

## Correlation Coefficient

### What Is the Correlation Coefficient?

The correlation coefficient is a statistical measure of the strength of the relationship between the relative movements of two variables. The values range between -1.0 and 1.0. A calculated number greater than 1.0 or less than -1.0 means that there was an error in the correlation measurement. A correlation of -1.0 shows a perfect [negative correlation](#), while a correlation of 1.0 shows a perfect [positive correlation](#). A correlation of 0.0 shows no linear relationship between the movement of the two variables.

### Understanding the Correlation Coefficient

There are several types of correlation coefficients, but the one that is most common is the Pearson correlation ( $r$ ). This measures the strength and direction of the [linear relationship](#) between two variables. It cannot capture nonlinear relationships between two variables and cannot differentiate between dependent and independent variables.

A value of exactly 1.0 means there is a perfect positive relationship between the two variables. For a positive increase in one variable, there is also a positive increase in the second variable. A value of -1.0 means there is a perfect negative relationship between the two variables. This shows that the variables move in opposite directions - for a positive increase in one variable, there is a decrease in the second variable. If the correlation between two variables is 0, there is no linear relationship between them.

The strength of the relationship varies in degree based on the value of the correlation coefficient. For example, a value of 0.2 shows there is a positive correlation between two variables, but it is weak and likely unimportant. Analysts in some fields of study do not consider correlations important until the value surpasses at least 0.8. However, a correlation coefficient with an absolute value of 0.9 or greater would represent a very strong relationship.

- Correlation coefficients are used to measure the strength of the relationship between two variables.
- Pearson correlation is the one most commonly used in statistics. This measures the strength and direction of a linear relationship between two variables.
- Values always range between -1 (strong negative relationship) and +1 (strong positive relationship). Values at or close to zero imply weak or no linear relationship.
- Correlation coefficient values less than +0.8 or greater than -0.8 are not considered significant.

### **Compatibility Specifying Procedure:**

1. Compute ( r ) correlation coefficient from the equation :

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{[\sum x^2 - \frac{(\sum x)^2}{n}][\sum y^2 - \frac{(\sum y)^2}{n}]}}$$

2. Specify the type and degree of correlation depending on the following table:

r	sign	Interpreting
1	+	Perfect positive correlation
	-	Perfect negative correlation
0.70-0.99	+	Good positive correlation
	-	Good negative correlation
0.50-0.69	+	Fair positive correlation
	-	Fair negative correlation
0.01-0.49	+	Poor positive correlation
	-	Poor negative correlation
0		No Correlation

**Example:** For the raw data shown in table , find out the degree of correlation coefficient .

Y	X
2.2	-3.5
3.5	-2.5
0.9	-4.5
3.8	-2.5
3.1	-3.5
1.4	-4.5

**Sol:**

X	Y	x <sup>2</sup>	x y	y <sup>2</sup>
-3.5	2.2	12.25	-7.7	4.84
-2.5	3.5	6.25	-8.75	12.25
-4.5	0.9	20.25	-4.05	0.81
-2.5	3.8	6.25	-9.5	14.44
-3.5	3.1	12.25	-10.85	9.61
-4.5	1.4	20.25	-6.3	1.96
$\sum x = -21$	$\sum y = 14.9$	$\sum x^2 = 77.5$	$\sum xy = -47.15$	$\sum y^2 = 43.91$

Sample Size (n) = 6

$$\text{Correlation Coefficient } r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{[\sum x^2 - \frac{(\sum x)^2}{n}][\sum y^2 - \frac{(\sum y)^2}{n}]}}$$

$$r = \frac{-47.15 - \frac{(-21)(14.9)}{6}}{\sqrt{\left[77.5 - \frac{(-21)^2}{6}\right]\left[43.91 - \frac{(14.9)^2}{6}\right]}} = 0.95 \text{ Good Positive Correlation}$$