

# How can I calculate the cumulative sum for each group in R?

Authored by  
**stats writer**

June 25, 2024

## RECOMMENDED CITATION

stats writer (2024). *How can I calculate the cumulative sum for each group in R?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=153055>

To calculate the cumulative sum for each group in R, one can use the "cumsum" function. This function takes a vector as an input and returns a vector of the same length, where each element is the cumulative sum of the elements in the input vector up to that point. Additionally, one can use the "group\_by" function from the dplyr package to group the data by a specific variable and then apply the "cumsum" function to calculate the cumulative sum for each group. This method is useful for analyzing trends and patterns within a dataset, and can be easily implemented in R programming.

## Calculate Cumulative Sum by Group in R

You can use the following methods to calculate a cumulative sum by group in R:

### Method 1: Use Base R

```
df$cum_sum <- ave(df$values_var, df$group_var,  
FUN=cumsum)
```

### Method 2: Use dplyr

```
library(dplyr)
```

```
df %>% group_by(group_var) %>% mutate(cum_sum =  
cumsum(values_var))
```

### Method 3: Use data.table

```
library(data.table)
```

## setDT(df)

The following examples show how to use each method in practice with the following data frame in R:

```
#create data frame
```

```
df <- data.frame(store=rep(c('A', 'B', 'C'), each=4),  
sales=c(3, 4, 4, 2, 5, 8, 9, 7, 6, 8, 3, 2))
```

```
#view data frame
```

```
df
```

```
store sales
```

```
1 A 3
```

```
2 A 4
```

```
3 A 4
```

```
4 A 2
```

```
5 B 5
```

```
6 B 8
```

```
7 B 9
```

```
8 B 7
```

```
9 C 6
```

```
10 C 8
```

```
11 C 3
```

```
12 C 2
```

## Example 1: Calculate Cumulative Sum by Group Using Base R

The following code shows how to use the `ave()` function from base R to calculate the cumulative sum of sales, grouped by store:

```
#add column to show cumulative sales by store  
df$cum_sales <- ave(df$sales, df$store, FUN=cumsum)
```

```
#view updated data frame
```

```
df
```

```
store sales cum_sales
```

```
1 A 3 3
```

```
2 A 4 7
```

```
3 A 4 11
```

```
4 A 2 13
```

```
5 B 5 5
```

```
6 B 8 13
```

```
7 B 9 22
```

```
8 B 7 29
```

```
9 C 6 6
```

```
10 C 8 14
```

```
11 C 3 17
```

```
12 C 2 19
```

The new column called `cum_sales` displays the cumulative sum of sales, grouped by store.

Example 2: Calculate Cumulative Sum by Group Using `dplyr`

The following code shows how to use various functions from the `dplyr` package in R to calculate the cumulative sum of sales, grouped by store:

```
library(dplyr)

#add column to show cumulative sales by store
df %>% group_by(store) %>% mutate(cum_sales =
cumsum(sales))

#view updated data frame
df

# A tibble: 12 x 3
# Groups:   store
store sales cum_sales

1 A 3 3
2 A 4 7
3 A 4 11
4 A 2 13
```

5 B 5 5

6 B 8 13

7 B 9 22

8 B 7 29

9 C 6 6

10 C 8 14

11 C 3 17

12 C 2 19

The new column called `cum_sales` displays the cumulative sum of sales, grouped by store.

Example 3: Calculate Cumulative Sum by Group Using `data.table`

The following code shows how to use various functions from the `data.table` package in R to calculate the cumulative sum of sales, grouped by store:

```
library(data.table)
```

```
#add column to show cumulative sales by store
```

```
setDT(df)
```

```
#view updated data frame
```

```
df
```

## store sales cum\_sales

1: A 3 3

2: A 4 7

3: A 4 11

4: A 2 13

5: B 5 5

6: B 8 13

7: B 9 22

8: B 7 29

9: C 6 6

10: C 8 14

11: C 3 17

12: C 2 19

The new column called `cum_sales` displays the cumulative sum of sales, grouped by store.

**Note:** All three methods produce the same result. However, the `dplyr` and `data.table` methods will tend to be quicker when working with extremely large data frames.

The following tutorials explain how to perform other common calculations in R: