How do I calculate the mean of a column in R with examples?

Authored by

stats writer

June 28, 2024

Calculating the mean of a column in R is a simple process that can be done using built-in functions and packages. The mean is a measure of central tendency that represents the average value of a set of numbers. In R, the mean can be calculated using the "mean()" function from the base package or the "mean()" function from the "dplyr" package. Both functions take a vector or a column from a data frame as input and return the mean value as the output. For example, to calculate the mean of a column "age" in a data frame "df", the syntax would be "mean(df$age)". Additionally, the "summary()" function can also be used to calculate the mean along with other descriptive statistics for multiple columns in a data frame. Overall, calculating the mean in R is a straightforward process that can be applied to any numerical data to gain insights into the data's central tendency.

Calculate the Mean of a Column in R (With Examples)

You can use one of the following methods to calculate the mean of a column in R:

#calculate mean using column name
mean(df$my_column)

#calculate mean using column name (ignore missing values)
mean(df$my_column, na.rm=TRUE)

#calculate mean using column position
mean(df)

#calculation mean of all numeric columns
colMeans(df)

The following examples show how to use each method with the following data frame in R:

```
#create data frame
df <- data.frame(team=c('A', 'A', 'A', 'B', 'B', 'B'),
points=c(99, 90, 93, 86, 88, 82),
assists=c(33, 28, 31, 39, NA, 30))

#view data frame
df

team points assists
1 A 99 33
2 A 90 28
3 A 93 31
4 B 86 39
5 B 88 NA
6 B 82 30
```


Example 1: Calculate Mean Using Column Name

The following code shows how to calculate the mean of the 'points' column using the column name:

```
#calculate mean of 'points' column
```

mean(df$points)

89.66667

The mean value in the 'points' column is 89.66667.

Example 2: Calculate Mean Using Column Name (Ignore Missing Values)

If we attempt to calculate the mean of a column that has missing values, we'll receive NA as a result:

#attempt to calculate mean of 'assists' column
mean(df$assists)

NA

We must use na.rm=TRUE to ignore missing values when calculating the column mean:

#calculate mean of 'assists' column and ignore missing values
mean(df$assists, na.rm=TRUE)

32.2

The mean value in the 'assists' column is 32.2.

Example 3: Calculate Mean Using Column Position

The following code shows how to calculate the mean of the column in index position 2:

```
#calculate mean of column in index position 2
mean(df)

89.66667
```

Example 4: Calculate Mean of All Numeric Columns

The following code shows how to calculate the mean of all numeric columns in the data frame:

```
#calculate mean of all numeric columns
colMeans(df, na.rm=TRUE)

points assists
89.66667 32.20000
```

The output displays the mean value of each numeric column in the data frame.

Additional Resources

The following tutorials explain how to calculate other

mean values in R: