Pearson correlation coefficient

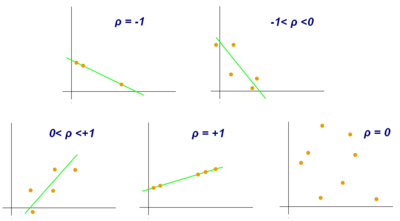
From Wikipedia, the free encyclopedia

[Jump to navigation](https://en.wikipedia.org/wiki/Pearson_correlation_coefficient#mw-head)[Jump to search](https://en.wikipedia.org/wiki/Pearson_correlation_coefficient#searchInput)

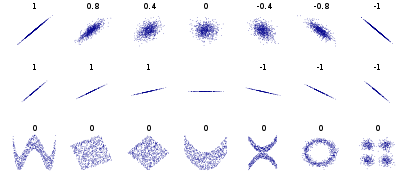
*Not to be confused with*[*Coefficient of determination*](https://en.wikipedia.org/wiki/Coefficient_of_determination)*.*

*For broader coverage of this topic, see*[*Correlation coefficient*](https://en.wikipedia.org/wiki/Correlation_coefficient)*.*

In [statistics](https://en.wikipedia.org/wiki/Statistics), the **Pearson correlation coefficient** (**PCC**, pronounced [/ˈpɪərsən/](https://en.wikipedia.org/wiki/Help:IPA/English)) ― also known as **Pearson's *r***, the **Pearson product-moment correlation coefficient** (**PPMCC**), the **bivariate correlation**,[[1]](https://en.wikipedia.org/wiki/Pearson_correlation_coefficient#cite_note-1) or colloquially simply as **the correlation coefficient**[[2]](https://en.wikipedia.org/wiki/Pearson_correlation_coefficient#cite_note-2) ― is a measure of [linear](https://en.wikipedia.org/wiki/Linear) [correlation](https://en.wikipedia.org/wiki/Correlation_and_dependence) between two sets of data. It is the ratio between the [covariance](https://en.wikipedia.org/wiki/Covariance)[[3]](https://en.wikipedia.org/wiki/Pearson_correlation_coefficient#cite_note-3)[[*circular reference*](https://en.wikipedia.org/wiki/Wikipedia:Verifiability#Wikipedia_and_sources_that_mirror_or_use_it)] of two variables and the product of their [standard deviations](https://en.wikipedia.org/wiki/Standard_deviation); thus, it is essentially a normalized measurement of the covariance, such that the result always has a value between −1 and 1. As with covariance itself, the measure can only reflect a linear correlation of variables, and ignores many other types of relationships or correlations. As a simple example, one would expect the age and height of a sample of teenagers from a high school to have a Pearson correlation coefficient significantly greater than 0, but less than 1 (as 1 would represent an unrealistically perfect correlation).

[](https://en.wikipedia.org/wiki/File:Correlation_coefficient.png)

Examples of scatter diagrams with different values of correlation coefficient (*ρ*)

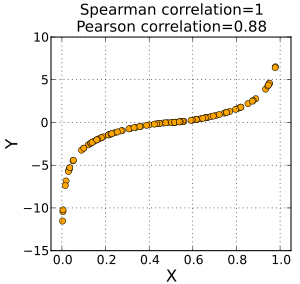
[](https://en.wikipedia.org/wiki/File:Correlation_examples2.svg)

Several sets of (*x*, *y*) points, with the correlation coefficient of *x* and *y* for each set. Note that the correlation reflects the strength and direction of a linear relationship (top row), but not the slope of that relationship (middle), nor many aspects of nonlinear relationships (bottom). N.B.: the figure in the center has a slope of 0 but in that case the correlation coefficient is undefined because the variance of *Y* is zero.

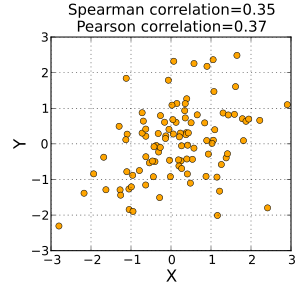
# Spearman's rank correlation coefficient

From Wikipedia, the free encyclopedia

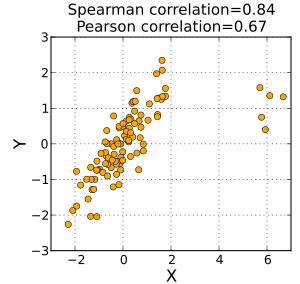
[Jump to navigation](https://en.wikipedia.org/wiki/Spearman%27s_rank_correlation_coefficient#mw-head)[Jump to search](https://en.wikipedia.org/wiki/Spearman%27s_rank_correlation_coefficient#searchInput)

[](https://en.wikipedia.org/wiki/File:Spearman_fig1.svg)

A Spearman correlation of 1 results when the two variables being compared are monotonically related, even if their relationship is not linear. This means that all data points with greater *x* values than that of a given data point will have greater *y* values as well. In contrast, this does not give a perfect Pearson correlation.

[](https://en.wikipedia.org/wiki/File:Spearman_fig2.svg)

When the data are roughly elliptically distributed and there are no prominent outliers, the Spearman correlation and Pearson correlation give similar values.

[](https://en.wikipedia.org/wiki/File:Spearman_fig3.svg)

The Spearman correlation is less sensitive than the Pearson correlation to strong outliers that are in the tails of both samples. That is because Spearman's *ρ* limits the outlier to the value of its rank.

In [statistics](https://en.wikipedia.org/wiki/Statistics), **Spearman's rank correlation coefficient** or **Spearman's *ρ***, named after [Charles Spearman](https://en.wikipedia.org/wiki/Charles_Spearman) and often denoted by the Greek letter [{\displaystyle \rho }](https://en.wikipedia.org/wiki/Rho_(letter)) (rho) or as {\displaystyle r\_{s}}, is a [nonparametric](https://en.wikipedia.org/wiki/Nonparametric_statistics) measure of [rank correlation](https://en.wikipedia.org/wiki/Rank_correlation) ([statistical dependence](https://en.wikipedia.org/wiki/Correlation_and_dependence) between the [rankings](https://en.wikipedia.org/wiki/Ranking) of two [variables](https://en.wikipedia.org/wiki/Variable_(mathematics)#Applied_statistics)). It assesses how well the relationship between two variables can be described using a [monotonic function](https://en.wikipedia.org/wiki/Monotonic_function).

The Spearman correlation between two variables is equal to the [Pearson correlation](https://en.wikipedia.org/wiki/Pearson_product-moment_correlation_coefficient) between the rank values of those two variables; while Pearson's correlation assesses linear relationships, Spearman's correlation assesses monotonic relationships (whether linear or not). If there are no repeated data values, a perfect Spearman correlation of +1 or −1 occurs when each of the variables is a perfect monotone function of the other.

Intuitively, the Spearman correlation between two variables will be high when observations have a similar (or identical for a correlation of 1) [rank](https://en.wikipedia.org/wiki/Ranking_(statistics)) (i.e. relative position label of the observations within the variable: 1st, 2nd, 3rd, etc.) between the two variables, and low when observations have a dissimilar (or fully opposed for a correlation of −1) rank between the two variables.

Spearman's coefficient is appropriate for both [continuous](https://en.wikipedia.org/wiki/Continuous_variable) and discrete [ordinal variables](https://en.wikipedia.org/wiki/Ordinal_variable).[[1]](https://en.wikipedia.org/wiki/Spearman%27s_rank_correlation_coefficient#cite_note-1)[[2]](https://en.wikipedia.org/wiki/Spearman%27s_rank_correlation_coefficient#cite_note-2) Both Spearman's {\displaystyle \rho } and [Kendall's {\displaystyle \tau }](https://en.wikipedia.org/wiki/Kendall_tau_rank_correlation_coefficient) can be formulated as special cases of a more [general correlation coefficient](https://en.wikipedia.org/wiki/General_correlation_coefficient).

