

How can I filter rows in a dataset using the “between” operator to select only those where a specific column falls within a certain range of values?

Authored by
stats writer

June 24, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I filter rows in a dataset using the “between” operator to select only those where a specific column falls within a certain range of values?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=151410>

The "between" operator is a useful tool for filtering rows in a dataset. It allows you to select only those rows where a specific column falls within a certain range of values. This can be done by specifying the minimum and maximum values for the column, and the operator will return all rows where the column value is within that range. This method is particularly helpful when working with large datasets, as it allows you to easily narrow down your results to a specific range of values without having to manually search through the entire dataset.

R: Filter Rows where Column is Between Two Values

You can use the following methods to filter a data frame in R where a specific column is between two values:

Method 1: Use Base R

```
df_new <- subset(df, points %in%100:120)
```

Method 2: Use dplyr

```
library(dplyr)
```

```
df_new <- df %>% filter(between(points, 100, 120))
```

Both of these examples filter a data frame to only contain the rows where the value in the points column is between 100 and 120.

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

```
#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
```

Example 1: Filter where Column is Between Two Values Using Base R

We can use the following syntax with the subset() function from base R to filter the data frame to only contain rows where the value in the points column is between 100 and 120:

#filter for rows where value in points column is between 100 and 120

```
df_new <- subset(df, points %in%100:120)
```

#view updated data frame

```
df_new
```

team points assists

1 Mavs 104 22

2 Pacers 110 30

3 Nets 114 20

Notice that only the rows where the value in the points column is between 100 and 120 are kept.

All other rows with a value outside of this range are dropped.

Example 2: Filter where Column is Between Two Values Using dplyr

We can use the following syntax with the filter() and between() functions from the dplyr package in R to filter the data frame to only contain rows where the value in the points column is between 100 and 120:

library(dplyr)

#filter for rows where value in points column is between 100 and 120

```
df_new <- df %>% filter(between(points, 100, 120))
```

#view updated data frame

```
df_new
```

```
team points assists
```

```
1 Mavs 104 22
```

```
2 Pacers 110 30
```

```
3 Nets 114 20
```

Notice that only the rows where the value in the points column is between 100 and 120 are kept.

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