

How can I calculate rolling correlation in R?

Authored by
stats writer

April 21, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I calculate rolling correlation in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=137793>

Rolling correlation in R refers to the process of calculating the correlation between two variables over a moving window of data points. This can be useful in identifying the relationship between two variables over time, rather than just at a single point. In R, this can be achieved by using the "rollapply" function from the "zoo" package, which allows for the calculation of the correlation for each window of data. By specifying the window size and the number of data points to be shifted, the rolling correlation can be calculated and the results can be plotted or further analyzed. This method provides a way to visualize and analyze the changing correlation between two variables, which can be helpful in identifying trends and patterns in the data.

Calculate Rolling Correlation in R

Rolling correlations are correlations between two time series on a rolling window. One benefit of this type of correlation is that you can visualize the correlation between two time series over time.

This tutorial explains how to calculate rolling correlations in R.

How to Calculate Rolling Correlations in R

Suppose we have the following data frame that display the total number of products sold for two different products (x and y) during a 15-month period:

```
#create data
```

```
data <- data.frame(month=1:15,
```

```
x=c(13, 15, 16, 15, 17, 20, 22, 24, 25, 26, 23, 24, 23, 22,  
20),
```

```
y=c(22, 24, 23, 27, 26, 26, 27, 30, 33, 32, 27, 25, 28, 26, 28))
```

```
#view first six rows
```

```
head(data)
```

```
month x y
```

```
1 1 13 22
```

```
2 2 15 24
```

```
3 3 16 23
```

```
4 4 15 27
```

```
5 5 17 26
```

```
6 6 20 26
```

To calculate a rolling correlation in R, we can use the rollapply() function from the zoo package.

This function uses the following syntax:

```
rollapply(data, width, FUN, by.column=TRUE)
```

where:

data: Name of the data frame
width: Integer specifying the window width for the rolling correlation
FUN: The function to be applied.
by.column: Specifies whether to

apply the function to each column separately. This is TRUE by default, but to calculate a rolling correlation we need to specify this to be FALSE.

Here's how to use this function to calculate the 3-month rolling correlation in sales between product x and product y:

```
#calculate 3-month rolling correlation between sales for  
x and y  
rollapply(data, width=3, function(x) cor(x,x),  
by.column=FALSE)
```

```
0.6546537 -0.6933752 -0.2401922 -0.8029551 0.8029551  
0.9607689  
0.9819805 0.6546537 0.8824975 0.8170572 -0.9449112  
-0.3273268  
-0.1889822
```

This function returns the correlation between the two product sales for the previous 3 months. For example:

The correlation in sales during months 1 through 3 was 0.6546537. The correlation in sales during months 2 through 4 was -0.6933752. The correlation in sales

during months 3 through 5 was -0.2401922.

And so on.

We can easily adjust this formula to calculate the rolling correlation for a different time period. For example, the following code shows how to calculate the 6-month rolling correlation in sales between the two products:

```
#calculate 6-month rolling correlation between sales for  
x and y
```

```
rollapply(data, width=6, function(x) cor(x,x),  
by.column=FALSE)
```

```
0.5587415 0.4858553 0.6931033 0.7564756 0.8959291  
0.9067715 0.7155418  
0.7173740 0.7684468 0.4541476
```

The correlation in sales during months 1 through 6 was 0.5587415. The correlation in sales during months 2 through 7 was 0.4858553. The correlation in sales during months 3 through 8 was 0.6931033.

And so on.

Notes

Keep the following in mind when using the `rollapply()` function:

The width (i.e. the rolling window) should be 3 or greater in order to calculate correlations. In the formulas above, we used `cor(x,x)` because the two columns that we wanted to calculate correlations between were in position 2 and 3. Adjust these numbers if the columns you're interested in are located in different positions.

How to Calculate Rolling Correlation in Excel