

How can I sum columns in R based on a condition?

Authored by
stats writer

May 5, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I sum columns in R based on a condition?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=143045>

Summing columns in R based on a condition involves using the conditional statements and functions in R to specify the criteria for selecting the columns and the values to be summed. This process can be achieved by first identifying the columns to be included in the sum, then creating a logical condition using the subset function, and finally using the sum function to calculate the total. By following this approach, the user can accurately and efficiently sum columns in R based on specific conditions.

Sum Columns Based on a Condition in R

You can use the following basic syntax to sum columns based on condition in R:

```
#sum values in column 3 where col1 is equal to 'A'  
sum(df)
```

The following examples show how to use this syntax in practice with the following data frame:

```
#create data frame  
df <- data.frame(conference = c('East', 'East', 'East',  
'West', 'West', 'East'),  
team = c('A', 'A', 'A', 'B', 'B', 'C'),  
points = c(11, 8, 10, 6, 6, 5),  
rebounds = c(7, 7, 6, 9, 12, 8))
```

```
#view data frame
```

```
df
```

conference team points rebounds

1 East A 11 7

2 East A 8 7

3 East A 10 6

4 West B 6 9

5 West B 6 12

6 East C 5 8

Example 1: Sum One Column Based on One Condition

The following code shows how to find the sum of the points column for the rows where team is equal to 'A':

```
#sum values in column 3 (points column) where team is  
equal to 'A'  
sum(df
```

29

The following code shows how to find the sum of the rebounds column for the rows where points is greater than 9:

```
#sum values in column 4 (rebounds column) where  
points is greater than 9
```

sum(df)

13

Example 2: Sum One Column Based on Multiple Conditions

The following code shows how to find the sum of the points column for the rows where team is equal to 'A' *and* conference is equal to 'East':

```
#sum values in column 3 (points column) where team is 'A' and conference is 'East'
```

```
sum(df)
```

29

Note that the **&** operator stands for "and" in R.

Example 3: Sum One Column Based on One of Several Conditions

The following code shows how to find the sum of the points column for the rows where team is equal to 'A' *or* 'C':

```
#sum values in column 3 (points column) where team is 'A' or 'C'
```

sum(df)

34

Note that the | operator stands for "or" in R.

The following tutorials explain how to perform other common functions in R:

ARABPSYCHOLOGY.COM