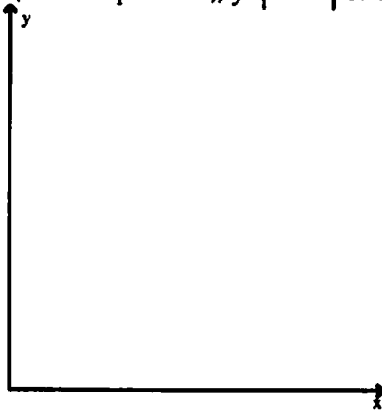


Construct a scatter diagram for the data.

- 1) In an area of the Great Plains, records were kept on the relationship between the rainfall (in inches) and the yield of wheat (bushels per acre). Find the correlation coefficient. What is the critical value? Is there a relationship between rainfall and yield?

Rainfall (in inches), x	10.5	8.8	13.4	12.5	18.8	10.3	7.0	15.6	16.0
Yield (bushels per acre), y	50.5	46.2	58.8	59.0	82.4	49.2	31.9	76.0	78.8

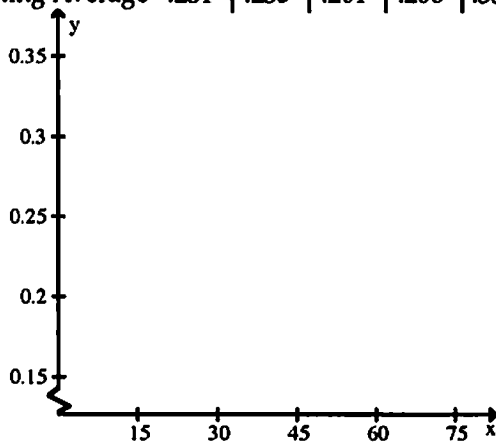


Provide an appropriate response.

- 2) An agricultural business wants to determine if the rainfall in inches can be used to predict the yield per acre on a wheat farm. Identify the predictor variable and the response variable.

- 3) The numbers of home runs that Mark McGwire hit in the first 13 years of his major league baseball career are listed below. (Source: Major League Handbook) Construct a scatter diagram for the data. Find the correlation coefficient. What is the critical value? Is there a relationship between the home runs and the batting averages?

Home Runs	33	39	22	42	9	9	39	52	58	70
Batting Average	.231	.235	.201	.268	.33	.252	.274	.312	.274	.299



4) A medical researcher wishes to determine if there is a relationship between the number of prescriptions written by medical professionals, per 100, children and the child's age. She surveys all the pediatrician's in a geographical region to collect her data. What is the explanatory variable?

5) Calculate the linear correlation coefficient for the data below. Is there a linear relationship?

x	-13	-11	-4	-7	-9	-10	-8	-6	-5	-12
y	6	-11	3	-8	-7	-4	0	-10	1	2

6) In an area of the Great Plains, records were kept on the relationship between the rainfall (in inches) and the yield of wheat (bushels per acre). Find the LSRL and calculate the linear correlation coefficient. What does the linear correlation coefficient tell us about the data?

Rainfall (in inches), x	13.4	11.7	16.3	15.4	21.7	13.2	9.9	18.5	18.9
Yield (bushels per acre), y	55.5	51.2	63.8	64	87.4	54.2	36.9	81	83.8

7) For a random sample of 30 countries, the linear correlation coefficient between the infant mortality rate and the average number of cars per capita was found to be $r = -0.717$. What does this imply? Does this suggest that if people buy more cars, this could lower the infant mortality rate? Why or why not? What is a likely lurking variable?

8) The data below are the final exam scores of 10 randomly selected history students and the number of hours they slept the night before the exam. Find the equation of the regression line for the given data. What would be the predicted score for a history student who slept 7 hours the previous night? Is this a reasonable question? Round the regression line values to the nearest hundredth, and round the predicted score to the nearest whole number.

Hours, x	3	5	2	8	2	4	4	5	6	3
Scores, y	65	80	60	88	66	78	85	90	90	71

9) For a given data set, the equation of the least squares regression line will always pass through

A) at least two point in the given data set.

B) the y -intercept and the slope.

C) every point in the given data set.

D) (\bar{x}, \bar{y}) .

10) Is there a relationship between the raises administrators at State University receive and their performance on the job? A faculty group wants to determine whether job rating (x) is a useful linear predictor of raise (y). Consequently, the group considered the straight-line regression model, $\hat{y} = \beta_0 + \beta_1 x$. Using the method of least squares, the faculty group obtained the following prediction equation, $\hat{y} = 14,000 + 2,000x$. Interpret the estimated slope of the line.

11) The regression line for the given data is $\hat{y} = 2.097x - 0.552$. Determine the residual of a data point for which $x = -2$ and $y = -6$.

x	-5	-3	4	1	-1	-2	0	2	3	-4
y	-10	-8	9	1	-2	-6	-1	3	6	-8

12) Calculate the coefficient of determination to the nearest thousandth, given that the linear correlation coefficient, r , is -0.625 . What does this tell you about the explained variation and the unexplained variation of the data about the regression line?

13) Each year a nationally recognized publication conducts its "Survey of America's Best Graduate and Professional Schools." An academic advisor wants to predict the typical starting salary of a graduate at a top business school using GMAT score of the school as a predictor variable. A simple linear regression of SALARY versus GMAT using 25 data points shown below.

$$b_0 = -92040 \quad b_1 = 228 \quad s = 3213 \quad R^2 = 0.66 \quad r = 0.81 \quad df = 23 \quad t = 6.67$$

Give a practical interpretation of $R^2 = 0.66$.

14) If the coefficient of determination is close to 1, then

- A) the least squares regression line equation has no explanatory value.
- B) the sum of the square residuals is large compared to the total variation.
- C) the least squares regression line equation explains most of the variation in the response variable.
- D) the linear correlation coefficient is close to zero.

15) The following data represent the living situation of newlyweds in a large metropolitan area and their annual household income. Find the marginal frequency for newlyweds who make between \$20,000 and \$35,000 per year.

	<\$20,000	\$20-35,000	\$35-50,000	\$50-75,000	>\$75,000
Own home	31	52	202	355	524
Rent home	67	66	52	23	11
Live w/family	89	69	30	4	2

16) The following data represent the living situation of newlyweds in a large metropolitan area and their annual household income. What percent of people who make less than \$20,000 per year own their own home? Round to the nearest tenth of a percent.

	<\$20,000	\$20-35,000	\$35-50,000	\$50-75,000	>\$75,000
Own home	31	52	202	355	524
Rent home	67	66	52	23	11
Live w/family	89	69	30	4	2

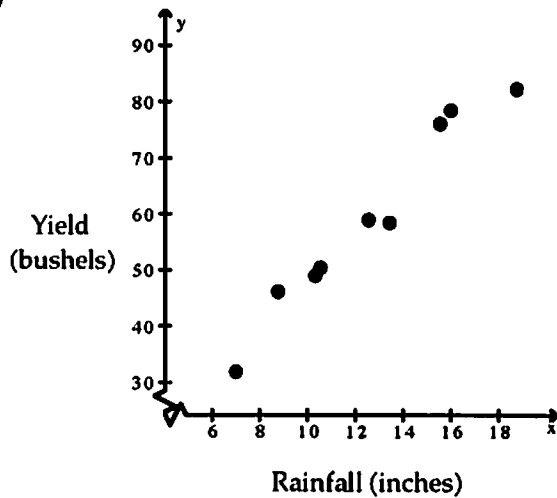
17) The following data represent the living situation of newlyweds in a large metropolitan area and their annual household income. What percent of people who rent their home make between \$35,000 and \$50,000 per year? Round to the nearest tenth of a percent.

	<\$20,000	\$20-35,000	\$35-50,000	\$50-75,000	>\$75,000
Own home	31	52	202	355	524
Rent home	67	66	52	23	11
Live w/family	89	69	30	4	2

Answer Key

Testname: MATH1040(CHPT4)REVIEW

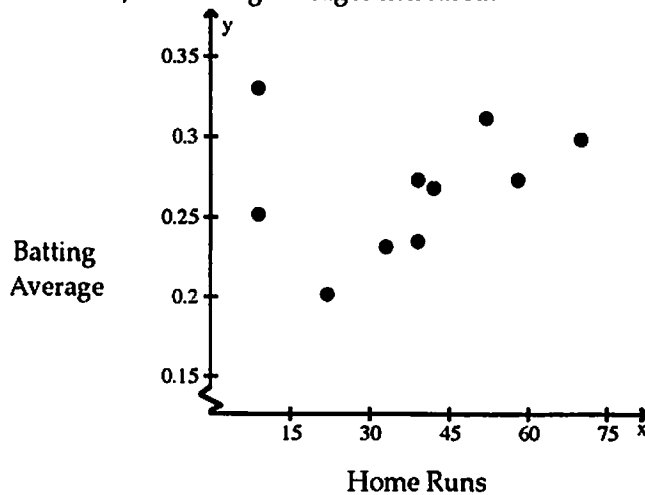
1)



$r = .98$ $cv = .666$ yes, $.98 > .666$ so there is a linear relationship.

2) predictor variable: rainfall in inches; response variable: yield per acre

3) In general, there appears to be a relationship between the home runs and batting averages. As the number of home runs increased, the batting averages increased. $r = .228$ $cv = .632$ No, $.228 < .632$ so there isn't a linear relationship.



4) Age of the child

5) -0.104 ; No the correlation is too low.

6) 0.981 ; There is a strong positive linear correlation

7) A negative correlation exists between the infant mortality rate and the number of cars per capita but this is an example of correlation not causation. If people buy more cars, this is unlikely to lead to a decrease in the infant mortality rate. A likely lurking variable is wealth and this lurking variable accounts for the negative correlation. More affluent countries tend to have both more cars per capita and lower infant mortality rates.

8) $\hat{y} = 5.04x + 56.11$; 91; Yes, it is reasonable.

9) D

10) For a 1-point increase in an administrator's rating, we estimate the administrator's raise to increase \$2,000.

11) -1.254

12) The coefficient of determination, $R^2 = 0.391$. That is, 39.1% of the variation is explained and 60.9% of the variation is unexplained.

13) 66% of the sample variation in SALARY can be explained by using GMAT in a straight-line model.

14) C

15) 187

16) 16.6%

17) 23.7%