

How to get the number of the most frequent value in a column?

[Ask Question](#)

I have a data frame and I would like to know how many times a given column has the most frequent value.

I try to do it in the following way:

```
items_counts = df['item'].value_counts()
max_item = items_counts.max()
```

As a result I get:

ValueError: cannot convert float NaN to integer

As far as I understand, with the first line I get series in which the values from a column are used as key and frequency of these values are used as values. So, I just need to find the largest value in the series and, because of some reason, it does not work. Does anybody know how this problem can be solved?

[pandas](#) [frequency](#) [series](#)

asked Feb 28 '13 at 15:11



[Roman](#)

22.9k 118 261 352

Are there `na` 's in your column? If so you should get rid of them with `dropna` or `fillna` . — [beardc](#) Feb 28 '13 at 15:26

3 Answers

It looks like you may have some nulls in the column. You can drop them with `df = df.dropna(subset=['item'])` . Then

This site uses cookies to deliver our services and to show you relevant ads and job listings. By using our site, you acknowledge that you have read and understand our [Cookie Policy](#), [Privacy Policy](#), and our [Terms of Service](#). Your use of Stack Overflow's Products and Services, including the Stack Overflow Network, is subject to these policies and terms.

answered Feb 28 '13 at 15:43



beardc

6,996 8 54 68

And... can I do it so that NA is counted as a value? I.e. I am happy to return NA if that is the most common value. –

[Full Decent](#) Jul 29 '17 at 22:47

@FullDecent use `.fillna()` instead of `.dropna()` – [beardc](#) Jul 31 '17 at 18:37



You may also consider using scipy's `mode` function which ignores NaN. A solution using it could look like:

```
from scipy.stats import mode
from numpy import nan
df = DataFrame({"a": [1,2,2,4,2], "b":
print mode(df)
```

The output would look like

```
(array([[ 2.,  3.]]), array([[ 3.,  2.]
```

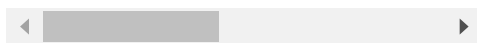
meaning that the most common values are 2 for the first columns and 3 for the second, with frequencies 3 and 2 respectively.

answered May 5 '15 at 22:00



jonathanrocher

1,032 6 8

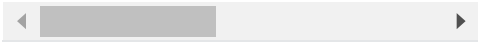


To continue to [@jonathanrocher](#) answer you could use `mode` in pandas DataFrame. It'll give a most frequent values (one or two) across the rows or columns:

```
import pandas as pd
import numpy as np
df = pd.DataFrame({"a": [1,2,2,4,2], '

In [2]: df.mode()
Out[2]:
   a    b
0  2  3.0
```

This site uses cookies to deliver our services and to show you relevant ads and job listings. By using our site, you acknowledge that you have read and understand our , and our . Your use of Stack Overflow's Products and Services, including the Stack Overflow Network, is subject to these policies and terms.



This site uses cookies to deliver our services and to show you relevant ads and job listings. By using our site, you acknowledge that you have read and understand our , , and our . Your use of Stack Overflow's Products and Services, including the Stack Overflow Network, is subject to these policies and terms.