

How can I select columns by name using dplyr in R?

Authored by
stats writer

June 29, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I select columns by name using dplyr in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=159832>

Dplyr is a popular R package used for data manipulation and analysis. With dplyr, you can easily select columns by their names from a data frame. This can be achieved using the `select()` function, where you specify the columns you want to select by their names as arguments. This method allows for a more efficient and streamlined approach to selecting columns, compared to traditional methods. By using dplyr, you can quickly and easily extract the necessary columns for further analysis or visualization.

Select Columns by Name Using dplyr

You can use the following methods to select columns of a data frame by name in R using the package:

Method 1: Select Specific Columns by Name

```
df %>% select(var1, var3)
```

Method 2: Select a Range of Columns by Name

```
df %>% select(var1:var3)
```

Method 3: Select All Columns Except Certain Columns

```
df %>% select(-c(var1, var3))
```

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

```
#create data frame
```

```
df <- data.frame(points=c(1, 5, 4, 5, 5, 7, 8),  
rebounds=c(10, 3, 3, 2, 6, 7, 12),  
assists=c(5, 5, 7, 6, 7, 9, 15),  
blocks=c(1, 1, 0, 4, 3, 2, 10))#view data frame  
df
```

```
points rebounds assists blocks
```

```
1 1 10 5 1
```

```
2 5 3 5 1
```

```
3 4 3 7 0
```

```
4 5 2 6 4
```

```
5 5 6 7 3
```

```
6 7 7 9 2
```

```
7 8 12 15 10
```

Example 1: Select Specific Columns by Name

We can use the following code to select only the points and assists columns:

```
library(dplyr)
```

```
#select only points and assists columns
```

```
df %>% select(points, assists)
```

```
points assists
```

```
1 1 5
2 5 5
3 4 7
4 5 6
5 5 7
6 7 9
7 8 15
```

Notice that only the points and assists columns are returned.

Example 2: Select a Range of Columns by Name

We can use the following code to select all columns between the names points and assists.

```
library(dplyr)
```

```
#select all columns between points and assists
df %>% select(points:assists)
```

```
points rebounds assists
```

```
1 1 10 5
2 5 3 5
3 4 3 7
4 5 2 6
```

5 5 6 7

6 7 7 9

7 8 12 15

A range of columns is returned, starting with the points column and ending with the assists column.

Example 3: Select All Columns Except Certain Columns

We can use the following code to select all columns except the points and assists columns.

```
library(dplyr)
```

```
#select all columns except points and assists columns  
df %>% select(-c(points, assists))
```

```
rebounds blocks
```

```
1 10 1
```

```
2 3 1
```

```
3 3 0
```

```
4 2 4
```

```
5 6 3
```

```
6 7 2
```

```
7 12 10
```

All of the columns are returned except the points and assists columns.

Note: You can find the complete documentation for the select function in dplyr .

Additional Resources

The following tutorials explain how to perform other common operations in dplyr:

ARABPSYCHOLOGY.COM