

How can data frames be merged by row names in R?

Authored by
stats writer

June 28, 2024

RECOMMENDED CITATION

stats writer (2024). *How can data frames be merged by row names in R?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=156686>

Data frames in R can be merged by their row names using the "merge" function. This function combines two data frames based on their shared row names, creating a new data frame with the rows from both original data frames. The resulting data frame will have all the columns from both data frames, with rows merged based on their matching row names. This allows for easy combination and comparison of data from multiple sources using a common identifier. The "merge" function also allows for customization of the merging process, such as handling of non-matching row names and duplicate row names. This method of merging data frames is useful in various data analysis and manipulation tasks in R.

Merge Data Frames by Row Names in R

You can use the following basic syntax to merge two data frames in R based on their rownames:

#inner join

```
merge(df1, df2, by=0)
```

#left join

```
merge(df1, df2, by=0, all.x=TRUE)
```

#outer join

```
merge(df1, df2, by=0, all=TRUE)
```

By using the argument `by=0`, we're able to tell R that we want to merge using the rownames of the data frames.

The following examples show how to use each method with the following two data frames:

#create first data frame

```
df1 <- data.frame(points=c(99, 90, 86, 88, 95),  
assists=c(33, 28, 31, 39, 34))
```

```
rownames(df1) <- c(1, 2, 3, 4, 5)
```

df1

points assists

1 99 33

2 90 28

3 86 31

4 88 39

5 95 34

#create second data frame

```
df2 <- data.frame(rebounds=c(17, 15, 22, 26, 25),  
blocks=c(7, 7, 15, 12, 14))
```

```
rownames(df2) <- c(3, 4, 5, 6, 7)
```

df2

rebounds blocks

3 17 7

4 15 7

5 22 15

6 26 12

7 25 14

Example 1: Perform Inner Join Using Row Names

The following code shows how to perform an inner join on two data frames using the row names:

```
#perform inner join using row names  
merge(df1, df2, by=0)
```

Row.names points assists rebounds blocks

1 3 86 31 17 7

2 4 88 39 15 7

3 5 95 34 22 15

Notice that only the rows whose row names belong in both data frames are kept in the final merged data frame.

Example 2: Perform Left Join Using Row Names

The following code shows how to perform a left join on two data frames using the row names:

```
#perform left join using row names  
merge(df1, df2, by=0, all.x=TRUE)
```

Row.names points assists rebounds blocks

```
1 1 99 33 NA NA  
2 2 90 28 NA NA  
3 3 86 31 17 7  
4 4 88 39 15 7  
5 5 95 34 22 15
```

Notice that all of the rows from the first data frame are kept in the final merged data frame.

Example 3: Perform Outer Join Using Row Names

The following code shows how to perform an outer join on two data frames using the row names:

```
#perform outer join using row names  
merge(df1, df2, by=0, all=TRUE)
```

Row.names points assists rebounds blocks

```
1 1 99 33 NA NA  
2 2 90 28 NA NA  
3 3 86 31 17 7  
4 4 88 39 15 7
```

5 5 95 34 22 15

6 6 NA NA 26 12

7 7 NA NA 25 14

Notice that all of the rows from both data frames are kept in the final merged data frame.

Additional Resources

The following tutorials explain how to perform other common functions related to data frames in R:

ARABPSYCHOLOGY.COM