



Pandas - Fixing Wrong Data

[< Previous](#)[Next >](#)

Wrong Data

"Wrong data" does not have to be "empty cells" or "wrong format", it can just be wrong, like if someone registered "199" instead of "1.99".

Sometimes you can spot wrong data by looking at the data set, because you have an expectation of what it should be.

If you take a look at our data set, you can see that in row 7, the duration is 450, but for all the other rows the duration is between 30 and 60.

It doesn't have to be wrong, but taking in consideration that this is the data set of someone's workout sessions, we conclude with the fact that this person did not work out in 450 minutes.

	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



Tutorials ▼

Exercises ▼

Services ▼



Sign Up

Log in

≡	.	CSS	JAVASCRIPT	SQL	PYTHON	JAVA	PHP	HOW TO	W3.CSS	C
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				
24		45	'2020/12/24'	105	132	246.0				
25		60	'2020/12/25'	102	126	334.5				
26		60	20201226	100	120	250.0				
27		60	'2020/12/27'	92	118	241.0				
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				

How can we fix wrong values, like the one for "Duration" in row 7?

Get Certified!

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

\$10 ENROLL



Replacing Values

One way to fix wrong values is to replace them with something else.

In our example, it is most likely a typo, and the value should be "45" instead of "450", and we could just insert "45" in row 7:



```
df.loc[7, 'Duration'] = 45
```

Try it Yourself »

For small data sets you might be able to replace the wrong data one by one, but not for big data sets.

To replace wrong data for larger data sets you can create some rules, e.g. set some boundaries for legal values, and replace any values that are outside of the boundaries.

Example

Loop through all values in the "Duration" column.

If the value is higher than 120, set it to 120:

```
for x in df.index:
    if df.loc[x, "Duration"] > 120:
        df.loc[x, "Duration"] = 120
```

Try it Yourself »

Removing Rows

Another way of handling wrong data is to remove the rows that contains wrong data.

This way you do not have to find out what to replace them with, and there is a good chance you do not need them to do your analyses.

Example

[Tutorials ▼](#)[Exercises ▼](#)[Services ▼](#)[Sign Up](#)[Log in](#)[CSS](#)[JAVASCRIPT](#)[SQL](#)[PYTHON](#)[JAVA](#)[PHP](#)[HOW TO](#)[W3.CSS](#)[C](#)

```
df.drop(x, inplace = True)
```

[Try it Yourself »](#)[◀ Previous](#)[Log in to track progress](#)[Next ▶](#)

COLOR PICKER



[Tutorials ▼](#)[Exercises ▼](#)[Services ▼](#)[Sign Up](#)[Log in](#)[CSS](#)[JAVASCRIPT](#)[SQL](#)[PYTHON](#)[JAVA](#)[PHP](#)[HOW TO](#)[W3.CSS](#)[C](#)[SPACES](#)[UPGRADE](#)[AD-FREE](#)[NEWSLETTER](#)[GET CERTIFIED](#)[REPORT ERROR](#)

Top Tutorials

[HTML Tutorial](#)
[CSS Tutorial](#)
[JavaScript Tutorial](#)
[How To Tutorial](#)
[SQL Tutorial](#)
[Python Tutorial](#)
[W3.CSS Tutorial](#)
[Bootstrap Tutorial](#)
[PHP Tutorial](#)
[Java Tutorial](#)
[C++ Tutorial](#)
[jQuery Tutorial](#)

Top Examples

[HTML Examples](#)
[CSS Examples](#)
[JavaScript Examples](#)
[How To Examples](#)
[SQL Examples](#)
[Python Examples](#)
[W3.CSS Examples](#)
[Bootstrap Examples](#)
[PHP Examples](#)
[Java Examples](#)
[XML Examples](#)
[jQuery Examples](#)

Top References

[HTML Reference](#)
[CSS Reference](#)
[JavaScript Reference](#)
[SQL Reference](#)
[Python Reference](#)
[W3.CSS Reference](#)
[Bootstrap Reference](#)
[PHP Reference](#)
[HTML Colors](#)
[Java Reference](#)
[Angular Reference](#)
[jQuery Reference](#)

Get Certified

[HTML Certificate](#)
[CSS Certificate](#)
[JavaScript Certificate](#)
[Front End Certificate](#)
[SQL Certificate](#)
[Python Certificate](#)
[PHP Certificate](#)
[jQuery Certificate](#)
[Java Certificate](#)
[C++ Certificate](#)
[C# Certificate](#)
[XML Certificate](#)

[FORUM](#)[ABOUT](#)

W3Schools is optimized for learning and training. Examples might be simplified to improve reading and learning. Tutorials, references, and examples are constantly reviewed to avoid errors, but we cannot warrant full correctness of all content. While using W3Schools, you agree to have read and accepted our [terms of use](#), [cookie and privacy policy](#).

Copyright 1999-2023 by Refsnes Data. All Rights Reserved. W3Schools is Powered by



Tutorials ▼

Exercises ▼

Services ▼



Sign Up

Log in



CSS

JAVASCRIPT

SQL

PYTHON

JAVA

PHP

HOW TO

W3.CSS

C