


3.23: Linear Relationships (4 of 4)

A link to an interactive elements can be found at the bottom of this page.

To see how an outlier affects the correlation, do the following:

1. Fill the scatterplot with a hypothetical positive linear relationship between X and Y (by clicking on the graph about a dozen times starting at the lower left and going up diagonally to the top right). Pay attention to the correlation coefficient calculated at the top left of the simulation. (Clicking on the garbage can lets you start over.)
2. Once you are satisfied with your hypothetical data, create an outlier by clicking on one of the data points in the upper right of the graph and dragging it down along the right side of the graph. Again, pay attention to what happens to the value of the correlation.

What did this activity illustrate? This activity illustrates that the correlation decreases when the outlier deviates from the pattern of the relationship. By dragging a data point from the upper right to the lower right, you created an outlier that does not fit the positive association in the rest of the data. This decreases the strength of the linear relationship and causes a decrease in .

In the next activity, you will see how the correlation increases when the outlier is consistent with the direction of the linear relationship.

Let's Summarize

- A special case of the relationship between two quantitative variables is the **linear** relationship in which a straight line simply and adequately summarizes the relationship.
- When the scatterplot displays a linear relationship, we supplement it with the correlation coefficient (r), which measures the *strength* and *direction* of a linear relationship between two quantitative variables. The correlation ranges between -1 and 1 . Values near -1 indicate a strong negative linear relationship, values near 0 indicate a weak linear relationship, and values near 1 indicate a strong positive linear relationship.
- The correlation is an appropriate numerical measure only for linear relationships and is sensitive to outliers. Therefore, the correlation should be used only as a supplement to a scatterplot (after we look at the data).

Contributors and Attributions

CC licensed content, Shared previously

- Concepts in Statistics. **Provided by:** Open Learning Initiative. **Located at:** <http://oli.cmu.edu>. **License:** [CC BY: Attribution](#)

This page titled [3.23: Linear Relationships \(4 of 4\)](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Lumen Learning](#).