

# How can I remove all duplicate rows in R so that none are left?

Authored by  
**stats writer**

June 26, 2024

## RECOMMENDED CITATION

stats writer (2024). *How can I remove all duplicate rows in R so that none are left?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=153132>

Removing duplicate rows in R is a process of eliminating all repeated rows from a data set, resulting in a dataset with no duplicate rows. This can be achieved using the `unique()` function, which identifies and removes all duplicate rows, leaving only distinct rows in the dataset. By removing duplicate rows, the data becomes more organized and easier to analyze, as there is no repetition of information. This process is useful in data cleaning and preprocessing, ensuring the accuracy and reliability of the data for further analysis or modeling.

## Remove Duplicate Rows in R so None are Left

You can use the following methods in R to remove duplicate rows from a data frame so that none are left in the resulting data frame:

### Method 1: Use Base R

```
new_df <- df
```

### Method 2: Use dplyr

```
library(dplyr)
```

```
new_df <- df %>%
```

```
group_by(across(everything())) %>%
```

```
filter(n() == 1)
```

The following examples show how to use each method in practice with the following data frame:

**#create data frame**

```
df <- data.frame(team=c('A', 'A', 'A', 'A', 'B', 'B', 'B', 'B'),  
points=c(20, 20, 28, 14, 13, 18, 27, 13))
```

**#view data frame**

**df**

**team points**

**1 A 20**

**2 A 20**

**3 A 28**

**4 A 14**

**5 B 13**

**6 B 18**

**7 B 27**

**8 B 13**

**Example 1: Use Base R**

The following code shows how to use functions from base R to remove duplicate rows from the data frame so that none are left:

**#create new data frame that removes duplicates so none are left**

```
new_df <- df
```

```
#view new data frame
```

```
new_df
```

```
team points
```

```
3 A 28
```

```
4 A 14
```

```
6 B 18
```

```
7 B 27
```

Notice that each of the duplicate rows have been removed from the data frame and none of the duplicates remain.

Example 2: Use dplyr

The following code shows how to use functions from the dplyr package in R to remove duplicate rows from the data frame so that none are left:

```
library(dplyr)
```

```
#create new data frame that removes duplicates so  
none are left
```

```
new_df <- df %>%
```

```
group_by(across(everything())) %>%  
filter(n()==1)
```

```
#view new data frame  
new_df
```

```
# A tibble: 4 x 2  
# Groups: team, points  
team points
```

```
1 A 28
```

```
2 A 14
```

```
3 B 18
```

```
4 B 27
```

Notice that each of the duplicate rows have been removed from the data frame and none of the duplicates remain.

Also notice that this produces the same result as the previous method.

**Note:** For extremely large data frames, the dplyr method will be faster than the base R method.

**The following tutorials explain how to perform other common functions in R:**

ARABPSYCHOLOGY.COM