How can the melt() function be used in R?

Authored by

stats writer

June 26, 2024

The melt() function in R is a useful tool for reshaping data frames from a wide format to a long format. This function allows the user to specify which columns should be kept as identification variables and which columns should be converted into a single column containing the values. This is particularly useful for data sets that have multiple measurements for each observation, as it allows for easier analysis and visualization. Additionally, the melt() function can be used for data aggregation and restructuring, making it a valuable function for data manipulation in R.

Use the melt() Function in R

You can use the melt() function from the reshape2 package in R to convert a data frame from a wide format to a long format.

A wide format contains values that *do not* repeat in the first column.

A long format contains values that *do* repeat in the first column.

For example, consider the following two datasets that contain the exact same data expressed in different formats:

Wide Format

Team	Points	Assists	Rebounds
A	88	12	22
B	91	17	28
C	99	24	30
D	94	28	31

Long Format

Team	Variable	Value
A	Points	88
A	Assists	12
A	Rebounds	22
B	Points	91
B	Assists	17
B	Rebounds	28
C	Points	99
C	Assists	24
C	Rebounds	30
D	Points	94
D	Assists	28
D	Rebounds	31

The melt() function uses the following basic syntax to convert a data frame in a wide format to a long format:

melt(df, id='team')

The id argument specifies which variable to use as the first column in the data frame whose values will be repeated.

The following example shows how to use this function in practice.

Example: How to Use melt() in R

Suppose we have the following data frame in R that is currently in a wide format:

```
#create data frame in wide format
df <- data.frame(team=c('A', 'B', 'C', 'D'),
points=c(88, 91, 99, 94),
assists=c(12, 17, 24, 28),
rebounds=c(22, 28, 30, 31))

#view data frame
df

team points assists rebounds
1 A 88 12 22
2 B 91 17 28
3 C 99 24 30
4 D 94 28 31
```

We can use the melt() function to quickly convert the data frame to a long format:

```
library(reshape2)

#use melt() to convert data frame from wide to long
```

format

long_df <- melt(df, id='team')

#view long data frame

long_df

team variable value

1 A points 88

2 B points 91

3 C points 99

4 D points 94

5 A assists 12

6 B assists 17

7 C assists 24

8 D assists 28

9 A rebounds 22

10 B rebounds 28

11 C rebounds 30

12 D rebounds 31

Notice that the data frame is now in a long format.

The columns points, assists, and rebounds have all been compressed into a single column called variable while their values have all been compressed into a

single column called values.

Feel free to rename the columns in the resulting data frame by using the names() function:

```
#rename columns in long_df
names(long_df) <- c('team', 'metric', 'amount')

#view updated data frame
long_df

team metric amount
1 A points 88
2 B points 91
3 C points 99
4 D points 94
5 A assists 12
6 B assists 17
7 C assists 24
8 D assists 28
9 A rebounds 22
10 B rebounds 28
11 C rebounds 30
12 D rebounds 31
```

Notice that the columns have been renamed.