Can different columns be combined from two data frames in R?

Authored by

stats writer

May 2, 2024

ARABPSYCHOLOGY.COM

RECOMMENDED CITATION

stats writer (2024). *Can different columns be combined from two data frames in R?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=141940

In R, it is possible to combine different columns from two separate data frames. This can be done by using functions such as cbind() or merge(), which allow for the merging of columns based on a shared identifier or index. This process is commonly used in data analysis and manipulation to merge data from multiple sources into a single data frame. The resulting combined data frame can then be used for further analysis and visualization.

Combine Two Data Frames in R with Different Columns

You can use the bind_rows() function from the package in R to quickly combine two data frames that have different columns:

```
library(dplyr)
bind_rows(df1, df2)
```

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

Example: Combine Two Data Frames with Different Columns

Suppose we have the following two data frames in R:

```
#define first data frame
df1 <- data.frame(A=c(1, 6, 3, 7, 5),
B=c(7, 9, 8, 3, 2),
C=c(3, 5, 2, 9, 9))

df1
```

ARABPSYCHOLOGY.COM

A B C

1 1 7 3

2 6 9 5

3 3 8 2

4 7 3 9

5 5 2 9

#define second data frame

df2 <- data.frame(B=c(1, 3, 3, 4, 5),

C=c(7, 7, 8, 3, 2),

D=c(3, 3, 6, 6, 8))

df2

B C D

1 1 7 3

2 3 7 3

3 3 8 6

4 4 3 6

5 5 2 8

Note that df1 has the following column names:

ABC

ARABPSYCHOLOGY.COM

And note that df2 has the following column names:

BCD

The column names don't match, so the function in R will throw an error if we attempt to use it.

```
#attempt to use rbind to row bind data frames
rbind(df1, df2)
```

```
Error in match.names(clabs, names(xi)) :
names do not match previous names
```

Instead, we can use the bind_rows() function from the dplyr package to combine these two data frames and simply fill in missing values in the resulting data frame with NA values:

```
library(dplyr)

#combine df1 and df2
bind_rows(df1, df2)
```

```
A B C D
1 1 7 3 NA
2 6 9 5 NA
```

ARABPSYCHOLOGY.COM

3 3 8 2 NA

4 7 3 9 NA

5 5 2 9 NA

6 NA 1 7 3

7 NA 3 7 3

8 NA 3 8 6

9 NA 4 3 6

10 NA 5 2 8

ARABPSYCHOLOGY.COM