How can I calculate cumulative sums in R, and what are some examples of how to do so?

Authored by

stats writer

April 20, 2024

Cumulative sums, also known as running totals, are a useful tool in data analysis to measure the accumulation of a particular variable over time or within a certain group. In R, calculating cumulative sums can be done using the `cumsum()` function, which returns a vector containing the sum of all values up to a given point in a data set. This function can be applied to any numerical vector, matrix, or data frame.

To calculate the cumulative sums in R, simply use the `cumsum()` function followed by the vector or data frame you want to operate on. For example, if you have a data frame with columns representing monthly sales data, you can use `cumsum()` to calculate the cumulative sales for each month. This will give you a new column with the cumulative sum of sales up to that month.

Another example of calculating cumulative sums in R is when analyzing stock market data. By using `cumsum()` on a vector of daily stock returns, you can obtain the cumulative return over a certain time period, providing insights into the overall performance of the stock.

Overall, the `cumsum()` function in R is a useful tool for tracking the progression of data over time or within groups, and can be applied to various types of data sets for different analytical purposes.

Calculate Cumulative Sums in R (With Examples)

You can use the cumsum() function from base R to easily calculate the cumulative sum of a vector of numeric values.

This tutorial explains how to use this function to calculate the cumulative sum of a vector along with how to visualize a cumulative sum.

How to Calculate a Cumulative Sum in R

The following code shows how to calculate the cumulative sum of sales for a given company over the course of 15 sales quarters:

```
#create dataset
data <- data.frame(quarter=1:15,
sales=c(1, 2, 2, 5, 4, 7, 5, 7, 6, 8, 5, 9, 11, 12, 4))

#create new column in dataset that contains cumulative sales
data$cum_sales <- cumsum(data$sales)

#view dataset
data
```

quarter sales cum_sales
1 1 1 1
2 2 2 3
3 3 2 5
4 4 5 10
5 5 4 14
6 6 7 21
7 7 5 26
8 8 7 33
9 9 6 39
10 10 8 47
11 11 5 52
12 12 9 61
13 13 11 72

14 14 12 84

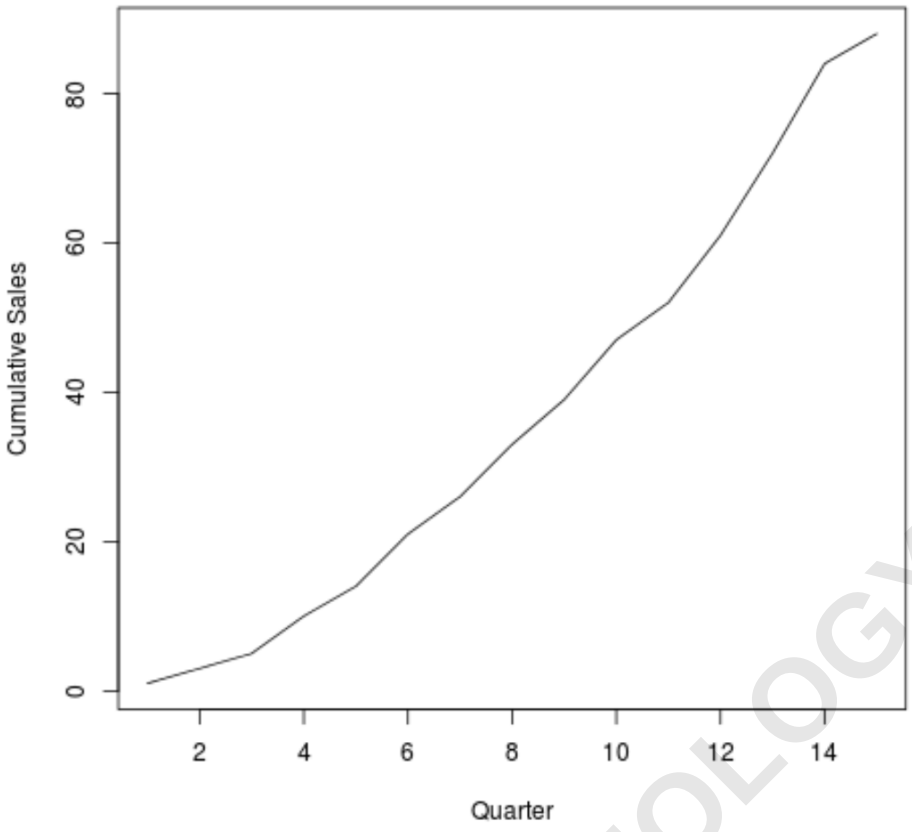
15 15 4 88

The values shown in the cum_sales column represent the total sales up to and including that quarter. For example, the cumulative sales in quarter 5 are calculated as: 1+2+2+5+4 = 14.

How to Visualize a Cumulative Sum in R

Once we've calculated the cumulative sales, we can create a simple line chart in base R to visualize the cumulative sales by quarter:

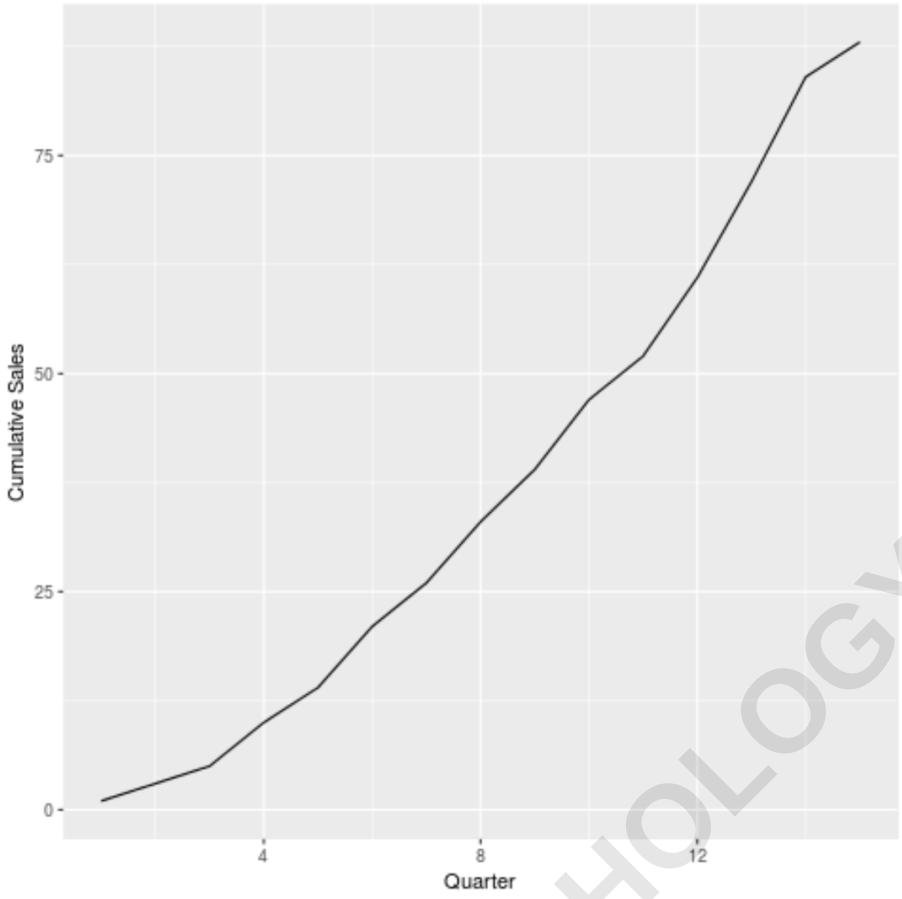
```
plot(data$cum_sales, type='l', xlab='Quarter', ylab='Cumulative Sales')
```



Alternatively, we can use the R visualization library ggplot2 to create the same line chart:

```
library(ggplot2)

ggplot(data, aes(x=quarter, y=cum_sales)) +
geom_line() +
labs(x='Quarter', y='Cumulative Sales')
```



How to Average Across Columns in R

How to Sum Specific Columns in R

How to Perform a COUNTIF Function in R