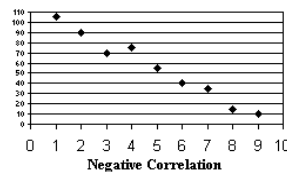
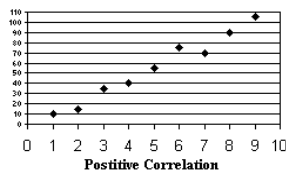


Study Guide

Estimate Line of Best Fit Scatter Plots 02/29/2012

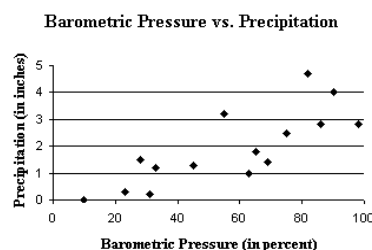
Estimate Line of Best Fit: Scatter Plot

A scatter plot is a graph that shows how much one variable is affected by another. One variable is plotted along the horizontal axis and the other variable is plotted along the vertical axis. The relationship between the two variables is called their correlation. Scatter plots usually consist of a large amount of data. The closer the data points come to making a straight line, the higher the correlation between the two variables (or the stronger the relationship). If the data points fall along a straight line from a low value to a high value, then the values are said to have a positive correlation. If the data points fall along a straight line going from a high value to a low value, then the variables are said to have a negative correlation. When looking for positive or negative correlation, look at the graph from left to right.



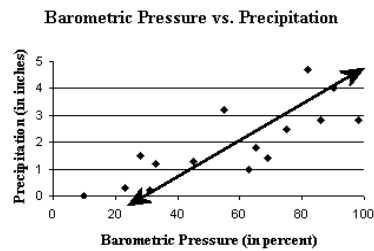
A line of best fit shows the relationship between two variables in a scatter plot. The line that is drawn represents an *average* of all of the data points, most of which will probably not lie on the line itself. Remember, the line is an estimation and not an exact measure. The points on each side of the line of best fit should be as close to the line as possible. Once the line of best fit is drawn, it is possible to estimate the value of one variable if the value of the other variable is known. This study guide will focus on drawing an estimated line of best fit for a scatter plot.

Example 1: Draw an estimated line of best fit for the Barometric Pressure vs. Precipitation scatter plot below.



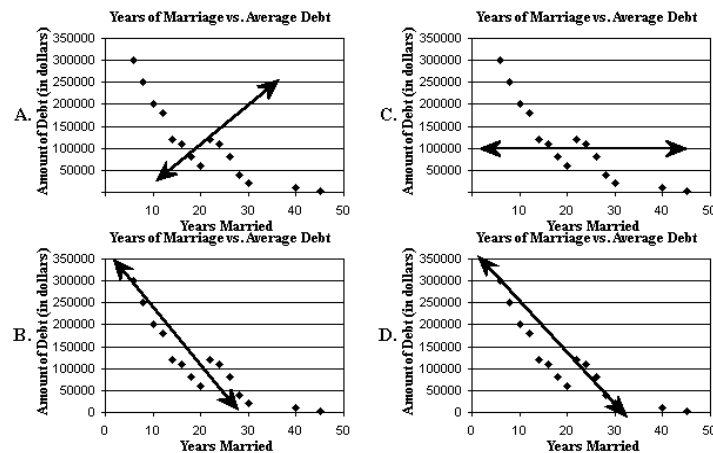
Step 1: Determine if there is a positive or a negative correlation between the two variables. Since the precipitation increases as the barometric pressure increases, there is a positive correlation. This means that the line of best fit will go up from left to right.

Step 2: Determine where to place the line of best fit. Recall that the line of best fit represents the *average* of all of the data points. This means that there should be close to the same number of data points above the line of best fit as there are below the line of best fit, and the points on each side of the line should be as close to the line as possible.



The placement of this line of best fit includes 7 points above the line and 6 points below the line. This is the estimated line of best fit.

Example 2: The graphs below show the relationship between the number of years a couple has been married and the amount of debt the couple is in. Which graph illustrates the line of best fit for this data?



Step 1: Determine whether there is a positive or a negative correlation between the two variables. Since the data points in this case get lower as the graph is read from left to right, there is a negative correlation. This means that the line of best fit will go down from left to right.

Step 2: Eliminate answer choices **A.** and **C.** because their lines of best fit do not go down from left to right.

Step 3: Of the two graphs left, determine which one represents the *average* of the two variables. Remember that this means that there should be approximately the same number of data points above the line of best fit as there are below the line of best fit and the points on each side of the line should be as close to the line as possible.

Answer: B. This graph shows the data points with a line of best fit that is drawn in the correct direction and has approximately the same number of data points above the line and below the line.

An activity that could help the student with this skill is to find scatter plots in newspapers, magazines, and on the Internet and have the student draw estimated lines of best fit for the data. Have the student justify the reasoning behind drawing the line by asking questions such as:

- Why did he or she draw a positive (or negative) correlation?
- Why did he or she place the line where it is?
- What does the line of best fit tell about the data?