How do you write a case statement in R and can you provide an example?

Authored by

stats writer

June 27, 2024

A case statement in R is a conditional statement that allows for the execution of different code based on the value of a specific variable. It follows the structure of "if-else" statements, where the code within the statement will only be executed if the specified condition is met. An example of a case statement in R is as follows:

#Defining a variable
age

Write a Case Statement in R (With Example)

A case statement is a type of statement that goes through conditions and returns a value when the first condition is met.

The easiest way to implement a case statement in R is by using the case_when() function from the dplyr package:

```
library(dplyr)

df %>%
mutate(new_column = case_when(
col1 < 9 ~ 'value1',
col1 < 12 ~ 'value2',
col1 < 15 ~ 'value3',
TRUE ~ 'Great'))
```

This particular function looks at the value in the column

called col1 and returns:

"value1" if the value in col1 is less than 9"value2" if the value in col1 is less than 12"value3" if the value in col2 is less than 15"value4" if none of the previous conditions are true

Note that TRUE is equivalent to an "else" statement.

The following example shows how to use this function in practice.

Example: Case Statement in R

Suppose we have the following data frame in R:

```
#create data frame
df <- data.frame(player=c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10),
points=c(6, 8, 9, 9, 12, 14, 15, 17, 19, 22))

#view data frame
df

player points
1 1 6
2 2 8
3 3 9
```

4 4 9

5 5 12

6 6 14

7 7 15

8 8 17

9 9 19

10 10 22

We can use the following syntax to write a case statement that creates a new column called class whose values are determined by the values in the points column:

```
library(dplyr)

#create new column using case statement
df %>%
mutate(class = case_when(
points < 9 ~ 'Bad',
points < 12 ~ 'OK',
points < 15 ~ 'Good',
TRUE ~ 'Great'))
```

player points class

1 1 6 Bad

2 2 8 Bad

3 3 9 OK

4 4 9 OK

5 5 12 Good

6 6 14 Good

7 7 15 Great

8 8 17 Great

9 9 19 Great

10 10 22 Great

The case statement looked at the value in the points column and returned:

"Bad" if the value in the points column was less than 9"OK" if the value in the points column was less than 12"Good" if the value in the points column was less than 15"Great" if none of the previous conditions are true

The new column is called class, since this is the name we specified in the mutate() function.