

4.3 Exercise

MATH 241

THOMPSON

1) The coefficient of determination, R^2 , measures the proportion of total variation in the response variable that is explained by the least squares regression line.

2) Total deviation = unexplained deviation + explained deviation

The deviation between the predicted and mean values of the response variable is called the explained deviation, so explained deviation is $\hat{y} - \bar{y}$. The deviation between the observed and predicted values of the response variable is called the unexplained deviation, so unexplained deviation is $y - \hat{y}$. Therefore, total deviation equals unexplained deviation plus explained deviation.

$$y - \bar{y} = (y - \hat{y}) + (\hat{y} - \bar{y})$$

3)

Match the coefficient of determination to the scatter diagram. The scales on the x-axis and y-axis are the same for each scatter diagram.

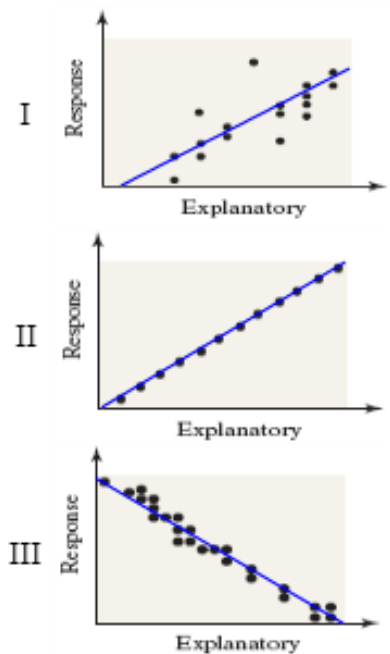
(a) $R^2 = 1$, (b) $R^2 = 0.98$, (c) $R^2 = 0.62$

(a) Scatter diagram II

(b) Scatter diagram III

(c) Scatter diagram I

The lower the percent the more scattered the points are on the



4)

The following data represent the time between eruptions and the length of eruption for 9 randomly selected geyser eruptions. The coefficient of determination is 80.3%. Provide an interpretation of this value.

Time, x	Length, y	Time, x	Length, y
12.12	1.83	11.69	1.81
11.73	1.72	12.16	1.87
11.97	1.79	11.64	1.75
12.11	1.84	11.74	1.79
11.31	1.65		

The least-squares regression line explains 80.3 % of the variation in length of eruption.

(Round to one decimal place as needed.) Given in problem

y variable

- 5) Suppose a doctor measures the height, x , and head circumference, y , of 11 children and obtains the data below. The correlation coefficient is 0.798 and the least squares regression line is $\hat{y} = 0.155x + 13.114$. Complete parts (a) and (b) below.

Height, x	27	25.5	26.75	25.75	28	26.5	25.75	27.25	27.25	27.25	27
Head Circumference, y	17.2	17.1	17.3	16.9	17.4	17.2	17.2	17.4	17.3	17.3	17.4

(a) Compute the coefficient of determination, R^2 .

$R^2 = 63.7\%$ (Round to one decimal place as needed.)

(b) Interpret the coefficient of determination.

Approximately 63.7 % of the variation in head circumference is explained by the least-squares regression model.
(Round to one decimal place as needed.) **appropriate**

STATCRUNCH
STAT – REGRESSION-SIMPLE LINEAR
CALCULATE
multiply by 100 to get % of R^2 then round

6)

The accompanying data represent the weights of various domestic cars and their gas mileages in the city. The linear correlation coefficient between the weight of a car and its miles per gallon in the city is $r = -0.974$. The least-squares regression line treating weight as the explanatory variable and miles per gallon as the response variable is $\hat{y} = -0.0066x + 43.3298$. Complete parts (a) and (b) below.

 Click the icon to view the data table.

(a) What proportion of the variability in miles per gallon is explained by the relation between weight of the car and miles per gallon?

The proportion of the variability in miles per gallon explained by the relation between weight of the car and miles per gallon is 94.9 %.

(Round to one decimal place as needed.)

(b) Interpret the coefficient of determination.

94.9 % of the variance in gas mileage is explained by the linear model.

(Round to one decimal place as needed.)

STATCRUNCH
STAT – REGRESSION-SIMPLE LINEAR
x is weight y is miles
multiply by 100 to get % of R^2 then round

- 7) What is the least-squares regression model between x and y for Data Set A?

Full Data Set: 

Data Set A		Data Set B		Data Set C	
x	y	x	y	x	y
3.6	8.9	3.1	8.9	2.8	8.9
8.3	15.0	9.4	15.0	8.1	15.0
0.5	4.8	1.2	4.8	2.0	4.8
1.4	6.0	5.4	6.0	1.0	6.0
8.2	14.9	5.9	14.9	5.9	14.9
5.9	11.9	4.3	11.9	4.3	11.9
4.3	9.8	3.4	9.8	3.4	9.8

STATCRUNCH

STAT – REGRESSION-SIMPLE LINE

x is first x y is first y

STATCRUNCH
STAT – REGRESSION-SIMPLE LINEAR
x is first x y is first y

What is the least-squares regression model between x and y for Data Set A, rounding to four decimal places? Choose the correct answer below.

- ☒ A. $\hat{y} = 1.2992x + 4.2159$
☐ B. $\hat{y} = 4.2159x + 1.1797$
☐ C. $\hat{y} = 0.4368x + 8.1464$
☐ D. $\hat{y} = 1.1797x + 4.8438$

9) What is the coefficient of determination for Data Set B?

Full Data Set: 

Data Set A		Data Set B		Data Set C	
x	y	x	y	x	y
3.6	8.9	3.1	8.9	2.8	8.9
8.3	15.0	8.4	15.0	8.4	15.0
0.5					4.8
1.4					6.0
8.2					14.9
5.9	11.9	5.0	11.9	1.4	11.9
4.3	9.8	3.4	9.8	1.0	9.8
8.3	15.0	7.4	15.0	7.9	15.0
0.3	4.7	0.1	4.7	5.9	4.7
6.8	13.0	7.5	13.0	5.0	13.0

(This is a reading assessment question. Be certain of your answer because you only get one attempt on this question.)

The coefficient of determination for Data Set B is 94.7 %.

(Type an integer or decimal rounded to the nearest tenth as needed.)

R² *make sure to use column B only
(second x and second y)