

Name: _____

Date: _____

Correlation Coefficient Algebra 1

In the last lesson we saw that for every data set relating two variables a **line of best fit** can be found. How well this line fits the data given, though, is questionable. To measure this fit, we use what is known as the **correlation coefficient**, commonly given the letter r .

Exercise #1: Turn your correlation coefficient on by doing the following steps.

Step 1 – Go to **CATALOG** and select **DiagnosticsOn**.

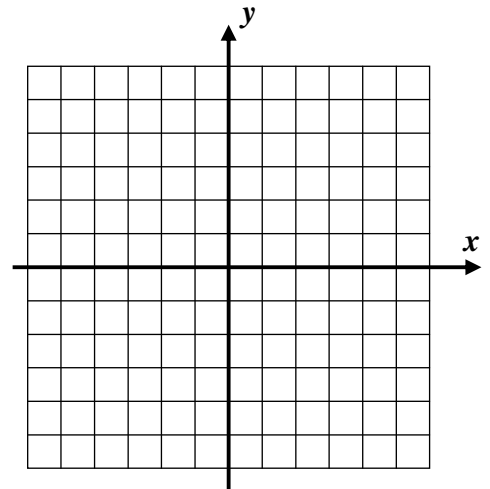
Step 2 – Hit **ENTER** twice

```
CATALOG [ ]
DependAuto
det(
DiagnosticOff
▶DiagnosticOn
dim(
Disp
DispGraph
```

Exercise #2: Consider the data set given to the right.

$$\{(-4, -2), (-2, 0), (1, 3), (3, 5)\}$$

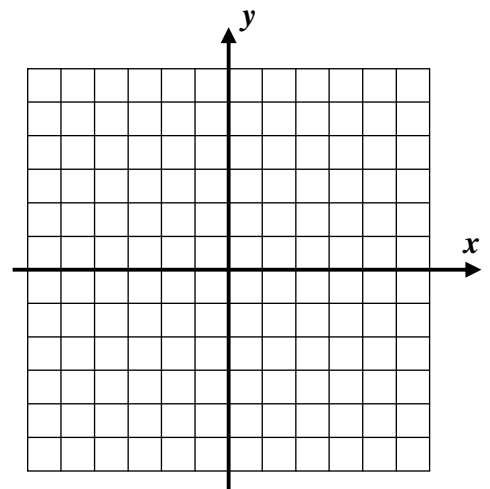
- (a) Graph the ordered pairs on the given grid.
- (b) Determine the equation for the line of best fit.
- (c) State the value of the correlation coefficient.



Exercise #3: Consider the data set given to the right.

$$\{(-2, 6), (0, 3), (2, 0), (4, -3)\}$$

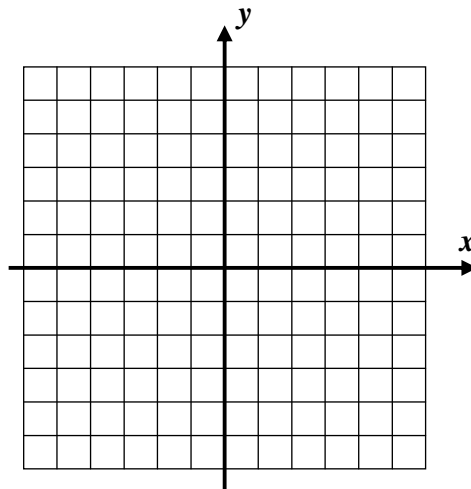
- (a) Graph the ordered pairs on the given grid.
- (b) Determine the equation for the line of best fit.
- (c) State the value of the correlation coefficient.



Exercise #4: Consider the data set given to the right.

$$\{(1,3), (3,4), (5,6), (6,5)\}$$

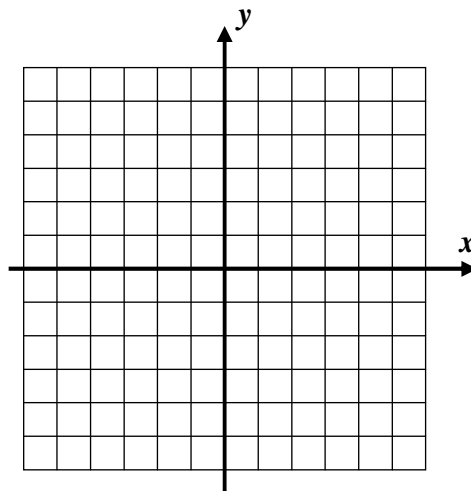
- (a) Graph the ordered pairs on the given grid.
- (b) Determine the equation for the line of best fit. Round coefficients to the *nearest hundredth*. Draw a rough sketch of the line of best fit on the axes.
- (c) State the value of the correlation coefficient. Round your answer to the nearest *hundredth*.
- (d) Why is the correlation coefficient not equal to one?



Exercise #5: Consider the data set given to the right.

$$\{(-4,5), (0,4), (2,-2), (5,-4)\}$$

- (a) Graph the ordered pairs on the given grid.
- (b) Determine the equation for the line of best fit. Round coefficients to the nearest *tenth*. Draw a rough sketch of the line of best fit on the axes.
- (c) State the value of the correlation coefficient. Round your answer to the nearest *hundredth*.
- (d) Why is the correlation coefficient negative?



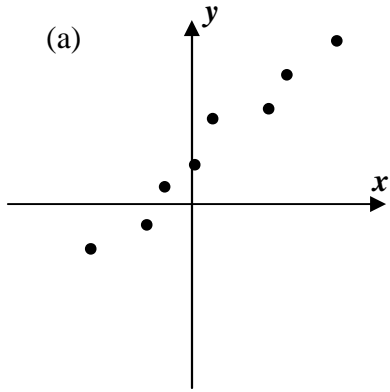
CORRELATION COEFFICIENT SUMMARY, r -VALUE

- | | |
|--|--|
| (1) r has a value on the interval $-1 \leq r \leq 1$. | (2) $r = \pm 1$ indicates a perfect linear fit |
| (3) $r = 0$ indicates no linear correlation | (4) the sign of r indicates a positive or negative slope |

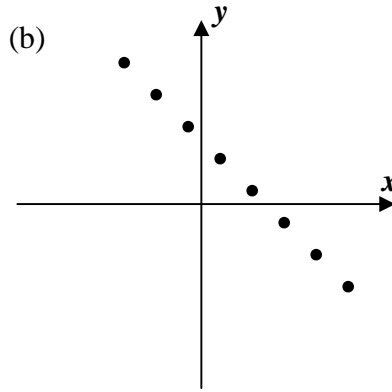
Correlation Coefficient Algebra 1 Homework

Skills

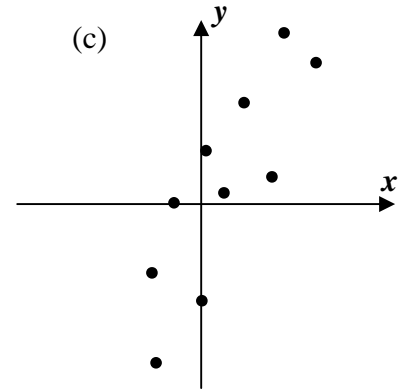
1. For each of the six scatterplots drawn below, sketch the line of best fit. Then, for each choose the most appropriate r -value given from the list below all of the figures.



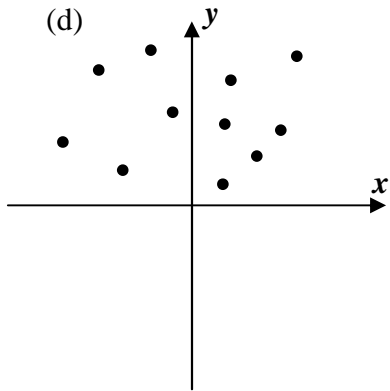
$r =$



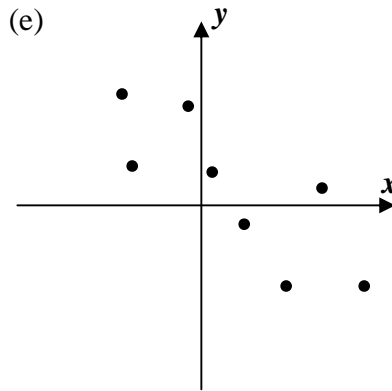
$r =$



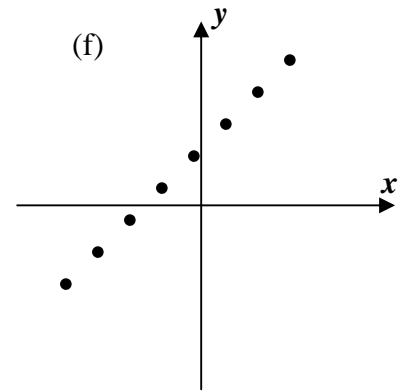
$r =$



$r =$



$r =$



$r =$

Possible r -values: -1.00, -0.65, 0.00, 0.75, 0.92, 1.00

2. Which of the following is closest to the correlation coefficient of the data set given below?

(1) 0.72

(3) 0.95

(2) 0.89

(4) -0.25

L1	L2	L3	Z
-2	5	-----	
0	5		
2	3		
4	11		
-----	-----		
L2(1) = -2			

3. Find the correlation coefficient, r , to the *nearest hundredths* place, for each list of ordered pairs.

(a) (-2, 1), (0, 3), (2, 5), (4, 7)

$r =$ _____

(b) (-7, 10), (-5, 6), (-3, 8), (-1, 2)

$r =$ _____

(c) (-8, 8), (2, -12), (-6, 8), (-4, 0), (3, -10), (1, 12)

$r =$ _____

Applications

4. The low and high temperatures for one week in Poughkeepsie, New York, are given below.

Low Temp, x	27	31	42	15	33	35	41
High Temp, y	42	44	58	31	50	62	65

(a) Write the best linear model, in $y = ax + b$ form, that gives the high temperature as a function of the low temperature. Round all coefficients to the nearest *tenth*.

(b) State the correlation coefficient of this data set to the nearest *hundredth*.

(c) If this model held for the following week, what was the high on a day when the low temperature was 44? Round your answer to the nearest *whole* degree.

(d) Another linear model predicted the high temperature based on the highest humidity of the day. For this linear model, the correlation coefficient was 0.85. Which model does a better job in predicting the high temperature? Justify.

Reasoning

5. The linear correlation coefficient provides two pieces of information about a data set and its line of best fit. What are they? Explain.