

How can I use the `dcast` function from the `data.table` package in R?

Authored by
stats writer

June 25, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I use the `dcast` function from the `data.table` package in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=151977>

The `dcast` function from the `data.table` package in R allows for efficient and flexible data reshaping by transforming data from long to wide format. This function takes in a data.table object, along with specified variables for row and column indices, and produces a new data.table with columns representing the unique values of the chosen variable(s). This can be useful for summarizing and organizing large datasets for further analysis. By utilizing the `dcast` function, users can easily manipulate their data and extract valuable insights.

Use dcast Function from data.table in R

You can use the dcast function from the data.table package in R to reshape a data frame from a to a wide format.

This function is particularly useful when you want to summarize specific variables in a data frame, grouped by other variables.

The following examples show how to use the dcast function in practice with the following data frame in R:

```
library(data.table)
```

```
#create data frame
```

```
df <- data.frame(team=c('A', 'A', 'A', 'A', 'B', 'B', 'B', 'B'),  
position=c('G', 'G', 'F', 'F', 'G', 'G', 'F', 'F'),  
points=c(18, 13, 10, 12, 16, 25, 24, 31),  
assists=c(9, 8, 8, 5, 12, 15, 10, 7))
```

#convert data frame to data table

dt <- setDT(df)

#view data table

dt

team position points assists

1: A G 18 9

2: A G 13 8

3: A F 10 8

4: A F 12 5

5: B G 16 12

6: B G 25 15

7: B F 24 10

8: B F 31 7

Example 1: Calculate Metric for One Variable, Grouped by Other Variables

The following code shows how to use the dcast function to calculate the mean points value, grouped by the team and position variables:

library(data.table)

#calculate mean points value by team and position

```
dt_new <- dcast(dt,  
team + position ~ .,  
fun.aggregate = mean,  
value.var = 'points')
```

```
#view results
```

```
dt_new
```

```
team position .
```

```
1: A F 11.0
```

```
2: A G 15.5
```

```
3: B F 27.5
```

```
4: B G 20.5
```

Example 2: Calculate Multiple Metrics for One Variable, Grouped by Other Variables

The following code shows how to use the `dcast` function to calculate the mean points value and the max points value, grouped by the team and position variables:

```
library(data.table)
```

```
#calculate mean and max points values by team and  
position
```

```
dt_new <- dcast(dt,  
team + position ~ .,  
fun.aggregate = list(mean, max),  
value.var = 'points')
```

```
#view results
```

```
dt_new
```

```
team position points_mean points_max
```

```
1: A F 11.0 12
```

```
2: A G 15.5 18
```

```
3: B F 27.5 31
```

```
4: B G 20.5 25
```

Example 3: Calculate Metric for Multiple Variables, Grouped by Other Variables

The following code shows how to use the `dcast` function to calculate the mean points value and mean assists value, grouped by the team and position variables:

```
library(data.table)
```

```
#calculate mean and max points values by team and  
position
```

```
dt_new <- dcast(dt,  
team + position ~ .,  
fun.aggregate = mean,  
value.var = c('points', 'assists'))
```

```
#view results
```

```
dt_new
```

```
team position points assists
```

```
1: A F 11.0 6.5
```

```
2: A G 15.5 8.5
```

```
3: B F 27.5 8.5
```

```
4: B G 20.5 13.5
```

The following tutorials provide additional information about data tables: