# How can I use the R function drop_na to drop rows with missing values?

Authored by

**stats writer**

June 27, 2024

The R function drop_na is a useful tool for removing rows from a data frame that contain missing values. This function allows users to easily filter out incomplete or invalid data, providing a clean and accurate dataset for analysis. By specifying the data frame and the desired columns to be checked, drop_na will identify and remove any rows that have missing values in those columns. This allows for a more efficient and precise analysis without having to manually remove these rows. Overall, the drop_na function provides a convenient solution for handling missing data in R.

## R: Use drop_na to Drop Rows with Missing Values

## You can use the drop_na() function from the package in R to drop rows with missing values in a data frame.

## There are three common ways to use this function:

## Method 1: Drop Rows with Missing Values in Any Column

## df %>% drop_na()

## Method 2: Drop Rows with Missing Values in Specific Column

## df %>% drop_na(col1)

## Method 3: Drop Rows with Missing Values in One of Several Specific Columns

## df %>% drop_na(c(col1, col2))

**The following examples show how to use each of these methods in practice with the following data frame:**

```
#create data frame
df <- data.frame(points=c(10, NA, 15, 15, 14, 16),
assists=c(4, NA, 4, NA, 9, 3),
rebounds=c(NA, 5, 10, 7, 7, NA))

#view data frame
df

points assists rebounds
1 10 4 NA
2 NA NA 5
3 15 4 10
4 15 NA 7
5 14 9 7
6 16 3 NA
```

### Example 1: Drop Rows with Missing Values in Any Column

**The following code shows how to use drop_na() to drop rows with missing values in any column:**

```
library(tidyr)
```

```
#drop rows with missing values in any column
df %>% drop_na()

points assists rebounds
1 15 4 10
2 14 9 7
```

The only rows left are the ones with no missing values in any column.

### Example 2: Drop Rows with Missing Values in Specific Column

The following code shows how to use drop_na() to drop rows with missing values in the rebounds column:

```
library(tidyr)

#drop rows with missing values in rebounds column
df %>% drop_na(rebounds)

points assists rebounds
1 NA NA 5
2 15 4 10
3 15 NA 7
4 14 9 7
```

## Example 3: Drop Rows with Missing Values in One of Several Specific Columns

**The following code shows how to use drop_na() to drop rows with missing values in the points or assists columns:**

```
library(tidyr)

#drop rows with missing values in the points or assists columns
df %>% drop_na(c(points, assists))
```

```
points assists rebounds
1 10 4 NA
2 15 4 10
3 14 9 7
4 16 3 NA
```

**The only rows left are the ones with no missing values in the points*or*assists columns.**

**Note: You can find the complete online documentation for the drop_na() method .**