Tutorials ▾   Exercises ▾   Services ▾   🔍   ◐   Sign Up   Log in

☰   .   CSS   JAVASCRIPT   SQL   PYTHON   JAVA   PHP   HOW TO   W3.CSS   C

Pandas - Removing Duplicates

❮ Previous   Next ❯

Discovering Duplicates

Duplicate rows are rows that have been registered more than one time.

```
     Duration          Date  Pulse  Maxpulse  Calories
0          60  '2020/12/01'    110       130     409.1
1          60  '2020/12/02'    117       145     479.0
2          60  '2020/12/03'    103       135     340.0
3          45  '2020/12/04'    109       175     282.4
4          45  '2020/12/05'    117       148     406.0
5          60  '2020/12/06'    102       127     300.0
6          60  '2020/12/07'    110       136     374.0
7         450  '2020/12/08'    104       134     253.3
8          30  '2020/12/09'    109       133     195.1
9          60  '2020/12/10'     98       124     269.0
10         60  '2020/12/11'    103       147     329.3
11         60  '2020/12/12'    100       120     250.7
12         60  '2020/12/12'    100       120     250.7
13         60  '2020/12/13'    106       128     345.3
14         60  '2020/12/14'    104       132     379.3
15         60  '2020/12/15'     98       123     275.0
16         60  '2020/12/16'     98       120     215.2
17         60  '2020/12/17'    100       120     300.0
18         45  '2020/12/18'     90       112       NaN
19         60  '2020/12/19'    103       123     323.0
20         45  '2020/12/20'     97       125     243.0
21         60  '2020/12/21'    108       131     364.2
22         45           NaN    100       119     282.0
23         60  '2020/12/23'    130       101     300.0
```

Tutorials ▾   Exercises ▾   Services ▾   🔍   ◐   Sign Up   Log in

☰   .   CSS   JAVASCRIPT   SQL   PYTHON   JAVA   PHP   HOW TO   W3.CSS   C

28	60	'2020/12/28'	103	132	NaN
29	60	'2020/12/29'	100	132	280.0
30	60	'2020/12/30'	102	129	380.3
31	60	'2020/12/31'	92	115	243.0

By taking a look at our test data set, we can assume that row 11 and 12 are duplicates.

To discover duplicates, we can use the `duplicated()` method.

The `duplicated()` method returns a Boolean values for each row:

Example   Get your own Python Server

Returns `True` for every row that is a duplicate, otherwise `False`:

```
print(df.duplicated())
```

Try it Yourself »

Get Certified!

Complete the Pandas modules, do the exercises, take the exam, and you will become w3schools certified!

$10 ENROLL

Removing Duplicates

To remove duplicates, use the `drop_duplicates()` method.

```
df.drop_duplicates(inplace = True)
```

Try it Yourself »

Remember: The `(inplace = True)` will make sure that the method does NOT return a *new* DataFrame, but it will remove all duplicates from the *original* DataFrame.

Test Yourself With Exercises

Exercise:

Insert the correct syntax for removing rows with empty cells.

```
df.        ()
```

Submit Answer »

Start the Exercise

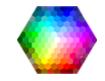❮ Previous   Log in to track progress   Next ❯

W3.CSS Tutorial   Bootstrap Reference
Bootstrap Tutorial   PHP Reference
PHP Tutorial   HTML Colors
Java Tutorial   Java Reference
C++ Tutorial   Angular Reference
jQuery Tutorial   jQuery Reference

Top Examples   ## Get Certified

HTML Examples   HTML Certificate
CSS Examples   CSS Certificate
JavaScript Examples   JavaScript Certificate
How To Examples   Front End Certificate
SQL Examples   SQL Certificate
Python Examples   Python Certificate
W3.CSS Examples   PHP Certificate
Bootstrap Examples   jQuery Certificate
PHP Examples   Java Certificate
Java Examples   C++ Certificate
XML Examples   C# Certificate
jQuery Examples   XML Certificate

[f]   [discord]   [in]   [instagram]   FORUM   ABOUT