Interpreting the Correlation Coefficient

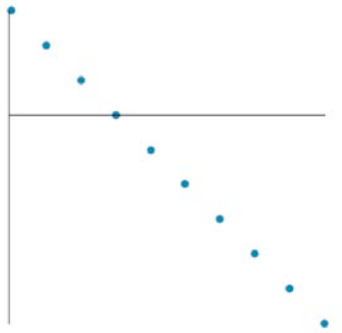
Use What You've Learned
Answers and Explanations

After watching the video, *Interpreting Correlation Coefficient,* complete the following problems.

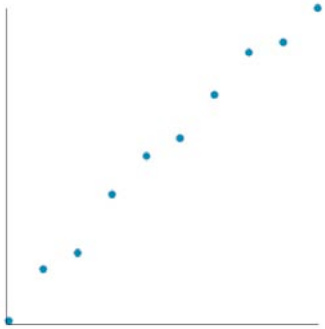
1. We've seen that some data have a positive correlation, meaning that their scatter plots look linear. We've also seen that some data have no correlation, meaning that they don't seem to be related in a linear way. Statisticians quantify the closeness of data points to a line by the "correlation coefficient." If the correlation coefficient is 1, then the points lie perfectly on a line. If it is zero, then they don't resemble a line (they may be related in other ways). Match the following scatter plots to the correlation coefficient. (Part of the fun will be figuring out what a negative correlation coefficient means!)

Correlation Coefficients:

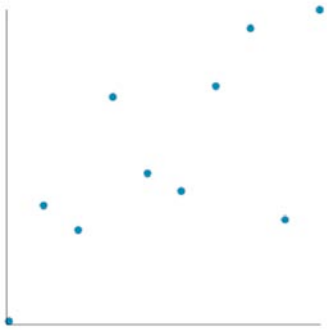
A) 1 B) 0.988 C) -0.8 D) 0 E) -0.2 F) 0.9 G) -1



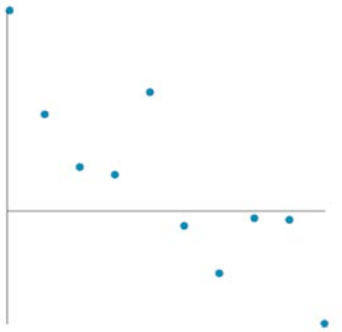
G



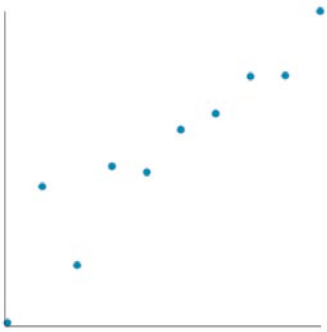
B



D



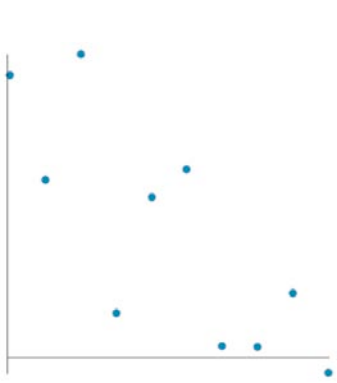
F



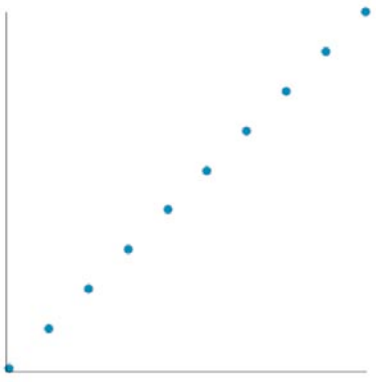
C

Interpreting the Correlation Coefficient

Use What You've Learned
Answers and Explanations



E

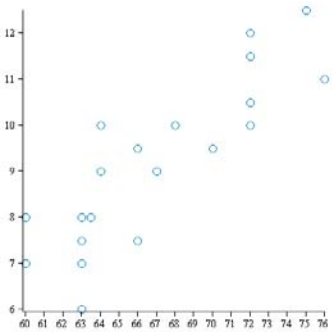


A

2. A study shows a strong negative correlation between hours of daylight, and admissions to a depression ward in a hospital. The more hours of daylight, the fewer the cases of depression. Is it possible that daylight prevents depression? Is it true that this study proves that daylight prevents depression (or that darkness causes depression)? If not, present some possible lurking variables that could explain this correlation.

   Yes, it is possible that daylight prevents depression. No, this study does not prove that daylight prevents depression. It may be that hours of daylight allow people to do more exercise, and THAT prevents depression. It could be that people get more depressed around the winter holidays, and there are fewer hours of daylight around the winter holidays as well. A lurking variable is anything that would cause an increase in depression that is ALSO correlated to decreased hours of daylight.

3. a. Do the two variables represented in the following scatter plot have a positive or a negative correlation?



   This scatter plot has a positive correlation coefficient because there is a general trend of the points up and to the right.

b. Which one of the following is the correlation coefficient, r, of the scatter plot above?

-1 -0.85 -0.15 0 0.15 0.85 1

0.85; Since it has a positive correlation, none of the negative values are reasonable. Both zero and 0.15 would have scatter plots that are much more scattered. Since the points are not on a straight line, the answer cannot be one. 0.85 means that there is a general pattern of a positive linear relationship but it is not a perfect.

4. Rank the following correlations from least to greatest (from most negative number to most positive number)
   a. The age of a randomly selected pet kitten, and the age of its owner
   b. The age of a Ford Escort, and its blue-book value
   c. The temperature in Dallas in degrees Fahrenheit, and the temperature in Dallas in degrees Celsius
   d. The height of a randomly selected man, and his weight
   e. The amount of our pie that I have eaten, and the amount left for you

   e < b < a < d < c

5. Is there a positive or negative correlation between area of a square and its side length, assuming the side length is between one and 20 inches? Is this correlation close to one or negative one? Explain your reasoning.

   There is a positive correlation - the longer the side length, the larger the area. But, the correlation coefficient is meaningless in this case because the graph of length vs. area is not linear. Correlation coefficient reflects the strength of a linear relationship.