

How can I select only the numeric columns in R using the dplyr package?

Authored by
stats writer

June 27, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I select only the numeric columns in R using the dplyr package?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=155033>

The dplyr package in R allows for efficient data manipulation and transformation. To select only the numeric columns in a dataset using this package, the `select()` function can be used. This function takes in the dataset as the first argument, followed by the column names to be selected. By using the `is.numeric()` function within the `select()` function, only the columns that have numeric data will be selected. This allows for a quick and simple way to filter out non-numeric columns in a dataset and work only with the desired numerical data.

Select Only Numeric Columns in R Using dplyr

You can use the following function from the package to select only numeric columns from a data frame in R:

```
df %>% select(where(is.numeric))
```

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

Example: Select Only Numeric Columns Using dplyr

Suppose we have the following data frame in R that contains information about various basketball players:

```
#create data frame
```

```
df <- data.frame(team=c('A', 'B', 'C', 'D', 'E'),  
points=c(22, 34, 30, 12, 18),  
assists=c(7, 9, 9, 12, 14),  
rebounds=c(5, 10, 10, 8, 8))
```

```
#view data frame
```

```
df
```

```
team points assists rebounds
```

```
1 A 22 7 5
```

```
2 B 34 9 10
```

```
3 C 30 9 10
```

```
4 D 12 12 8
```

```
5 E 18 14 8
```

We can use the following syntax to select only the numeric columns from the data frame:

```
library(dplyr)
```

```
#select only the numeric columns from the data frame
```

```
df %>% select(where(is.numeric))
```

```
points assists rebounds
```

```
1 22 7 5
```

```
2 34 9 10
```

```
3 30 9 10
```

```
4 12 12 8
```

```
5 18 14 8
```

Notice that only the three numeric columns have been selected - points, assists, and rebounds.

We can verify that these columns are numeric by using the `str()` function to display the data type of each variable in the data frame:

```
#display data type of each variable in data frame  
str(df)
```

```
'data.frame': 5 obs. of 4 variables:
```

```
$ team : chr "A" "B" "C" "D" ...
```

```
$ points : num 22 34 30 12 18
```

```
$ assists : num 7 9 9 12 14
```

```
$ rebounds: num 5 10 10 8 8
```

From the output we can see that team is a character variable while points, assists, and rebounds are all numeric.

Related:

The following tutorials explain how to perform other common tasks using dplyr: