

# How can I filter rows that contain a certain string using dplyr?

Authored by  
**stats writer**

April 18, 2024

## RECOMMENDED CITATION

stats writer (2024). *How can I filter rows that contain a certain string using dplyr?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=136690>

Dplyr is a popular R package used for data manipulation. It offers various functions that allow users to filter, select, and arrange data in a data frame. One of its useful features is the ability to filter rows that contain a certain string. This can be achieved by using the "filter" function and specifying the condition using the "contains" argument. This will return a new data frame with only the rows that contain the specified string. This functionality can be helpful in data cleaning and analysis tasks, as it allows for efficient data filtering based on specific criteria.

## Filter Rows that Contain a Certain String Using dplyr

Often you may want to filter rows in a data frame in R that contain a certain string. Fortunately this is easy to do using the `filter()` function from the `dplyr` package and the `grep()` function in Base R.

This tutorial shows several examples of how to use these functions in practice using the following data frame:

```
#create data frame
```

```
df <- data.frame(player = c('P Guard', 'S Guard', 'S  
Forward', 'P Forward', 'Center'),  
points = c(12, 15, 19, 22, 32),  
rebounds = c(5, 7, 7, 12, 11))
```

```
#view data frame
```

```
df
```

```
player points rebounds
```

```
1 P Guard 12 5
2 S Guard 15 7
3 S Forward 19 7
4 P Forward 22 12
5 Center 32 11
```

Example 1: Filter Rows that Contain a Certain String

The following code shows how to filter rows that contain a certain string:

```
#load dplyr package
```

```
library(dplyr)
```

```
#filter rows that contain the string 'Guard' in the player
column
```

```
df %>% filter(grepl('Guard', player))
```

```
player points rebounds
```

```
1 P Guard 12 5
```

```
2 S Guard 15 7
```

Comparing grep() vs. grepl() in R: What's the Difference?

## Example 2: Filter Rows that Contain at Least One String

The following code shows how to filter rows that contain 'Guard' or 'Forward' in the player column:

```
#filter rows that contain 'Guard' or 'Forward' in the  
player column
```

```
df %>% filter(grepl('Guard|Forward', player))
```

```
player points rebounds
```

```
1 P Guard 12 5
```

```
2 S Guard 15 7
```

```
3 S Forward 19 7
```

```
4 P Forward 22 12
```

The following code shows how to filter rows that contain 'P' or 'Center' in the player column:

```
#filter rows that contain 'P' or 'Center' in the player  
column
```

```
df %>% filter(grepl('P|Center', player))
```

```
player points rebounds
```

```
1 P Guard 12 5
```

```
2 P Forward 22 12
```

### 3 Center 32 11

#### Example 3: Filter Out Rows that Contain a Certain String

The following code shows how to filter out (i.e. remove) rows that contain 'Guard' in the player column:

```
#filter out rows that contain 'Guard' in the player column
```

```
df %>% filter(!grepl('Guard', player))
```

```
player points rebounds
```

```
1 S Forward 19 7
```

```
2 P Forward 22 12
```

```
3 Center 32 11
```

The following code shows how to filter out (i.e. remove) rows that contain 'Guard' or 'Center' in the player column:

```
#filter out rows that contain 'Guard' or 'Center' in the player column
```

```
df %>% filter(!grepl('Guard|Center', player))
```

```
player points rebounds
```

```
1 S Forward 19 7
```

## 2 P Forward 22 12

ARABPSYCHOLOGY.COM