

How can rows be selected based on a condition in R, and what are some examples of doing so?

Authored by
stats writer

May 14, 2024

RECOMMENDED CITATION

stats writer (2024). *How can rows be selected based on a condition in R, and what are some examples of doing so?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=144108>

Rows can be selected based on a condition in R by using the "subset" function or by using a logical vector as an index to the data frame. The subset function allows for the selection of rows based on a logical condition, such as filtering for rows that meet a certain criteria or have a specific value in a particular column. A logical vector can also be used to subset rows by specifying which rows to keep and which to exclude.

For example, to select rows in a data frame where the value in the "age" column is greater than 30, the subset function can be used as follows:

```
subset(dataframe, age > 30)
```

To select rows based on a logical vector, the vector must have the same length as the number of rows in the data frame. For instance, if the data frame has 10 rows, the logical vector should also have 10 elements. If we want to select the first, third, and fifth rows, we can create a logical vector as follows:

```
c(TRUE, FALSE, TRUE, FALSE, TRUE, FALSE, FALSE, FALSE, FALSE, FALSE)
```

This will select the rows with index 1, 3, and 5 and exclude the rest.

In conclusion, selecting rows based on a condition in R allows for the customization and filtering of data frames based on specific criteria, providing more control and accuracy in data analysis.

Select Rows by Condition in R (With Examples)

You can use one of the following methods to select rows by condition in R:

Method 1: Select Rows Based on One Condition

df

Method 2: Select Rows Based on Multiple Conditions

df

Method 3: Select Rows Based on Value in List

df

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

```
#create data frame
```

```
df <- data.frame(points=c(1, 2, 4, 3, 4, 8),  
assists=c(6, 6, 7, 8, 8, 9),  
team=c('A', 'A', 'A', 'B', 'C', 'C'))
```

```
#view data frame
```

df

```
points assists team
```

```
1 1 6 A
```

```
2 2 6 A
```

```
3 4 7 A
```

```
4 3 8 B
```

```
5 4 8 C
```

```
6 8 9 C
```

Method 1: Select Rows Based on One Condition

The following code shows how to select rows based on

one condition in R:

```
#select rows where team is equal to 'A'  
df
```

```
points assists team
```

```
1 1 6 A
```

```
2 2 6 A
```

```
3 4 7 A
```

Notice that only the rows where the team is equal to 'A' are selected.

We can also use `!=` to select rows that are not equal to some value:

```
#select rows where team is not equal to 'A'  
df
```

```
points assists team
```

```
4 3 8 B
```

```
5 4 8 C
```

```
6 8 9 C
```

Method 2: Select Rows Based on Multiple Conditions

The following code shows how to select rows based on multiple conditions in R:

```
#select rows where team is equal to 'A' and points is greater than 1
```

```
df
```

```
points assists team
```

```
2 2 6 A
```

```
3 4 7 A
```

Notice that only the rows where the team is equal to 'A' *and* where points is greater than 1 are selected.

Method 3: Select Rows Based on Value in List

The following code shows how to select rows where the value in a certain column belongs to a list of values:

```
#select rows where team is equal to 'A' or 'C'
```

```
df
```

Notice that only the rows where the team is equal to 'A' or 'C' are selected.