

How can I count the number of observations within each group in R?

Authored by
stats writer

April 18, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I count the number of observations within each group in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=136644>

The process of counting the number of observations within each group in R involves using the "table" function to create a frequency table, which displays the number of occurrences of each unique value within a group. This function can be applied to a vector, factor, or data frame in R, allowing for easy calculation of the number of observations within each group. Additionally, the "aggregate" function can be used to group data and calculate the number of observations within each group simultaneously. This method is particularly useful for data analysis and organization in R.

Count Observations by Group in R

Often you may be interested in counting the number of observations by group in R.

Fortunately this is easy to do using the count() function from the dplyr library.

This tutorial explains several examples of how to use this function in practice using the following data frame:

```
#create data frame
```

```
df <- data.frame(team = c('A', 'A', 'A', 'B', 'B', 'B', 'B', 'B',  
'C', 'C', 'C', 'C'),  
position = c('G', 'G', 'F', 'G', 'F', 'F', 'F', 'G', 'G', 'F', 'F', 'F'),  
points = c(4, 13, 7, 8, 15, 15, 17, 9, 21, 22, 25, 31))
```

```
#view data frame
```

```
df
```

team position points

1 A G 4

2 A G 13

3 A F 7

4 B G 8

5 B F 15

6 B F 15

7 B F 17

8 B G 9

9 C G 21

10 C F 22

11 C F 25

12 C F 31

Example 1: Count by One Variable

The following code shows how to count the total number of players by team:

```
library(dplyr)
```

```
#count total observations by variable 'team'
```

```
df %>% count(team)
```

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

team n

1 A 3

2 B 5

3 C 4

From the output we can see that:

Team A has 3 players Team B has 5 players Team C has 4 players

This single count() function gives us a nice idea of the distribution of players by team.

Note that we can also sort the counts if we'd like:

```
#count total observations by variable 'team'  
df %>% count(team, sort=TRUE)
```

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

team n

1 B 5

2 C 4

3 A 3

Example 2: Count by Multiple Variables

We can also sort by more than one variable:

```
#count total observations by 'team' and 'position'  
df %>% count(team, position)
```

```
# A tibble: 6 x 3  
team position n
```

```
1 A F 1  
2 A G 2  
3 B F 3  
4 B G 2  
5 C F 3  
6 C G 1
```

From the output we can see that:

Team A has 1 player at the 'F' (forward) position and 2 players at the 'G' (guard) position. Team B has 3 players at the 'F' (forward) position and 2 players at the 'G' (guard) position. Team C has 3 players at the 'F' (forward) position and 1 player at the 'G' (guard) position.

Example 3: Weighted Count

We can also "weight" the counts of one variable by another variable. For example, the following code shows how to count the total observations per team, using the variable 'points' as the weight:

```
df %>% count(team, wt=points)
```

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

```
team n
```

```
1 A 24
```

```
2 B 64
```

```
3 C 99
```

You can find the complete documentation for the `count()` function [here](#).