Covariance & Correlation

Mathematics

Two measures of joint variation, and the dependence neither one sees.

Covariance measures joint variation in the original units; correlation rescales it to lie between −1 and 1.

$$\mathrm{Cov}(X, Y) = E[XY] - E[X]E[Y] \cdot \rho = \mathrm{Cov}/(\sigma X\ \sigma Y)$$

$\rho = 1$ — **Exact positive linear fit**

All points on a line.

X and Y independent — **$\mathrm{Cov}(X, Y) = 0$**

The converse fails.

$Y = X^2$ on $[-1, 1]$ — **$\rho = 0$**

Dependent, uncorrelated.

THE COMMON MISTAKE

× **Reading $\rho = 0$ as independence**

✓ **$\rho = 0$ rules out linear association only**

Correlation measures straight-line association. A perfect parabola through the origin has correlation zero and complete dependence.

Correlation is not causation, and it is not a slope either. It is scale-free; a regression coefficient is not. Educational reference.