

How can I use the n() function in R? Can you provide examples of how it can be used?

Authored by
stats writer

June 25, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I use the n() function in R? Can you provide examples of how it can be used?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=152717>

The n() function in R is a useful tool for obtaining the number of observations in a data set. It can be used to quickly and accurately count the number of rows or elements in a given data frame, vector, or list. This function is particularly helpful for data analysis and statistical calculations where the size of the data set is an important factor. To use the n() function, simply input the desired data set as the argument and the function will return the corresponding number of observations. For example, n(my_data_frame) will return the number of rows in the data frame called "my_data_frame". In addition, the n() function can also be used in combination with other functions to perform calculations and filtering operations on data sets. Overall, the n() function is a valuable tool for managing and analyzing data in R.

Use n() Function in R (With Examples)

You can use the n() function from the package in R to count the number of observations in a group.

Here are three common ways to use this function in practice:

Method 1: Use n() to Count Observations by Group

```
df %>%  
  group_by(group_variable) %>%  
  summarise(count = n())
```

Method 2: Use n() to Add Column that Shows Observations by Group

```
df %>%  
  group_by(group_variable) %>%
```

```
mutate(count = n())
```

Method 3: Use n() to Filter Based on Observations by Group

```
df %>%  
group_by(group_variable) %>%  
filter(n() > 15)
```

The following examples show how to use each method in practice with the following data frame in R that contains information about various basketball players:

```
#create data frame  
df <- data.frame(team=c('A', 'A', 'A', 'B', 'B', 'C'),  
points=c(22, 25, 25, 20, 29, 13),  
assists=c(10, 12, 9, 4, 11, 10),  
rebounds=c(9, 8, 5, 10, 14, 12))
```

```
#view data frame
```

```
df
```

```
team points assists rebounds  
1 A 22 10 9  
2 A 25 12 8
```

3 A 25 9 5

4 B 20 4 10

5 B 29 11 14

6 C 13 10 12

Example 1: Use n() to Count Observations by Group

The following code shows how to use the n() function along with the summarise() function to count the number of observations by team:

```
library(dplyr)
```

```
#count number of observations by team
```

```
df %>%
```

```
group_by(team) %>%
```

```
summarise(count = n())
```

```
# A tibble: 3 x 2
```

```
team count
```

```
1 A 3
```

```
2 B 2
```

```
3 C 1
```

From the output we can see:

Team A occurs 3 times Team B occurs 2 times Team C occurs 1 time

Example 2: Use n() to Add Column that Shows Observations by Group

The following code shows how to use the n() function along with the mutate() function to add a column to the date frame that contains the number of observations by team:

```
library(dplyr)
```

```
#add new column that shows number of observations  
by team
```

```
df %>%
```

```
group_by(team) %>%
```

```
mutate(count = n())
```

```
# A tibble: 6 x 5
```

```
# Groups: team
```

```
team points assists rebounds count
```

```
1 A 22 10 9 3
```

2 A 25 12 8 3

3 A 25 9 5 3

4 B 20 4 10 2

5 B 29 11 14 2

6 C 13 10 12 1

The new column called count contains the team count for each row in the data frame.

Example 3: Use n() to Filter Based on Observations by Group

The following code shows how to use the n() function along with the filter() function to filter the data frame to only show rows where the team occurs greater than one time:

```
library(dplyr)
```

```
#filter rows where team count is greater than 1
```

```
df %>%
```

```
group_by(team) %>%
```

```
filter(n() > 1)
```

```
# A tibble: 5 x 4
```

```
# Groups: team
```

team points assists rebounds

1 A 22 10 9

2 A 25 12 8

3 A 25 9 5

4 B 20 4 10

5 B 29 11 14

Notice that the resulting data frame only contains rows where the team is "A" or "B" because these are the only teams that have a count greater than one.