

How can I calculate the mode by group in R, and can you provide some examples?”

Authored by
stats writer

June 27, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I calculate the mode by group in R, and can you provide some examples?”*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=154856>

Calculating the mode by group in R refers to finding the most frequently occurring value within a specific group or category in a dataset. This can be achieved by using the "by" function in R, which allows for grouping the data based on a particular variable and then applying the "Mode" function to that group. The result will be the mode for each group within the dataset. For example, if we have a dataset of students' grades in different subjects, we can use the "by" function to group the grades by subject and then calculate the mode for each subject. This will provide us with the most common grade for each subject.

Calculate Mode by Group in R (With Examples)

The **mode** of a dataset represents the most frequently occurring value.

The statistical software R does not have a built-in function to calculate the mode of a dataset, but you can use the following function to calculate the mode:

```
find_mode <- function(x) {  
  u <- unique(x)  
  tab <- tabulate(match(x, u))  
  u[tab == max(tab)]  
}
```

The following examples show how to use this function to calculate the mode by group in R.

Example 1: Calculate Mode by Group in R (One Mode)

Suppose we have the following data frame in R that shows the points scored by basketball players on various teams:

```
#define data frame
```

```
df <- data.frame(team=c('A', 'A', 'A', 'A', 'B', 'B', 'B', 'B'),  
points=c(5, 7, 7, 9, 12, 12, 10, 14))
```

```
#view data frame
```

```
df
```

```
team points
```

```
1 A 5
```

```
2 A 7
```

```
3 A 7
```

```
4 A 9
```

```
5 B 12
```

```
6 B 12
```

```
7 B 10
```

```
8 B 14
```

We can use the following code to calculate the mode of points, grouped by team:

```
library(dplyr)
```

#define function to calculate mode

```
find_mode <- function(x) {  
  u <- unique(x)  
  tab <- tabulate(match(x, u))  
  u  
}
```

#calculate mode of 'points' by 'team'

```
df %>%  
  group_by(team) %>%  
  summarize(mode_points = find_mode(points))
```

A tibble: 2 x 2

team mode_points

1 A 7

2 B 12

From the results we can see:

The mode of points for team A is 7. The mode of points for team B is 12.

Example 2: Calculate Mode by Group in R (Multiple Modes)

Suppose we have the following data frame in R:

```
#define data frame
```

```
df <- data.frame(team=c('A', 'A', 'A', 'A', 'B', 'B', 'B', 'B'),  
points=c(5, 7, 7, 9, 12, 12, 10, 10))
```

```
#view data frame
```

```
df
```

```
team points
```

```
1 A 5
```

```
2 A 7
```

```
3 A 7
```

```
4 A 9
```

```
5 B 12
```

```
6 B 12
```

```
7 B 10
```

```
8 B 10
```

We can use the following code to calculate the mode of points, grouped by team:

```
library(dplyr)
```

```
#define function to calculate mode
```

```
find_mode <- function(x) {
```

```
u <- unique(x)
```

```
tab <- tabulate(match(x, u))
u
}

#calculate mode of 'points' by 'team'
df %>%
group_by(team) %>%
summarize(mode_points = find_mode(points))

# A tibble: 3 x 2
# Groups: team
team mode_points

1 A 7
2 B 12
3 B 10
```

From the results we can see:

The mode of points for team A is 7. The mode of points for team B is 12 and 10.

In this example, there were two points values that occurred most frequently for team B, so each of these mode values is returned on a separate line for team B in

the output.

The following tutorials explain how to calculate other descriptive statistics in R:

ARABPSYCHOLOGY.COM