

How can I add a column to a data frame in R based on other columns?

Authored by
stats writer

May 4, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I add a column to a data frame in R based on other columns?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=142623>

To add a new column to a data frame in R based on existing columns, you can use the "mutate" function from the "dplyr" package. This allows you to create a new column by performing operations on the existing columns. Additionally, you can use the "cbind" function to combine two data frames, adding a new column to the existing data. These methods provide efficient ways to add new information to a data frame and update it according to the desired criteria.

R: Add Column to Data Frame Based on Other Columns

You can use the following basic syntax to add a column to a data frame in R based on the values in other columns:

```
#add new column 'col3' with values based on columns 1  
and 2  
df$col3 <- with(df, ifelse(col1 > col2, value_if_true,  
value_if_false))
```

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

Example 1: Add Character Column Based on Other Columns

The following code shows how to add a new character column based on the values in other columns of the data frame:

```
#create data frame  
df <- data.frame(team=c('Mavs', 'Cavs', 'Spurs', 'Nets'),
```

```
scored=c(99, 90, 84, 96),  
allowed=c(95, 80, 87, 95))
```

```
#view data frame
```

```
df
```

```
team scored allowed
```

```
1 Mavs 99 95
```

```
2 Cavs 90 80
```

```
3 Spurs 84 87
```

```
4 Nets 96 95
```

```
#add 'result' column based on values in 'scored' and  
'allowed' columns
```

```
df$result <- with(df, ifelse(scored > allowed, 'Win',  
'Loss'))
```

```
#view updated data frame
```

```
df
```

```
team scored allowed result
```

```
1 Mavs 99 95 Win
```

```
2 Cavs 90 80 Win
```

```
3 Spurs 84 87 Loss
```

```
4 Nets 96 95 Win
```

And the following code shows how to add a new character column that combines two `ifelse()` functions to produce three potential values in a new column:

```
#create data frame
```

```
df <- data.frame(team=c('Mavs', 'Cavs', 'Spurs', 'Nets'),  
scored=c(99, 90, 84, 96),  
allowed=c(95, 80, 87, 95))
```

```
#view data frame
```

```
df
```

```
team scored allowed
```

```
1 Mavs 99 95
```

```
2 Cavs 90 80
```

```
3 Spurs 84 87
```

```
4 Nets 96 95
```

```
#add 'quality' column based on values in 'scored' and  
'allowed' columns
```

```
df$quality <- with(df, ifelse(scored > 95, 'great',  
ifelse(scored > 85, 'good', 'bad')))
```

```
#view updated data frame
```

```
df
```

team scored allowed quality

1 Mavs 99 95 great

2 Cavs 90 80 good

3 Spurs 84 87 bad

4 Nets 96 95 great

Example 2: Add Numeric Column Based on Other Columns

The following code shows how to add a new numeric column to a data frame based on the values in other columns:

```
#create data frame
```

```
df <- data.frame(team=c('Mavs', 'Cavs', 'Spurs', 'Nets'),  
scored=c(99, 90, 84, 96),  
allowed=c(95, 80, 87, 95))
```

```
#view data frame
```

```
df
```

team scored allowed

1 Mavs 99 95

2 Cavs 90 80

3 Spurs 84 87

4 Nets 96 95

```
#add 'lower_score' column based on values in 'scored'  
and 'allowed' columns
```

```
df$lower_score <- with(df, ifelse(scored > allowed,  
allowed, scored))
```

```
#view updated data frame
```

```
df
```

```
team scored allowed lower_score
```

```
1 Mavs 99 95 95
```

```
2 Cavs 90 80 80
```

```
3 Spurs 84 87 84
```

```
4 Nets 96 95 95
```

ARABPSYCHOLOGY.COM