

How can we subset a data frame in R using multiple conditions?

Authored by
stats writer

June 26, 2024

RECOMMENDED CITATION

stats writer (2024). *How can we subset a data frame in R using multiple conditions?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=153254>

Subsetting a data frame in R using multiple conditions allows us to extract specific rows or columns from a data frame that meet a set of criteria. This can be achieved by using the `subset()` function, which takes in the data frame, followed by the conditions to be met. Multiple conditions can be specified using logical operators such as "AND" (&) and "OR" (|). This technique is useful for data analysis and allows for efficient extraction of relevant information from large data sets.

Subset Data Frame in R by Multiple Conditions

You can use the following methods to subset a data frame by multiple conditions in R:

Method 1: Subset Data Frame Using "OR" Logic

```
df_sub <- subset(df, team == 'A' | points < 20)
```

This particular example will subset the data frame for rows where the team column is equal to 'A' or the points column is less than 20.

Method 2: Subset Data Frame Using "AND" Logic

```
df_sub <- subset(df, team == 'A' & points < 20)
```

This particular example will subset the data frame for rows where the team column is equal to 'A' and the points column is less than 20.

This tutorial explains how to use each method in practice with the following data frame:

```
#create data frame
```

```
df <- data.frame(team=c('A', 'A', 'A', 'B', 'B', 'B'),  
position=c('Guard', 'Guard', 'Forward',  
'Guard', 'Forward', 'Forward'),  
points=c(22, 25, 19, 22, 12, 35))
```

```
#view data frame
```

```
df
```

```
team position points
```

```
1 A Guard 22
```

```
2 A Guard 25
```

```
3 A Forward 19
```

```
4 B Guard 22
```

```
5 B Forward 12
```

```
6 B Forward 35
```

Example 1: Subset Data Frame Using "OR" Logic

The following code shows how to subset the data frame for rows where the team column is equal to 'A' or the points column is less than 20:

#subset data frame where team is 'A' or points is less than 20

```
df_sub <- subset(df, team == 'A' | points <20)
```

#view subset

```
df_sub
```

team position points

1 A Guard 22

2 A Guard 25

3 A Forward 19

5 B Forward 12

Each of the rows in the subset either have a value of 'A' in the team column or have a value in the points column less than 20.

Note: The | symbol represents "OR" in R.

In this example, we only included one "OR" symbol in the subset() function but we could include as many as we like to subset based on even more conditions.

Example 2: Subset Data Frame Using "AND" Logic

The following code shows how to subset the data frame

for rows where the team column is equal to 'A' and the points column is less than 20:

```
#subset data frame where team is 'A' and points is less than 20
```

```
df_sub <- subset(df, team == 'A' & points <20)
```

```
#view subset
```

```
df_sub
```

```
team position points
```

```
3 A Forward 19
```

Notice that the resulting subset only contains one row.

This is because only one row has a value of 'A' in the team column and has a value in the points column less than 20.

Note: The & symbol represents "AND" in R.

In this example, we only included one "AND" symbol in the subset() function but we could include as many as we like to subset based on even more conditions.