

Variable Relationships

3 min

Aggregating data is a way of exploring variable relationships. We specifically looked at relationships between a numeric variable and a categorical variable, but we should also examine relationships between two numeric

Preview: Docs Loading link description

[variables](#)

.

For example, we might wonder: Does musician income vary with years of experience? To start, we can take a look at a **scatter plot** with experience on the x-axis and income on the y-axis. Each point in the plot represents a musician, and the coordinates of that point are the musician's experience (x) and income (y).

The cloud of points in the plot has a pattern. The points move from the lower left to the upper right part of the plot. In other words, lower levels of experience tend to be associated with lower incomes, and higher levels of experience tend to be associated with higher incomes. The points don't form a perfect line though — there is some variation.

We can describe this relationship more precisely by measuring the **correlation coefficient**. This number ranges from -1 to +1 and tells us two things about a linear relationship:

1. **Direction:** A positive coefficient means that higher values in one variable are associated with higher values in the other. A negative coefficient means higher values in one variable are associated with lower values of the other.
2. **Strength:** The farther the coefficient is from 0, the stronger the relationship and the more the points in a scatter plot look like a line.

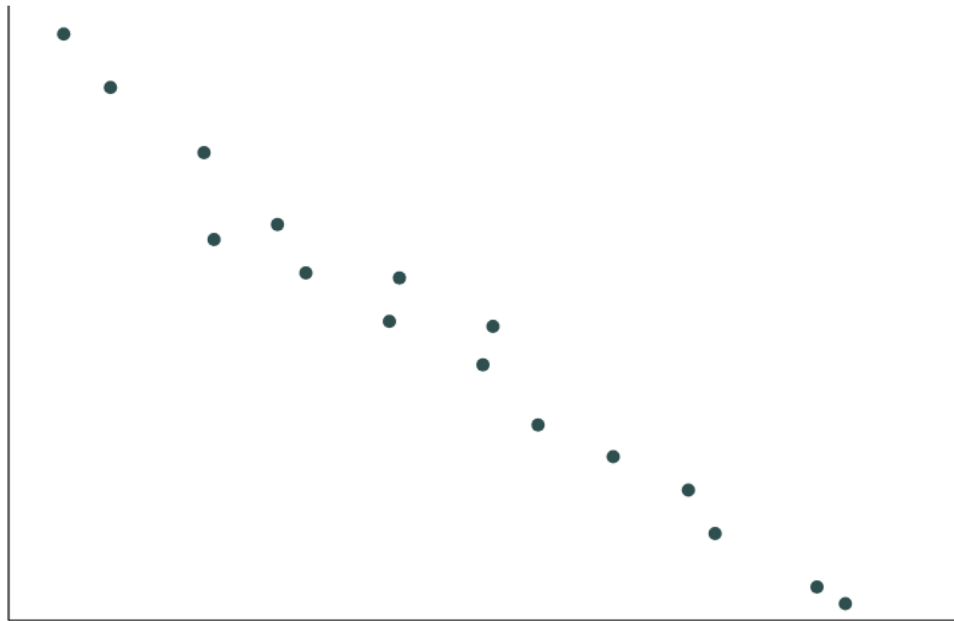
The correlation coefficient for income and experience is 0.74 — the relationship is positive and moderately strong.

Instructions

Check out the interactive art in the learning environment. As you move the slider, the scatter plot and corresponding correlation coefficient change. What kind of correlation coefficient do you think occurs when the points form a curved line rather than a straight line?

correlation coefficient: -1.00

strong correlation

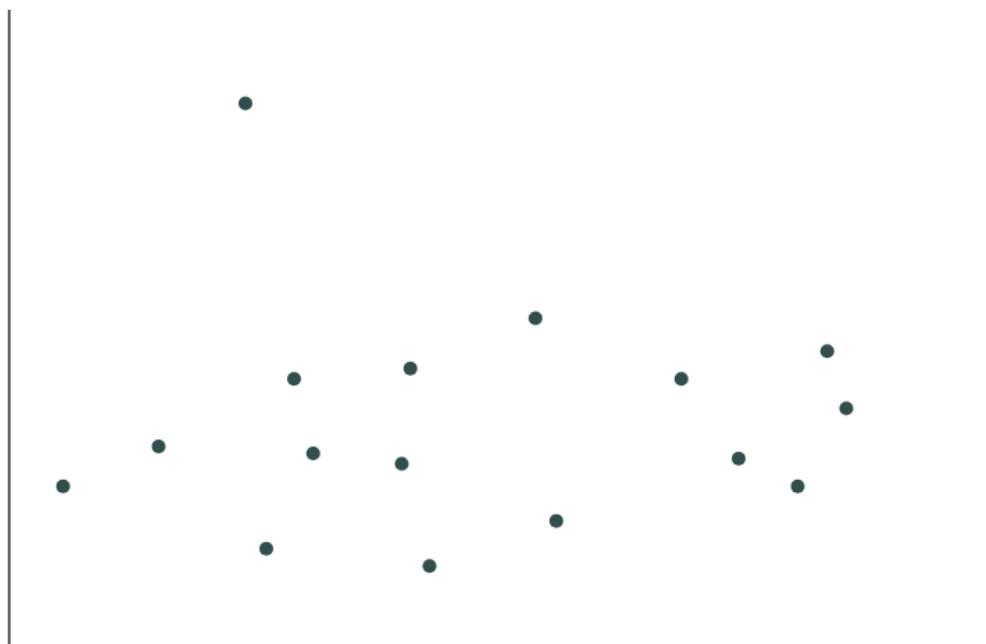


correlation coefficient: -.50



correlation coefficient: 0

weak correlation



correlation coefficient: .50



correlation coefficient: 1.00

strong correlation

