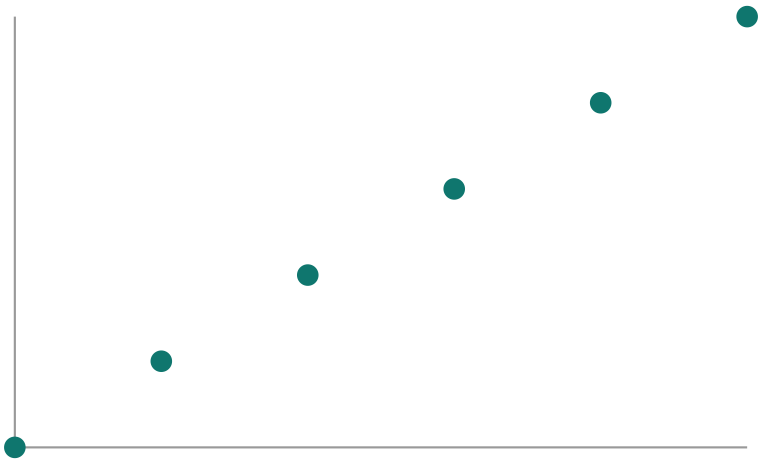
Pearson vs Spearman Correlation

Pearson and Spearman correlation both summarize a relationship between two paired variables, but they answer slightly different questions. Select one of four fixed patterns below to see both coefficients computed on the exact same data.

StatReason

Pearson vs Spearman Correlation

Pattern



Pearson r	1
Spearman rho	1

Pearson r = 1, Spearman rho = 1

What this answers

This page answers "when do Pearson and Spearman correlation agree, and when do they tell a genuinely different story about the same data?" Use the scatter plot before choosing either coefficient, since a single number can look identical for very different underlying patterns.

What each coefficient measures

Pearson correlation measures the strength and direction of a linear relationship. Spearman correlation measures the strength and direction of a monotonic relationship after values are converted to ranks. A Pearson r near 1 or minus 1 indicates a strong linear pattern. A Pearson r near zero indicates little linear association, but it does not rule out a curved relationship, exactly the situation the U-shaped pattern above demonstrates.

Reading the four patterns

The linear pattern gives both coefficients a value at or near 1, since a straight-line relationship is also perfectly monotonic. The monotonic-curve pattern (values that consistently rise but not in a straight line) gives Spearman a higher value than Pearson, since Spearman only needs the ranks to move consistently, not the raw values to fall on a line. The U-shaped pattern gives both coefficients a value near zero, correctly reflecting that neither a straight line nor a consistently rising rank pattern describes this relationship, even though the two variables are clearly related. The outlier pattern shows how a single extreme point can pull Pearson's r toward it much more strongly than Spearman's rank-based rho, since one extreme raw value has less leverage once it is converted to a rank.

Worked example

Consider a data set where sleep hours increase from five to nine and alertness ratings also increase, but the relationship curves rather than following a straight line. Spearman may show a strong monotonic association even when Pearson is smaller, because the straight-line model is a poor description of the actual curve. A U-shaped pattern, where alertness first drops and then rises again, could give both correlations a value near zero even though the variables are strongly related; only the scatter plot reveals a relationship a single coefficient cannot capture.

Neither coefficient identifies a cause

A positive association can arise because X affects Y, Y affects X, a third variable affects both, a selected sample creates a pattern, or pure chance produces an unusual sample. Time order, assignment, measurement quality, and confounding require design knowledge beyond what any correlation calculator can supply; see Correlation Does Not Prove Causation for the specific alternative explanations to rule out first.

Source

This comparison follows the correlation guidance in the NIST/SEMATECH e-Handbook of Statistical Methods and the shared statistical reasoning contract every StatReason engine is built against.

Limitations

Report the coefficient, sample size, and a plain description of direction; avoid universal strong-or-weak labels unless you name a field-specific convention. Neither coefficient is immune to every problem: many tied values, a very small sample, or a sharply curved pattern can still make a single correlation misleading even after this comparison.

Next action: use the **Scatter Plot Maker** before interpreting a correlation. Use the **Pearson Correlation Calculator** for a linear relationship and the **Spearman Rank Correlation Calculator** for a monotonic rank relationship.