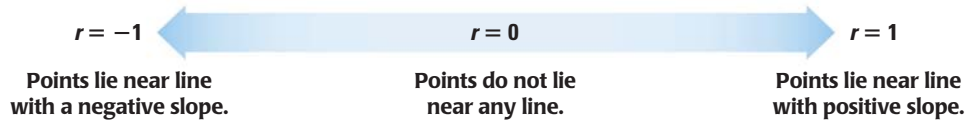
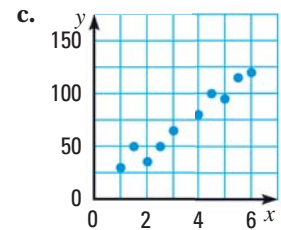
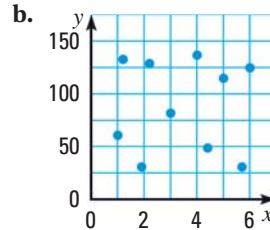
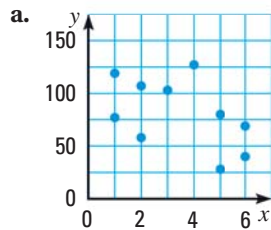


CORRELATION COEFFICIENTS A **correlation coefficient**, denoted by r , is a number from -1 to 1 that measures how well a line fits a set of data pairs (x, y) . If r is near 1 , the points lie close to a line with positive slope. If r is near -1 , the points lie close to a line with negative slope. If r is near 0 , the points do not lie close to any line.



EXAMPLE 2 Estimate correlation coefficients

Tell whether the correlation coefficient for the data is closest to -1 , -0.5 , 0 , 0.5 , or 1 .

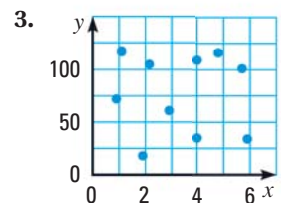
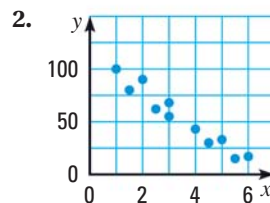
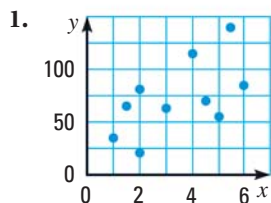


Solution

- a. The scatter plot shows a clear but fairly weak negative correlation. So, r is between 0 and -1 , but not too close to either one. The best estimate given is $r = -0.5$. (The actual value is $r \approx -0.46$.)
- b. The scatter plot shows approximately no correlation. So, the best estimate given is $r = 0$. (The actual value is $r \approx -0.02$.)
- c. The scatter plot shows a strong positive correlation. So, the best estimate given is $r = 1$. (The actual value is $r \approx 0.98$.)

GUIDED PRACTICE for Examples 1 and 2

For each scatter plot, (a) tell whether the data have a *positive correlation*, a *negative correlation*, or *approximately no correlation*, and (b) tell whether the correlation coefficient is closest to -1 , -0.5 , 0 , 0.5 , or 1 .



BEST-FITTING LINES If the correlation coefficient for a set of data is near ± 1 , the data can be reasonably modeled by a line. The **best-fitting line** is the line that lies as close as possible to all the data points. You can approximate a best-fitting line by graphing.