

Pandas - Cleaning Empty Cells

Previous

Next >

Empty Cells

Empty cells can potentially give you a wrong result when you analyze data.

Remove Rows

One way to deal with empty cells is to remove rows that contain empty cells.

This is usually OK, since data sets can be very big, and removing a few rows will not have a big impact on the result.

Example

Get your own Python Server

Return a new Data Frame with no empty cells:

```
import pandas as pd

df = pd.read_csv('data.csv')

new_df = df.dropna()

print(new_df.to_string())
```



Note: By default, the dropna() method returns a *new* DataFrame, and will not change the original.

If you want to change the original DataFrame, use the inplace = True argument:

Example

Remove all rows with NULL values:

```
import pandas as pd

df = pd.read_csv('data.csv')

df.dropna(inplace = True)

print(df.to_string())

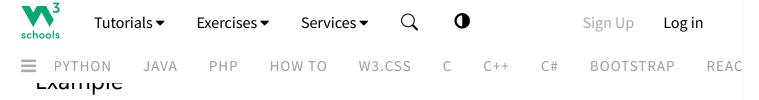
Try it Yourself »
```

Note: Now, the dropna(inplace = True) will NOT return a new DataFrame, but it will remove all rows containing NULL values from the original DataFrame.

Replace Empty Values

Another way of dealing with empty cells is to insert a *new* value instead.

This way you do not have to delete entire rows just because of some empty cells.



Replace NULL values with the number 130:

```
import pandas as pd

df = pd.read_csv('data.csv')

df.fillna(130, inplace = True)

Try it Yourself »
```

Replace Only For Specified Columns

The example above replaces all empty cells in the whole Data Frame.

To only replace empty values for one column, specify the *column name* for the DataFrame:

Example

Replace NULL values in the "Calories" columns with the number 130:

```
import pandas as pd

df = pd.read_csv('data.csv')

df.fillna({"Calories": 130}, inplace=True)

Try it Yourself »
```

Replace Using Mean, Median, or Mode



values for a specified column:

Example

Calculate the MEAN, and replace any empty values with it:

```
import pandas as pd

df = pd.read_csv('data.csv')

x = df["Calories"].mean()

df.fillna({"Calories": x}, inplace=True)

Try it Yourself »
```

Mean = the average value (the sum of all values divided by number of values).

Example

Calculate the MEDIAN, and replace any empty values with it:

```
import pandas as pd

df = pd.read_csv('data.csv')

x = df["Calories"].median()

df.fillna({"Calories": x}, inplace=True)
```



Median = the value in the middle, after you have sorted all values ascending.

Example

Try it Yourself »

Calculate the MODE, and replace any empty values with it:

```
import pandas as pd

df = pd.read_csv('data.csv')

x = df["Calories"].mode()[0]

df.fillna({"Calories": x}, inplace=True)
```

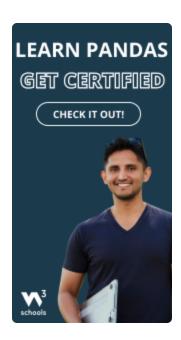
Mode = the value that appears most frequently.

⟨ Previous
Next >

Track your progress - it's free!

Sign Up Log in





COLOR PICKER

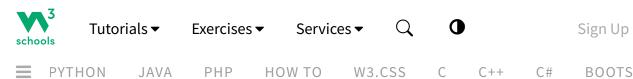






PLUS

SPACES



FOR BUSINESS

CONTACT US

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 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

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 ¡Query Examples**

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**







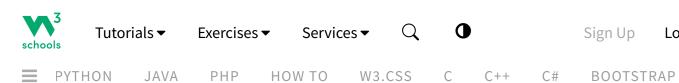




Log in

REAC

BOOTSTRAP



warrant full correctness

of all content. While using W3Schools, you agree to have read and accepted our <u>terms of use</u>, <u>cookie and privacy policy</u>.

Copyright 1999-2025 by Refsnes Data. All Rights Reserved. W3Schools is Powered by W3.CSS.

Log in

REAC