How can I filter rows in R?

Authored by

stats writer

April 18, 2024

Filtering rows in R is a process of selecting and extracting specific rows from a data set based on certain conditions or criteria. This can be done using the `filter()` function in R, which allows users to specify the desired conditions for selecting rows. The `filter()` function uses logical operators such as equal to, not equal to, greater than, and less than, to define the conditions for filtering. Additionally, the `filter()` function can be combined with other functions such as `mutate()` to create more complex filtering conditions. Overall, filtering rows in R is a useful technique for extracting relevant data from a larger data set and is commonly used in data analysis and manipulation tasks.

Filter Rows in R

Often you may be interested in subsetting a data frame based on certain conditions in R. Fortunately this is easy to do using the filter() function from the dplyr package.

library(dplyr)

This tutorial explains several examples of how to use this function in practice using the built-in dplyr dataset called starwars:

#view first six rows of starwars dataset
head(starwars)

 # A tibble: 6 x 13
name height mass hair_color skin_color eye_color birth_year gender homeworld

1 Luke~ 172 77 blond fair blue 19 male Tatooine
2 C-3PO 167 75 <NA> gold yellow 112 <NA> Tatooine
3 R2-D2 96 32 <NA> white, bl~ red 33 <NA> Naboo
4 Dart~ 202 136 none white yellow 41.9 male Tatooine
5 Leia~ 150 49 brown light brown 19 female Alderaan
6 Owen~ 178 120 brown, gr~ light blue 52 male Tatooine
... with 4 more variables: species , films , vehicles ,
starships

Example 1: Filter Rows Equal to Some Value

The following code shows how to filter the dataset for rows where the variable 'species' is equal to Droid.

```
starwars %>% filter(species == 'Droid')
```

```
# A tibble: 5 x 13
name height mass hair_color skin_color eye_color
birth_year gender homeworld
```

1 C-3PO 167 75 gold yellow 112 Tatooine
2 R2-D2 96 32 white, bl~ red 33 Naboo
3 R5-D4 97 32 white, red red NA Tatooine
4 IG-88 200 140 none metal red 15 none
5 BB8 NA NA none none black NA none

 # ... with 4 more variables: species , films , vehicles ,
 # starships

We can see that 5 rows in the dataset met this condition, as indicated by *#A tibble: 5 x 13*.

Example 2: Filter Rows Using 'And'

We can also filter for rows where the species is Droid and the eye color is red:

starwars %>% filter(species == 'Droid' & eye_color == 'red')

 # A tibble: 3 x 13
name height mass hair_color skin_color eye_color birth_year gender homeworld

1 R2-D2 96 32 <NA> white, bl~ red 33 <NA> Naboo
2 R5-D4 97 32 <NA> white, red red NA <NA> Tatooine
3 IG-88 200 140 none metal red 15 none <NA>
 # ... with 4 more variables: species , films , vehicles ,
 # starships

We can see that 3 rows in the dataset met this condition.

Example 3: Filter Rows Using 'Or'

We can also filter for rows where the species is Droid or the eye color is red:

```
starwars %>% filter(species == 'Droid' | eye_color == 'red')

# A tibble: 7 x 13
name height mass hair_color skin_color eye_color birth_year gender homeworld

1 C-3PO 167 75 <NA> gold yellow 112 <NA> Tatooine
2 R2-D2 96 32 <NA> white, bl~ red 33 <NA> Naboo
3 R5-D4 97 32 <NA> white, red red NA <NA> Tatooine
4 IG-88 200 140 none metal red 15 none <NA>
5 Bossk 190 113 none green red 53 male Trandosha
6 Nute~ 191 90 none mottled g~ red NA male Cato Nei~
7 BB8 NA NA none none black NA none <NA>
# ... with 4 more variables: species , films , vehicles ,
# starships
```

We can see that 7 rows in the dataset met this condition.

Example 4: Filter Rows with Values in a List

```
starwars %>% filter(eye_color %in% c('blue', 'yellow', 'red'))

# A tibble: 35 x 13
name height mass hair_color skin_color eye_color birth_year gender

1 Luke~ 172 77 blond fair blue 19 male
2 C-3PO 167 75 <NA> gold yellow 112 <NA>
3 R2-D2 96 32 <NA> white, bl~ red 33 <NA>
4 Dart~ 202 136 none white yellow 41.9 male
5 Owen~ 178 120 brown, gr~ light blue 52 male
6 Beru~ 165 75 brown light blue 47 female
7 R5-D4 97 32 <NA> white, red red NA <NA>
8 Anak~ 188 84 blond fair blue 41.9 male
9 Wilh~ 180 NA auburn, g~ fair blue 64 male
10 Chew~ 228 112 brown unknown blue 200 male
# ... with 25 more rows, and 5 more variables: homeworld , species ,
# films , vehicles , starships
```

We can see that 35 rows in the dataset had an eye color of blue, yellow, or red.

How to Use %in% Operator in R (With Examples)

Example 5: Filter Rows Using Less Than or Greater Than

We can also filter rows using less than or greater than operations on numeric variables:

```
#find rows where height is greater than 250
starwars %>% filter(height > 250)

# A tibble: 1 x 13
name height mass hair_color skin_color eye_color
birth_year gender homeworld

1 Yara~ 264 NA none white yellow NA male Quermia
# ... with 4 more variables: species , films , vehicles ,
# starships

#find rows where height is between 200 and 220
starwars %>% filter(height > 200 & height < 220)

# A tibble: 5 x 13
name height mass hair_color skin_color eye_color
birth_year gender homeworld

1 Dart~ 202 136 none white yellow 41.9 male Tatooine
2 Rugo~ 206 NA none green orange NA male Naboo
```

3 Taun~ 213 NA none grey black NA female Kamino
4 Grie~ 216 159 none brown, wh~ green, y~ NA male Kalee
5 Tion~ 206 80 none grey black NA male Utapau
... with 4 more variables: species , films , vehicles ,
starships

#find rows where height is above the average height
starwars %>% filter(height >mean(height, na.rm = TRUE))

A tibble: 51 x 13
name height mass hair_color skin_color eye_color birth_year gender

1 Dart~ 202 136 none white yellow 41.9 male
2 Owen~ 178 120 brown, gr~ light blue 52 male
3 Bigg~ 183 84 black light brown 24 male
4 Obi-~ 182 77 auburn, w~ fair blue-gray 57 male
5 Anak~ 188 84 blond fair blue 41.9 male
6 Wilh~ 180 NA auburn, g~ fair blue 64 male
7 Chew~ 228 112 brown unknown blue 200 male
8 Han ~ 180 80 brown fair brown 29 male
9 Jabb~ 175 1358 <NA> green-tan~ orange 600 herma~
10 Jek ~ 180 110 brown fair blue NA male

```
# ... with 41 more rows, and 5 more variables:
homeworld , species ,
# films , vehicles , starships
```

You can find the complete documentation for the filter() function here.