

How can I select rows in a data frame based on specific values in a vector using the “R” programming language?

Authored by
stats writer

June 28, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I select rows in a data frame based on specific values in a vector using the “R” programming language?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=156507>

In the "R" programming language, rows can be selected in a data frame based on specific values in a vector by using the subset function. This function allows for the selection of rows that match the values specified in the vector, providing a convenient way to filter and extract data from a data frame. By using this method, users can easily manipulate and analyze data based on specific criteria, making it a powerful tool for data analysis and manipulation.

R: Select Rows in Data Frame Based on Values in Vector

You can use one of the following methods to select rows from a data frame in R based on values in a vector:

Method 1: Use Base R

```
new_df <- df
```

Method 2: Use dplyr Package

```
library(dplyr)
```

```
new_df <- df %>% filter(column_name %in% values_vector)
```

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

#create data frame

```
df <- data.frame(division=c('West', 'West', 'East', 'East',  
'North'),  
points=c(120, 100, 104, 98, 105),  
assists=c(30, 35, 64, 28, 23))
```

#view data frame

df

division points assists

```
1 West 120 30  
2 West 100 35  
3 East 104 64  
4 East 98 28  
5 North 105 23
```

Example 1: Use Base R to Select Rows Based on Values in Vector

We can use the following code to select only the rows from the original data frame where the value in the division column is equal to 'West' or 'North.'

#define values of interest

```
my_values <- c('West', 'North')
```

#select rows that contain 'West' or 'North' in division

column

```
new_df <- df
```

```
#view results
```

```
new_df
```

```
division points assists
```

```
1 West 120 30
```

```
2 West 100 35
```

```
5 North 105 23
```

The new data frame only contains the rows where the value in the division column is equal to 'West' or 'North.'

Example 2: Use dplyr to Select Rows Based on Values in Vector

We can also use the filter() function from the package in R select only the rows from the original data frame where the value in the division column is equal to 'West' or 'North.'

```
library(dplyr)
```

```
#define values of interest
```

```
my_values <- c('West', 'North')
```

#select rows that contain 'West' or 'North' in division column

```
new_df <- df %>% filter(division %in% my_values)
```

#view results

```
new_df
```

division points assists

1 West 120 30

2 West 100 35

3 North 105 23

The new data frame only contains the rows where the value in the division column is equal to 'West' or 'North.'

Note: The base R and dplyr methods produce the same results. However, the dplyr method will tend to be faster when working with extremely large data frames.

Additional Resources