

How do you compare two columns in R, including examples?

Authored by
stats writer

April 20, 2024

RECOMMENDED CITATION

stats writer (2024). *How do you compare two columns in R, including examples?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=137421>

Comparing two columns in R involves examining the values within each column and determining if there are any similarities or differences between them. This can be done using various functions and operators in R, such as the "==" (equality) operator or the "identical()" function. For example, if we have two columns "A" and "B" with values 1, 2, and 3 in the first and 1, 4, and 3 in the second, using the "==" operator would return a logical vector of TRUE, FALSE, and TRUE, indicating that the values in column A and B are equal in the first and third rows, but not in the second row. The "identical()" function, on the other hand, would return a single value of FALSE, as it compares the entire columns rather than individual values. By comparing two columns, we can gain insights into the relationships between different variables in a dataset.

Compare Two Columns in R (With Examples)

Often you may want to compare two columns in R and write the results of the comparison to a third column.

You can easily do this by using the following syntax:

```
df$new_col <- ifelse(df$col1 > df$col2, 'A',  
ifelse(df$col1 < df$col2, 'B', 'C'))
```

This single line of code does the following:

If column 1 is greater than column 2 then write 'A' as the output to the third column. Otherwise, if column 1 is less than column 2 then write 'B' as the output. Otherwise, write 'C' as the output.

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

Example: Compare Two Columns in R

Suppose we have the following data frame that shows the number of goals scored by two soccer teams in five different matches:

```
#create data frame
```

```
df <- data.frame(A_points=c(1, 3, 3, 3, 5),  
B_points=c(4, 5, 2, 3, 2))
```

```
#view data frame
```

```
df
```

```
A_points B_points
```

```
1 1 4
```

```
2 3 5
```

```
3 3 2
```

```
4 3 3
```

```
5 5 2
```

We can use the following code to compare the number of goals by row and output the winner of the match in a third column:

```
#compare A_points and B_points and output results to
```

new column titled *winner*

```
df$winner <- ifelse(df$A_points > df$B_points, 'A',  
ifelse(df$A_points < df$B_points, 'B', 'Tie'))
```

#view data frame

df

```
A_points B_points winner
```

```
1 1 4 B
```

```
2 3 5 B
```

```
3 3 2 A
```

```
4 3 3 Tie
```

```
5 5 2 A
```

The results of the comparison are shown in the new column called *winner*.

How to Stack Data Frame Columns in R

How to Combine Two Columns into One in R

How to Loop Through Column Names in R