pearson Correlation Coefficient, PCC

# library link

<https://github.com/scipy/scipy/blob/main/scipy/stats/_continuous_distns.py>

# basic description

the Pearson correlation coefficient (PCC, pronounced /ˈpɪərsən/) ― also known as Pearson's r, the Pearson product-moment correlation coefficient (PPMCC), the bivariate correlation,[1] or colloquially simply as the correlation coefficient[2] ― is a measure of linear correlation between two sets of data. It is the ratio between the covariance of two variables and the product of their standard deviations; thus it is essentially a normalised 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 relationship or correlation. 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/Pearson_correlation_coefficient>

# version

* NumPy >= 1.14.6 (pip install numpy)
* Scipy >= 1.1.0 (pip install scipy)
* pandas >= 1.2.4 (pip install pandas)

# dataset

* Use sklearn built-in data (see homepage below)
* <https://scikit-learn.org/stable/install.html>

# code description

* Among the diabetes patient data in sklearn, the correlation between the bmi level and the target level is expressed as a scatterplot and the values are expressed as Pearson correlation coefficients..

# validation

* x