For each addition absence a student's score in statistics falls, on average, by 1.9%.

4. give the y-intercept (include units) 84.6%

5. Interpret (if possible) the y-intercept *with respect to this situation and this regression*

   The expected mean score for a student with no absences is 84.6%.

6. Find the correlation coefficient, r $r = -\sqrt{.82} \approx -.9055$ −.91

7. Find the predicted final score for a student who had 10 absences (A = 10).

   $\hat{S}(10) = 84.6 - 1.9(10)$

   $84.6 - 19 = 65.6$ 65.6

8. Donald had 10 absences and his final score for the course was 61.4%, what was the residual for this student's score.

   $61.4 - 65.6$

   −4.2%

VI. (Just one example) Based on EPA data for 2017 passenger vehicles, gasoline mileage (M in mpg) is related to weight of the car, W in thousands of pounds – *K-lbs.*) by the regression equation: $M = 44.4 - 7.2W$. The predicted gas mileage of a passenger vehicle that weights 6500 lbs is –2.4 mpg, which of course is impossible. What went wrong in making this prediction with our regression equation?

This is an example of extrapolation, 6500 lbs (6.5 K lbs) is out of the viable range of W.