How can I use the %in% operator in R to filter for rows with a specific value in a list?

Authored by

stats writer

June 24, 2024

The %in% operator in R is a useful tool for filtering rows in a dataset that have a specific value within a list. This operator allows for efficient and precise filtering, eliminating the need for multiple conditional statements. By using the %in% operator, users can specify a list of values and the operator will return the rows that contain those values. It is a simple and effective method for data manipulation and can be easily incorporated into data analysis workflows.

R: Use %in% to Filter for Rows with Value in List

You can use the following basic syntax with the %in% operator in R to filter for rows that contain a value in a list:

```
library(dplyr)

#specify team names to keep
team_names <- c('Mavs', 'Pacers', 'Nets')

#select all rows where team is in list of team names to keep
df_new <- df %>% filter(team %in% team_names)
```

This particular syntax filters a data frame to only keep the rows where the value in the team column is equal to one of the three values in the team_names vector that we specified.

The following example shows how to use this syntax in

practice.

Example: Using %in% to Filter for Rows with Value in List

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

```
#create data frame
df <- data.frame(team=c('Mavs', 'Pacers', 'Mavs', 'Celtics', 'Nets', 'Pacers'),
points=c(104, 110, 134, 125, 114, 124),
assists=c(22, 30, 35, 35, 20, 27))

#view data frame
df

team points assists
1 Mavs 104 22
2 Pacers 110 30
3 Mavs 134 35
4 Celtics 125 35
5 Nets 114 20
6 Pacers 124 27
```

Suppose we would like to filter the data frame to only contain rows where the value in the team column is

equal to one of the following team names:

MavsPacersNets

We can use the following syntax with the %in% operator to do so:

```
library(dplyr)

#specify team names to keep
team_names <- c('Mavs', 'Pacers', 'Nets')

#select all rows where team is in list of team names to keep
df_new <- df %>% filter(team %in% team_names)

#view updated data frame
df_new

team points assists
1 Mavs 104 22
2 Pacers 110 30
3 Mavs 134 35
4 Nets 114 20
5 Pacers 124 27
```

Notice that only the rows with a value of Mavs, Pacers or Nets in the team column are kept.

If you would like to filter for rows where the team name is not in a list of team names, simply add an exclamation point (!) in front of the column name:

```
library(dplyr)

#specify team names to not keep
team_names <- c('Mavs', 'Pacers', 'Nets')

#select all rows where team is not in list of team names to keep
df_new <- df %>% filter(!team %in% team_names)

#view updated data frame
df_new

team points assists
1 Celtics 125 35
```

Notice that only the rows with a value *not equal* to Mavs, Pacers or Nets in the team column are kept.

Note: You can find the complete documentation for the

filter function in dplyr .