How can I calculate the difference between rows in R?

Authored by

stats writer

May 5, 2024

Calculating the difference between rows in R involves using the diff() function, which calculates the difference between consecutive values in a vector or data frame. This function can be applied to a single column or multiple columns, allowing for the comparison of values within a row or between rows. Additionally, the diff() function can be used with different parameters to customize the type of difference calculation, such as absolute differences or percentage changes. This method is useful for analyzing trends and changes over time in data sets, as well as identifying any outliers or discrepancies between rows.

Calculate Difference Between Rows in R

You can use the diff() function to calculate the difference between rows of a data frame in R:

#find difference between rows in every column of data frame
diff(as.matrix(df))

#find difference between rows of specific column
diff(df$column_name)

The following examples show how to use this syntax in practice.

Example 1: Find Difference Between Rows of Every Column

The following code shows how to calculate the difference between rows for every column in a data frame:

```
#create data frame
df <- data.frame(day=c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10),
sales=c(7, 8, 8, 12, 10, 9, 13, 16, 11, 7))

#view data frame
df

day sales
1 1 7
2 2 8
3 3 8
4 4 12
5 5 10
6 6 9
7 7 13
8 8 16
9 9 11
10 10 7

#calculate difference between rows for each column
diff(as.matrix(df))

day sales
1 1
1 0
```

1 4

1 -2

1 -1

1 4

1 3

1 -5

1 -4

Example 2: Find Difference Between Rows of Specific Column

The following code shows how to calculate the difference between rows for a specific column in a data frame:

```
#create data frame
df <- data.frame(day=c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10),
sales=c(7, 8, 8, 12, 10, 9, 13, 16, 11, 7))

#calculate difference between rows in 'sales' column
diff(df$sales)
```

1 0 4 -2 -1 4 3 -5 -4

Example 3: Find Difference Between Rows & Append New Column

The following code shows how to calculate the

difference between rows for a specific column in a data frame and then append those differences as a new column at the end of the data frame:

```
#create data frame
df <- data.frame(day=c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10),
sales=c(7, 8, 8, 12, 10, 9, 13, 16, 11, 7))

#calculate difference between rows in 'sales' column
sales_diff <- diff(df$sales)

#append NA to beginning of differences vector
sales_diff <- c(NA, sales_diff)

#append differences vector as new column
df$sales_diff <- sales_diff

#view updated data frame
df

day sales sales_diff
1 1 7 NA
2 2 8 1
3 3 8 0
4 4 12 4
5 5 10 -2
```

6 6 9 -1

7 7 13 4

8 8 16 3

9 9 11 -5

10 10 7 -4

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