How can I use the Spread function in R, and can you provide some examples?

Authored by

stats writer

May 1, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I use the Spread function in R, and can you provide some examples?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=141645

ARABPSYCHOLOGY.COM

The Spread function in R is a useful tool for organizing data in a wide format. It allows the user to convert data from a long format, where each observation has its own row, to a wide format, where each variable has its own column. This can be particularly helpful when dealing with large datasets that have multiple observations for each variable. Some examples of using the Spread function in R include organizing survey data, converting time series data into a more manageable format, and creating pivot tables for analysis. Overall, the Spread function in R provides a convenient and efficient way to restructure and manipulate data for various analytical purposes.

Use Spread Function in R (With Examples)

The spread() function from the package can be used to "spread" a key-value pair across multiple columns.

This function uses the following basic syntax:

spread(data, key value)

where:

data: Name of the data framekey: Column whose values will become variable namesvalue: Column where values will fill under new variables created from key

The following examples show how to use this function in practice.

Example 1: Spread Values Across Two Columns

Suppose we have the following data frame in R:

ARABPSYCHOLOGY.COM

```
#create data frame
df <- data.frame(player=rep(c('A', 'B'), each=4),
year=rep(c(1, 1, 2, 2), times=2),
stat=rep(c('points', 'assists'), times=4),
amount=c(14, 6, 18, 7, 22, 9, 38, 4))

#view data frame
df

player year stat amount
1 A 1 points 14
2 A 1 assists 6
3 A 2 points 18
4 A 2 assists 7
5 B 1 points 22
6 B 1 assists 9
7 B 2 points 38
8 B 2 assists 4
```

We can use the spread() function to turn the values in the stat column into their own columns:

```
library(tidyr)
#spread stat column across multiple columns
```

ARABPSYCHOLOGY.COM

spread(df, key=stat, value=amount)

player year assists points
1 A 1 6 14
2 A 2 7 18
3 B 1 9 22
4 B 2 4 38

Example 2: Spread Values Across More Than Two Columns

Suppose we have the following data frame in R:

```
#create data frame
df2 <- data.frame(player=rep(c('A'), times=8),
year=rep(c(1, 2), each=4),
stat=rep(c('points', 'assists', 'steals', 'blocks'), times=2),
amount=c(14, 6, 2, 1, 29, 9, 3, 4))

#view data frame
df2
```

player year stat amount
1 A 1 points 14
2 A 1 assists 6
3 A 1 steals 2
4 A 1 blocks 1

ARABPSYCHOLOGY.COM

5 A 2 points 29

6 A 2 assists 9

7 A 2 steals 3

8 A 2 blocks 4

We can use the spread() function to turn the four unique values in the stat column into four new columns:

```
library(tidyr)
#spread stat column across multiple columns
spread(df2, key=stat, value=amount)
```

player year assists blocks points steals

1 A 1 6 1 14 2

2 A 2 9 4 29 3

Every column is a variable.Every row is an observation.Every cell is a single value.

The tidyr package uses four core functions to create tidy data:

1. The spread() function.

2. The function.

ARABPSYCHOLOGY.COM

3. The function.

4. The function.

If you can master these four functions, you will be able to create "tidy" data from any data frame.

ARABPSYCHOLOGY.COM