

How do you perform a correlation test in R, and what are some examples of using it?

Authored by
stats writer

April 21, 2024

RECOMMENDED CITATION

stats writer (2024). *How do you perform a correlation test in R, and what are some examples of using it?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=137610>

A correlation test in R is a statistical method used to measure the strength and direction of the relationship between two variables. This test is commonly used to determine if there is a linear relationship between two continuous variables or to identify patterns and trends in data.

To perform a correlation test in R, the user must first load the necessary packages, such as "stats" or "psych," which contain the functions for conducting the test. Then, the user can use the "cor.test()" function to specify the variables and desired significance level for the test. The output of the test includes the correlation coefficient, p-value, and a scatterplot with the regression line.

There are various examples of using a correlation test in R, such as analyzing the relationship between income and education level, or the correlation between temperature and ice cream sales. It can also be used to determine the correlation between stock prices and economic indicators, or to investigate the relationship between customer satisfaction and sales. Overall, a correlation test in R is a valuable tool for understanding relationships between variables and making informed decisions based on data analysis.

Perform a Correlation Test in R (With Examples)

One way to quantify the relationship between two variables is to use the Pearson correlation coefficient, which is a measure of the linear association between two variables.

It always takes on a value between -1 and 1 where:

**-1 indicates a perfectly negative linear correlation between two variables
0 indicates no linear correlation between two variables
1 indicates a perfectly positive linear correlation between two variables**

To determine if a correlation coefficient is statistically

significant, you can calculate the corresponding t-score and p-value.

The formula to calculate the t-score of a correlation coefficient (r) is:

$$t = r * \sqrt{n-2} / \sqrt{1-r^2}$$

The p-value is calculated as the corresponding two-sided p-value for the t-distribution with n-2 degrees of freedom.

Example: Correlation Test in R

To determine if the correlation coefficient between two variables is statistically significant, you can perform a correlation test in R using the following syntax:

```
cor.test(x, y, method=c("pearson", "kendall",  
"spearman"))
```

where:

x, y: Numeric vectors of data.
method: Method used to calculate correlation between two vectors. Default is "pearson."

For example, suppose we have the following two vectors in R:

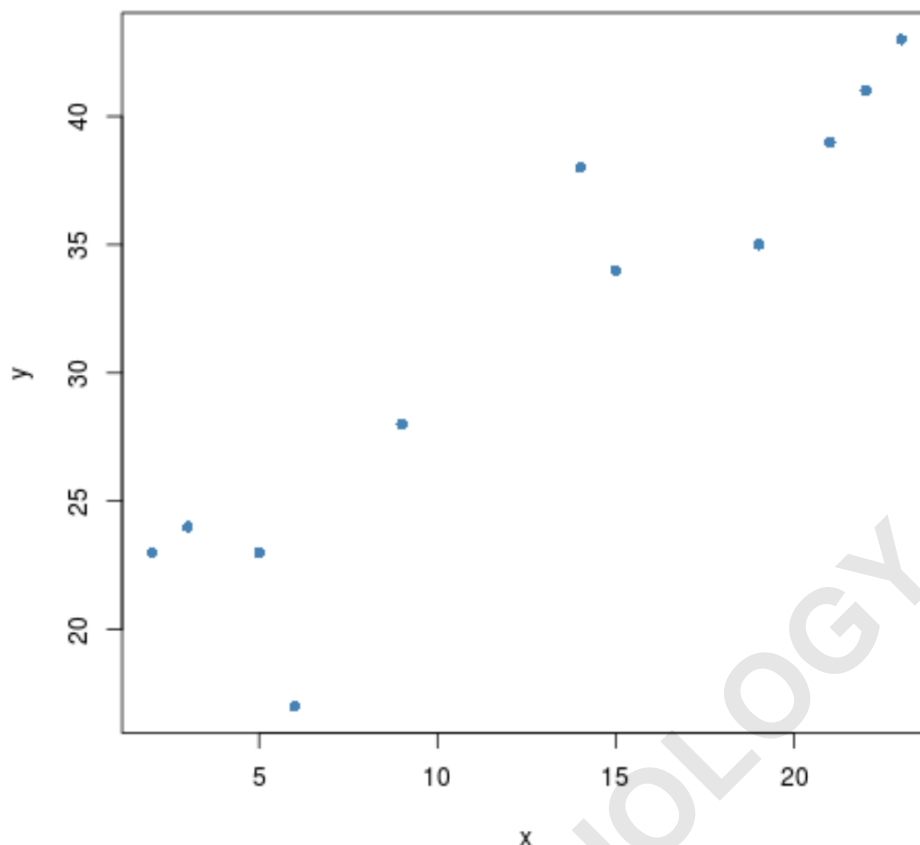
```
x <- c(2, 3, 3, 5, 6, 9, 14, 15, 19, 21, 22, 23)
```

```
y <- c(23, 24, 24, 23, 17, 28, 38, 34, 35, 39, 41, 43)
```

Before we perform a correlation test between the two variables, we can create a quick scatterplot to view their relationship:

```
#create scatterplot
```

```
plot(x, y, pch=16)
```



To see if this correlation is statistically significant, we can perform a correlation test:

```
#perform correlation test between the two vectors  
cor.test(x, y)
```

Pearson's product-moment correlation

data: x and y

t = 7.8756, df = 10, p-value = 1.35e-05

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

0.7575203 0.9799783

sample estimates:

cor

0.9279869

The correlation coefficient between the two vectors turns out to be 0.9279869.

The test statistic turns out to be 7.8756 and the corresponding p-value is 1.35e-05.

Since this value is less than .05, we have sufficient evidence to say that the correlation between the two variables is statistically significant.

The following tutorials provide additional information about correlation coefficients:

[An Introduction to the Pearson Correlation Coefficient](#)