Assignment

LESSON 2: Gotta Keep It Correlatin'

Write

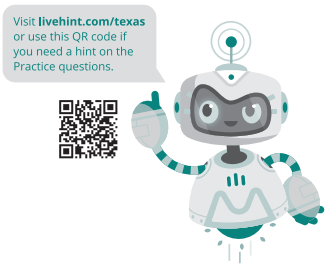
Complete each sentence.

1. A correlation is a _____ for causation, but a correlation is not a _____ for causation.
2. A _____ is when some other reason may cause the same result.
3. _____ is when one event causes a second event.
4. A _____ is when there are other variables that are unknown or unobserved.
5. The _____ is a value between −1 and 1 that indicates how close the data are to forming a straight line.
6. The percentage of variation of the observed values of the data points from their predicted values is represented by the _____.

Remember

Sets of data can frequently be modeled by using a linear function called a *regression equation*. A value called the *correlation coefficient* can also be calculated to assist in determining how well the regression equation fits the data.

Practice



1. The table shows the percent of the United States population who did not receive needed dental care services due to cost.

Year	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Percent	7.9	8.1	8.7	8.6	9.2	10.7	10.7	10.8	10.5	12.6	13.3

   a. Do you think a linear regression equation would best describe this situation? Why or why not?
   b. Determine the linear regression equation for these data. Interpret the equation in terms of this problem situation.
   c. Compute and interpret the correlation coefficient of this data set. Does it seem appropriate to use a line of best fit? Explain your reasoning.
   d. Sketch a scatter plot of the data. Then, plot the equation of the regression line on the same grid. Do you still think a linear regression is appropriate? Explain your answer.

2. A teacher claims that students who study will receive good grades.
   a. Do you think that studying is a necessary condition for a student to receive good grades?
   b. Do you think that studying is a sufficient condition for a student to receive good grades?
   c. Do you think that there is a correlation between students who study and students who receive good grades?
   d. Do you think that it is true that studying will cause a student to receive good grades?
   e. List two or more confounding variables that could have an effect on this claim.
3. For each situation, decide whether the correlation implies causation. List reasons why or why not.
   a. The number of violent video games sold in the U.S. is highly correlated to crime rates in real life.
   b. The number of newspapers sold in a city is highly correlated to the number of runs scored by the city's professional baseball team.
   c. The number of mouse traps found in a person's house is highly correlated to the number of mice found in their house.

Stretch

Consider the points: (1, 2), (2, 3), (3, 2), (4, 5), (5, 2.5), (6, 6), (7, 3), (8, 7). The line of best fit for the graph of the points is $y = 0.5x + 1.4$.

1. Complete the table to determine the predicted values of y for each value of x using the line of best fit, and the values of the differences between the observed y-values from the points and the predicted y-values from the line of best fit.

x	Observed y-Value	Predicted y-Value	Observed y-Value − Predicted y-Value
1	2	1.9	0.1
2	3		
3	2		
4	5		
5	2.5		
6	6		
7	3		
8	7		

2. Determine whether there is a pattern in the differences between the y-values from the completed table. Explain what this might indicate about using the line of best fit to make predictions.

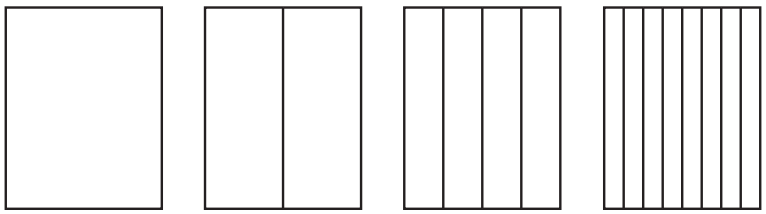
Review

1. The table shows the highest maximum temperature for the month of October in Philadelphia, Pennsylvania, over ten years.

Year	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Highest Maximum Temperature (°F)	64.9	53.1	61	54	63	68	61	57.9	64.9	66.9

   a. Identify the independent and dependent quantities and their units of measure.
   b. Use the data table and graphing technology to generate a line of best fit. What is the slope and *y*-intercept of the line and what do they represent?

2. Harrison draws a rectangle, and then in each successive figure he splits the rectangles into two rectangles as shown.



   a. Analyze the number of rectangles in each figure. Describe the pattern.
   b. Write the number of rectangles in each of the first six figures as a numeric sequence.

3. Determine the slope, *x*-intercept, and *y*-intercept of the graph.