pandas : update value if condition in 3 columns are met

Ask Question

25

5

I have a dataframe like this:

```
In[1]: df
Out[1]:
      A      B       C            D
1   blue    red    square        NaN
2  orange  yellow  circle        NaN
3  black   grey    circle        NaN
```

and I want to update column D when it meets 3 conditions. Ex:

```
df.ix[ np.logical_and(df.A=='blue', df.B=='red', df.C=='square'), ['D'] ] =
'succeed'
```

It works for the first two conditions, but it doesn't work for the third, thus:

```
df.ix[ np.logical_and(df.A=='blue', df.B=='red', df.C=='triangle'), ['D'] ] =
'succeed'
```

has exactly the same result:

```
In[1]: df
Out[1]:
      A      B       C            D
1   blue    red    square        succeed
2  orange  yellow  circle        NaN
3  black   grey    circle        NaN
```

python pandas

asked Jan 21 '14 at 15:57

Eduardo Oliveira

126 1 3 4

1 does `df.ix[(df.A=='blue') & (df.B=='red') & (df.C=='square'),'D']='success'` work? – EdChum Jan 21 '14 at 16:03

4 Answers

23

You could try this instead:

```
df[ (df.A=='blue') & (df.B=='red')
```

answered Jan 21 '14 at 16:02

Tim
914 2 13 35

Yes, it works. Though I still don't get the difference between "numpy.logical_and" and "&". Thank you – Eduardo Oliveira Jan 21 '14 at 16:30

also you can use `or` in a query like `df[ (df.A=='blue') | (df.B=='red') ]` – Tim Jan 21 '14 at 17:11

16 If getting warning `A value is trying to be set on a copy of a slice from a DataFrame.` while using above solution then do this: `df.loc[ (df.A=='blue') & (df.B=='red') & (df.C=='square'), 'D'] = 'succeed'` – Shams Mar 15 '17 at 5:57

1 I get the `A value is trying to be set on a copy of a slice from a DataFrame.` error aa well. Is there another method to do this? – Meghna Natraj Jul 2 '17 at 3:40

2 Just as a side note to anyone who runs into this and thinks they're going to trim it down...you MUST have the parentheses around the conditions or operator precedence will cause issues. Voice of experience. :-P – Brent Writes Code Apr 2 '18 at 2:45

20

Using

```
df[ (df.A=='blue') & (df.B=='red')
```

Gives below warning

```
/usr/local/lib/python2.7/dist-pack
SettingWithCopyWarning:
A value is trying to be set on a c
Try using .loc[row_indexer,col_ind
```

Better way of achieving this seems to be answered Aug 12 '17 at 14:38

Praveen
918 8 15

1 df.loc is much faster way than the previous one. – R.K. Apr 19 '18 at 18:36

4

You could try:

```
df['D'] = np.where((df.A=='blue')
```

This answer might provide a detailed answer to the your question: [Update row values where certain condition is met in pandas]()

edited Nov 1 '17 at 10:41

Aryan Firouzian
1,102 5 17 26

answered Nov 1 '17 at 10:04

theSanjeev
106 1 9

2

The third parameter of [logical_and]() is to assign the array used to store the result.

Currently, the method @TimRich provided might be the best. In pandas 0.13 (in development), there's a new experimental [query]() method. Try it!

answered Jan 21 '14 at 16:56

waitingkuo
38.6k 17 87 100