Ch 1 Quiz 1 Statistics Name:

1. In 1997, there were 12,298,000 undergraduate students in U.S. colleges. According to the U.S. Department of Education, there were 127,000 American Indian or Alaskan Native, 737,000 Asian or Pacific Islander, 1,380,000 non-Hispanic black, 1,108,000 Hispanic, and 8,682,000 non-Hispanic white students. In addition, 265,000 foreign undergraduates were enrolled in U.S. colleges.
   (a) Each number, including the total, is rounded to the nearest thousand. Separate rounding may cause roundoff errors, so that the sum of the counts does not equal the total given. Are roundoff errors present in these data? Explain.

   (b) Present the data in a bar graph (Label axes!!!).

2. A Richmond television station used a questionnaire to gather data on the ages of viewers of *ACTION*, a program aimed at a young audience. Here is a stemplot of the data:

```
0   566899
1   0012223445667789
2   00123346
3   0135
4   28
5   025
6   05
7
8
9
10
11
12  0
```

...2.

(a) Represent the data in a histogram using intervals of 10 (i.e. 0-9, 10-19, etc.).

(b) Describe the shape, center, and spread of the distribution. Comment on any other feature of the plot that needs to be addressed.

3. What would be the best type of graph to use in order to display a house's market value over a span of 10 years? Explain your answer and state how the graph would be labeled.

Ch 1 Quiz 2 Statistics Name:

Use data sets A and B below for questions 1-4.

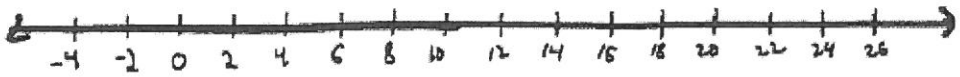
A: 5, 7, 9, 11, 13, 15, 17 B: 5, 5, 5, 11, 17, 17, 17

1) Find the five number summaries for each distribution.

(A): Min =
Q1 =
Med =
Q3 =
Max =

(B): Min =
Q1 =
Med =
Q3 =
Max =

2) Make box-and-whisker plots for both distributions under the same number line below.



(A)

(B)

3) Find the mean and standard deviation for both distributions A and B.

(A): $\bar{X}$ =
s =

(B): $\bar{X}$ =
s =

4) Compare the shape, center, and spread of the two distributions.

For problems 5 and 6, use the distribution of IQ scores shown below.

107, 110, 115, 119, 125, 132, 135, 140, 150, 175

5) Find the Interquartile Range (IQR) of the distribution of IQ scores.

IQR =

6) Indicate any and all outliers in the distribution of IQ scores. If there are no outliers, say so.

True or False

_____ 7) The mean of a distribution is a measure of spread, or dispersion.

_____ 8) The standard deviation of a distribution is a measure of spread, or dispersion.

_____ 9) Mrs. Butterworth decides to curve her students' test scores by adding 10 points to each score. The mean (or average) score will not be changed by the curve.

_____ 10) The CEO of a PricewaterhouseCoopers firm decides to give each employee in the firm a 5% raise. Both the mean and the standard deviation will increase by 5%.

Ch 1 Test Statistics Name: _____

True or False

_____ 1) In a distribution, any value greater than 1.5 x IQR is considered to be an outlier.

_____ 2) The mean (or average) is resistant to any outliers.

_____ 3) Mrs. Butterworth's cooking class averages a 77 on a quiz. She decides to add 3 points to each student's grade. After doing so, the class average will still be a 77.

_____ 4) Given a skewed distribution with extreme outliers, the most appropriate measure of center and spread to use would be the mean and standard deviation.

_____ 5) As the spread of a distribution increases, the standard deviation also increases.

_____ 6) A bar graph displays a categorical variable, while a histogram display only quantitative variables.

Problems

Use the distribution of yards per passing play from a football game shown below for problems 7-11.

0, 5, 63, 0, 0, 0, 0, 0, 30, 19, 0, 17

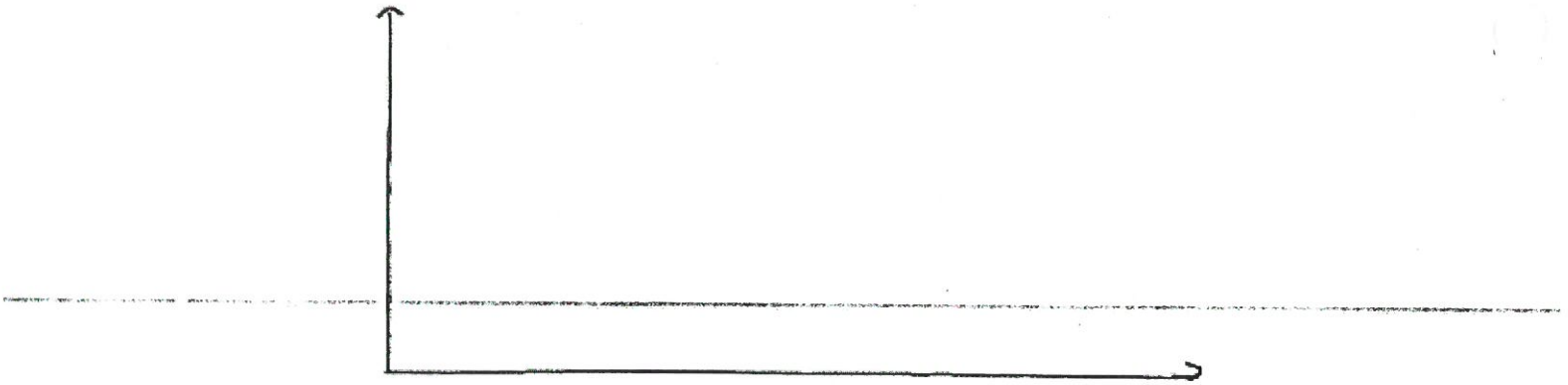
7) Make a stem-and-leaf plot of the distribution of yards per passing play. Then find the 5 number summary.

Min =
Q_1 =
Med =
Q_3 =
Max

IQR =

8) Indicate any and all outliers in the distribution. If there are no outliers, say so.

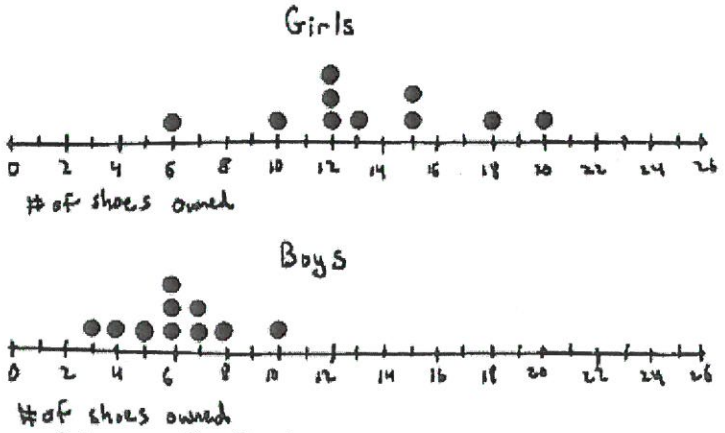
9) Make a histogram of the distribution using intervals of 10 (0-9, 10-19, etc.). Make sure to use labels!



10) Describe the shape of the distribution of yards per pass play.

11) What measures of center and spread would be the most appropriate for this distribution? Find them.

The stacked dotplots below represent the number of shoes owned by 10 girls and 10 boys. Use the dotplots to answer problems 12-14.



12) Compare the shapes of the two distributions.

13) Compare the centers (typical values) of the two distributions. Use appropriate statistical measures to support your comparison.

14) Compare the spreads of the two distributions. Use appropriate statistical measures to support your comparison.

Ch 2 Quiz 1 Statistics Name:

True or False

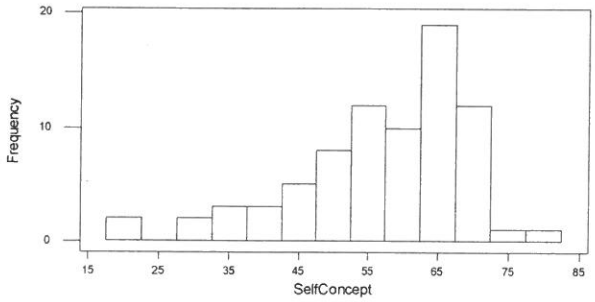
_____1) In a perfectly symmetric density curve, the mean equals the median.

_____2) A z-score of 2.5 means that the value is 2.5 standard deviations below average.

_____ 3) A value at the 72nd percentile is higher or equal to 72% of other values in a distribution.

Problems

A group of 78 third-grade students in a midwestern elementary school took a "self-concept" test that measured how well they felt about themselves. Higher scores indicate more positive self-concepts. A histogram and some summary statistics from Minitab for these students' self-concept scores are given below.



Descriptive Statistics: SelfConcept

Variable	N	Mean	Median	StDev
SelfConc	78	56.96	59.50	12.41

Variable	Minimum	Maximum	Q1	Q3
SelfConc	20.00	80.00	51.00	66.25

4) Draw an appropriate density curve for summarizing the histogram on the graph above. How would you describe the <u>shape</u> of this density curve?

5) One student scored a 75 on the Self-Concept Test. Describe the location of this value within the distribution by finding the standardized value (z-score).

6) One student in the group had a self-concept score of 62.

a) Find the standardized value (z-score).

b) Thirty-four students scored higher than 62. At what percentile did the student score?

The stemplot below represents the heights (in inches) of 25 students on a soccer team. Each leaf =1 in.

6	3 4 4
6	5 6 6 7 7 7 8 8 8 8 8 9
7	0 0 1 1 1 2 2 4
7	5 6

7) Ben is on the soccer team and is 6'2 (or 74 inches) tall. Ben's height is at what percentile of the soccer team?

8) Jack is on the soccer team and is 5'6 (or 66 inches) tall. Jack's height is at what percentile of the soccer team?

9) Find the height of the soccer player at the 60^{th} percentile on the soccer team.

Ch 2 Quiz 2 Statistics Name:

True or False

(1 point each)

_____1) The mean of a *standard* Normal distribution is always zero.

_____2) In a Normal distribution, 100% of the data is within 1 standard deviation of the mean.

_____ 3) The Normal curve is only used to model distributions that are skewed.

Problems

4) Use Table A to find the proportion of observations from a standard Normal distribution that satisfies $-1.51 < Z < 0.84$. Sketch the Normal curve and shade the area under the curve that is the answer to the question. (2 points)

5) The weights of babies born at a certain hospital average 129 oz, with a standard deviation f 12 oz. Assuming that the weights are normally distributed, find the percent of babies with weights that are: (Show your work and sketch the Normal curve and shade the specified area under the curve for each.)

a) More than 141 oz? (2 points)

b) Less than 105 oz? (2 points)

c) Between 117 oz and 141 oz? (2 points)

6) Use the information provided in problem 5 to determine the weight for which the smallest 20% of babies born is less than. (1 point)

BONUS… Give an example of a quantitative variable that does *not* have a Normal distribution. Justify your answer. (1 point)

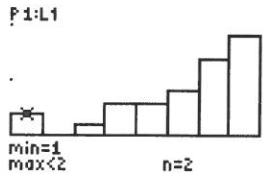
Ch 2 Test Statistics Name:

Multiple Choice

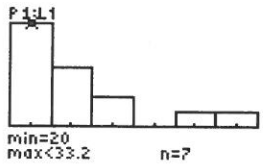
(1 point each)

_____ 1) Which of the following histograms would be best approximated by a Normal distribution?

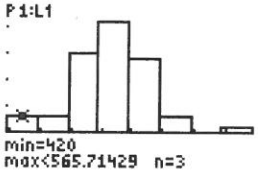
(a)



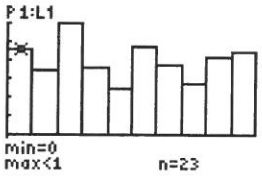
(b)



(c)



(d)



(e) All of (a) through (d)

_____ 2) Many professional schools require applicants to take a standardized test. Suppose that 1000 students take the test, and you find that your mark of 63 (out of 100) is the 73rd percentile.

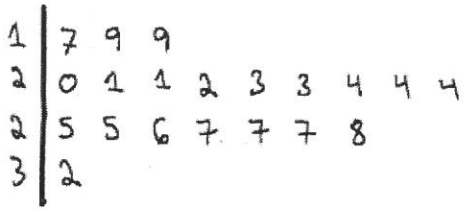
This means that
(a) at least 73% of the people scored 63 or better.
(b) at least 270 people scored 73 or better.
(c) at least 730 people scored 73 or better.
(d) at least 27% of the people scored 73 or worse.
(e) at least 73% of the people scored 63 or worse.

_____ 3) The area under the standard Normal curve corresponding to $-0.3 < Z < 1.6$ is
(a) 0.3273
(b) 0.4713
(c) 0.5631
(d) 0.9542
(e) none of the above

_____ 4) In a standard Normal distribution, the mean always equals
(a) 1
(b) 0
(c) .5
(d) 100
(e) none of the above

Problems

Use the stemplot to answer questions 5-7. The stemplot below represents the vertical jumps (in inches) of 20 students on a basketball team. Each leaf =1 in



5) Michael is on the basketball team and has a vertical of 26 inches. Michael's vertical is at what percentile of the basketball team? (2 points)

6) Jimmy is on the basketball team and has a vertical of 22 inches. Jimmy's vertical is at what percentile of the basketball team? (2 points)

7) Find the vertical jump of the basketball player at the 65^{th} percentile on the team. (1 point)

8) The heights of American men aged 18 to 24 are approximately Normally distributed with mean 68 inches and standard deviation 2.5 inches. Find the percent of heights described below.
(Show your work and sketch the Normal curve and shade the specified area under the curve for each.)

a) Less than 60 in? (2 points)

b) More than 72 in? (2 points)

c) Between 66 and 72? (2 points)

9) Use the information in problem 8 to determine how short a guy must be in order to fall in the shortest 10% of American men between the ages of 18 – 24? (*Bonus 2 points*)

Ch 3 Quiz 1 Statistics Name:

Matching

Write the letter of the graph that best matches the given correlation (r) below.

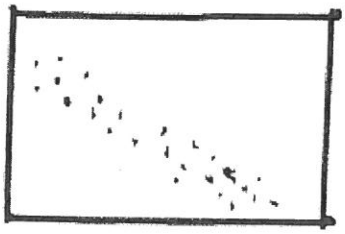
_____1) r = –0.254

_____2) r = –0.878

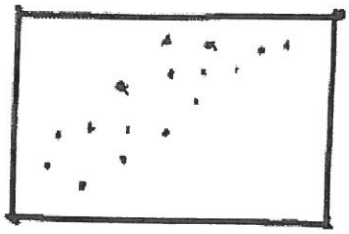
_____ 3) r = 0.702

_____ 4) r = 0.321

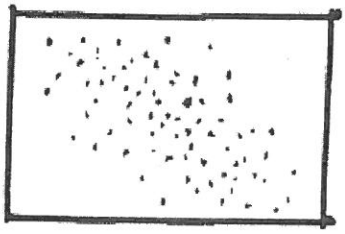
A



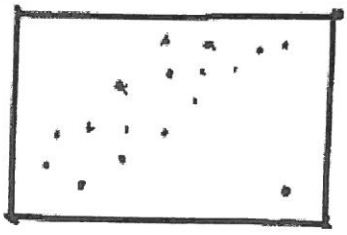
B



C

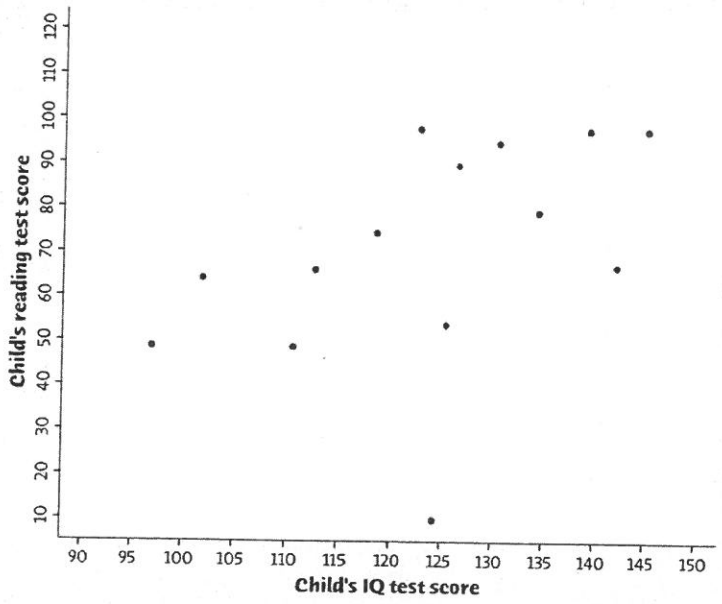


D



Problems

Here is a scatterplot of reading test scores against IQ test scores for 14 fifth-grade children. Refer to this scatterplot for problems 5-10.



5) What is the explanatory variable in the scatter plot?

6) What is the response variable in the scatterplot?

7) Does the scatterplot show a positive association, negative association, or neither? Explain.

8) Which of the following is the closest to the correlation between reading test scores and IQ score for the group of 15 children? **(Circle the best choice)**

$r = 0.9$ $r = 0.4$ $r = -0.9$ $r = -0.4$

9) Does there appear to be any "unusual points" on the graph? If so, circle it on the scatterplot.

10) What effect would removing the "unusual point" have on the correlation? Justify your answer.

A student wonders if tall women tend to date taller men than do short women. She measures herself, her dormitory roommate, and the women in the adjoining rooms; then she measures the next man each woman dates. Here are the data (heights in inches):

Women	66	64	66	65	70	65
Men	72	68	70	68	71	65

11) In the problem above what is the explanatory variable?

12) Make a well-labeled scatterplot of these data.

13) Based on the scatterplot, do you expect the correlation to be positive or negative? Near ± 1 or not? Explain.

14) Use your calculator to find the correlation r between the heights of the men and women.

15) Do the data show that taller women tend to date taller men? Explain.

16) Thanksgiving is just around the corner. What is something for which you are thankful?

Ch 3 Quiz 2 Statistics Name:

___1) In a statistics course, a linear regression equation was computed to predict a student's final exam score from his/her score on the first test. The equation of the least-squares regression line was

$$\hat{y} = 10 + 0.9x$$

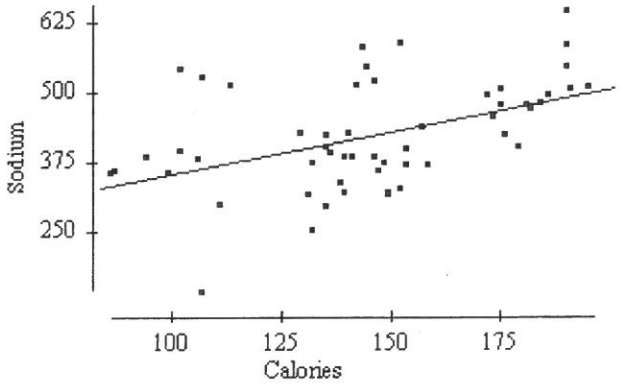
where y represents the final exam score and x is the score on the first exam. Suppose Joe scores a 90 on the first exam. What would be the predicted value of his score on the final exam?

A) 91.
B) 90.
C) 89.
D) 81.
E) It cannot be determined from the information given. We also need to know the correlation coefficient.

___2) The fraction of the variation in the values of y that is explained by the least-squares regression of y on x is

A) the correlation coefficient.
B) the slope of the least-squares regression line.
C) the coefficient of determination (which is the square of the correlation coefficient).
D) the intercept of the least-squares regression line.
E) the residual.

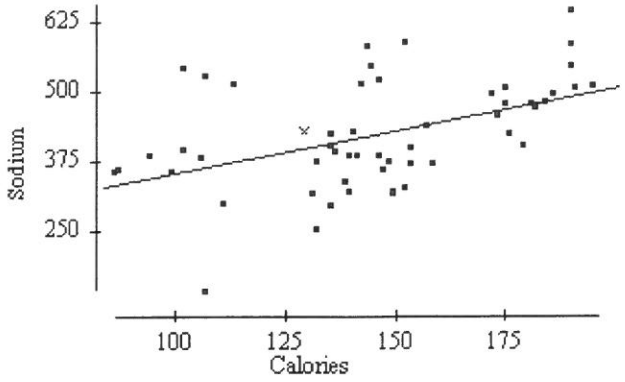
___3) Below is a scatterplot of the calories and sodium content (in milligrams) of several brands of meat hot dogs. The least-squares regression line has been drawn on the plot.



Based on the least-squares regression line in this scatterplot, one would predict that a hot dog containing 100 calories would have a sodium content (in milligrams) of about

A) 70. B) 350. C) 375. D) 400. E) 600.

___4) A scatterplot of the calories and sodium content of several brands of meat hot dogs is shown below. The least-squares regression line has been drawn on the plot.



Referring to this scatterplot, the value of the residual for the point labeled x

A) is about 40.
B) is about 125.
C) is about 425.
D) is about 1300.
E) cannot be determined from the information given.

___5) A least-squares regression line is fitted to a set of data. If one of the data points has a positive residual, then

A) the correlation between the values of the response and explanatory variables must be positive.
B) the point must lie above the least-squares regression line.
C) the slope of the least-squares regression line must be positive.
D) the point must lie near the right edge of the scatterplot.
E) all of the above.

Ch 3 Test Statistics Name:

___1) In a statistics course, a linear regression equation was computed to predict a student's final exam score from his/her score on the first test. The equation of the least-squares regression line was

$$\hat{y} = 10 + 0.9x$$

where y represents the final exam score and x is the score on the first exam. Suppose Sally scores a 80 on the first exam. What would be the predicted value of her score on the final exam?

A) 80.
B) 81.
C) 79.
D) 82.
E) It cannot be determined from the information given. We also need to know the correlation coefficient.

____2) Refer to the previous problem. On the final exam Sally actually scored 88. What is the value of her residual?

(a) –6
(b) 0
(c) 6
(d) 8.2

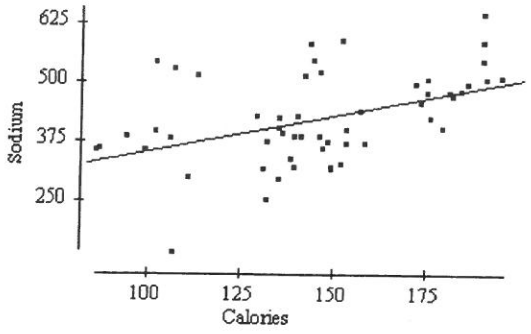
___3) If 75% of the variation in the values of y is explained by the least-squares regression of y on x, then the value of the correlation, r, is

A) .87
B) .75
C) 75
D) .56

____4) You have data for many families on the parents' income and the years of education their eldest child completes. When you make a scatterplot,

(a) the explanatory variable is parents' income, and you expect to see a negative association.
(b) the explanatory variable is parents' income, and you expect to see a positive association.
(c) the explanatory variable is parents' income, and you expect to see very little association.
(d) the explanatory variable is years of education, and you expect to see a negative association.
(e) the explanatory variable is years of education, and you expect to see a positive association.

___5) Below is a scatterplot of the calories and sodium content (in milligrams) of several brands of meat hot dogs. The least-squares regression line has been drawn on the plot.



Based on the least-squares regression line in this scatterplot, one would predict that a hot dog containing 150 calories would have a sodium content (in milligrams) of about
A) 270. B) 350. C) 375. D) 425. E) 600.

___6) A community college announces that the correlation between college entrance exam grades and scholastic achievement was found to be –1.08. On the basis of this you would tell the college that
(a) the entrance exam is a good predictor of success.
(b) the exam is a poor predictor of success.
(c) students who do best on this exam will be poor students.
(d) students at this school are underachieving.
(e) the college should hire a new statistician.

A certain psychologist counsels people who are getting divorced. A random sample of five of her patients provided the following data where x = number of years of courtship before marriage, and y = number of years of marriage before divorce.

x	3	0.5	2	1.5	5
y	9	6	14	10	20

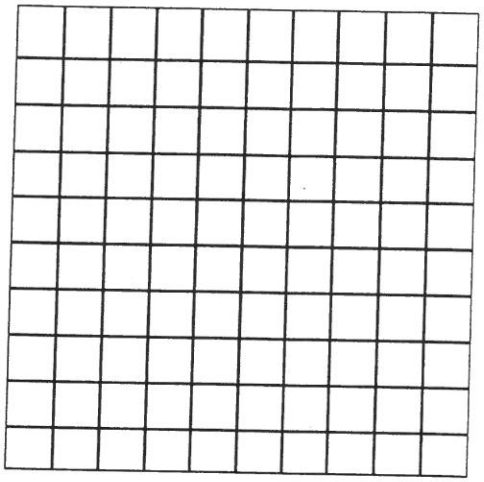
7) Construct a scatterplot on the grid provided (Label axes):

8) Describe the form, direction, and strength of the relationship.

Form:

Direction:

Strength:



9) Use your calculator to determine the least-squares regression line. Write the equation, and plot this line on your graph. (Be sure to use two actual points from the regression equation for an accurate sketch.)

10) Interpret the slope of the regression line in the context of the problem.

11) Interpret the y-intercept of the regression line in the context of the problem.

12) Clearly label the residual for the data point (3, 9) on the grid above and give its value.

13) How well does the linear model you calculated in Question 11 fit the data? (Interpret r^2 in the context of the problem).