

How can we calculate the mean of multiple columns in R?

Authored by
stats writer

April 20, 2024

RECOMMENDED CITATION

stats writer (2024). *How can we calculate the mean of multiple columns in R?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=137423>

Calculating the mean of multiple columns in R is a process of finding the average value of a set of numerical data from different columns. This can be achieved by using the "mean()" function in R, which takes in multiple columns as arguments and returns the mean value for each column. The mean is calculated by adding all the values in a column and dividing it by the total number of values in that column. By using the "mean()" function, this calculation can be done efficiently and accurately for multiple columns at once, providing a comprehensive overview of the data set. This process is particularly useful in analyzing and summarizing large data sets in an organized and efficient manner.

Calculate the Mean of Multiple Columns in R

Often you may want to calculate the mean of multiple columns in R. Fortunately you can easily do this by using the `colMeans()` function.

`colMeans(df)`

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

Using `colMeans()` to Find the Mean of Multiple Columns

The following code shows how to use the `colMeans()` function to find the mean of every column in a data frame:

```
#create data frame
```

```
df <- data.frame(var1=c(1, 3, 3, 4, 5),
```

```
var2=c(7, 7, 8, 3, 2),
```

```
var3=c(3, 3, 6, 6, 8),  
var4=c(1, 1, 2, 8, 9))
```

```
#find mean of each column  
colMeans(df)
```

```
var1 var2 var3 var4  
3.2 5.4 5.2 4.2
```

We can also specify *which* columns to find the mean for:

```
#find the mean of columns 2 and 3  
colMeans(df)
```

```
var2 var3  
5.4 5.2
```

```
#find the mean of the first three columns  
colMeans(df)
```

```
var1 var2 var3  
3.2 5.4 5.2
```

If there happen to be some columns that aren't numeric,

you can use `sapply()` to specify that you'd only like to find the mean of columns that are numeric:

```
#create data frame
```

```
df <- data.frame(var1=c(1, 3, 3, 4, 5),  
var2=c(7, 7, 8, 3, 2),  
var3=c(3, 3, 6, 6, 8),  
var4=c(1, 1, 2, 8, 9),  
var5=c('a', 'a', 'b', 'b', 'c'))
```

```
#find mean of only numeric columns
```

```
colMeans(df)
```

```
var1 var2 var3 var4
```

```
3.2 5.4 5.2 4.2
```

And if there happen to be missing values in any columns, you can use the argument `na.rm=TRUE` to ignore missing values when calculating the means:

```
#create data frame with some missing values
```

```
df <- data.frame(var1=c(1, 3, NA, NA, 5),  
var2=c(7, 7, 8, 3, 2),  
var3=c(3, 3, 6, 6, 8),  
var4=c(1, 1, 2, 8, NA))
```

**#find mean of each column and ignore missing values
colMeans(df, na.rm=TRUE)**

var1 var2 var3 var4

3.0 5.4 5.2 3.0

How to Loop Through Column Names in R

How to Sum Specific Columns in R

ARABPSYCHOLOGY.COM