How to do a conditional count after groupby on a Pandas Dataframe?

Asked 2 years, 11 months ago   Active 4 months ago   Viewed 37k times

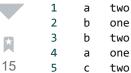I have the following dataframe:

```
     key1  key2
0     a    one
1     a    two
2     b    one
3     b    two
4     a    one
5     c    two
```

25

15

Now, I want to group the dataframe by the `key1` and count the column `key2` with the value `"one"` to get this result:

```
     key1
0     a    2
1     b    1
2     c    0
```

I just get the usual count with:

```
df.groupby(['key1']).size()
```

But I don't know how to insert the condition.

I tried things like this:

```
df.groupby(['key1']).apply(df[df['key2'] == 'one'])
```

But I can't get any further. How can I do this?

python   pandas   dataframe   pandas-groupby

edited Jul 12 '19 at 23:05   asked Aug 18 '17 at 8:58

EloyRD   Sethias
84   11   **270**   1   3   6

6 Answers

Active | Oldest | Votes

I think you need add condition first:

```
(x=='one').sum()).reset_index(name='count')
print (df11)
  key1  count
0   a      2
1   b      1
2   c      0
```

Or use [categorical](#) with `key1` , then missing value is added by `size` :

```
df['key1'] = df['key1'].astype('category')
df1 = df[df['key2'] == 'one'].groupby(['key1']).size().reset_index(name='count')
print (df1)
  key1  count
0   a      2
1   b      1
2   c      0
```

If need all combinations:

```
df2 = df.groupby(['key1', 'key2']).size().reset_index(name='count')
print (df2)
  key1 key2  count
0   a  one      2
1   a  two      1
2   b  one      1
3   b  two      1
4   c  two      1

df3 = df.groupby(['key1', 'key2']).size().unstack(fill_value=0)
print (df3)
key2  one  two
key1
a      2    1
b      1    1
c      0    1
```

edited Aug 18 '17 at 9:15   answered Aug 18 '17 at 9:00

jezrael
525k 46   657   677

oh i forgot that. I got that as well. But then the Key1 with value C is not respected – Sethias Aug 18 '17 at 9:01 ✎

You can count the occurence of 'one' for the groupby dataframe, in the column 'key2' like this:

```
df.groupby('key1')['key2'].apply(lambda x: x[x == 'one'].count())
```

7

yield

```
key1
a    2
b    1
```

answered Aug 18 '17 at 9:05

Florian Mutel
739   4   11

Option 1

4

```
df.set_index('key1').key2.eq('one').sum(level=0).astype(int).reset_index()
```

```
   key1  key2
0    a     2
1    b     1
2    c     0
```

Option 2

```
df.key2.eq('one').groupby(df.key1).sum().astype(int).reset_index()
```

```
   key1  key2
0    a     2
1    b     1
2    c     0
```

Option 3

```
f, u = df.key1.factorize()
pd.DataFrame(dict(key1=u, key2=np.bincount(f, df.key2.eq('one')).astype(int)))
```

```
   key1  key2
0    a     2
1    b     1
2    c     0
```

Option 4

```
pd.crosstab(df.key1, df.key2.eq('one'))[True].rename('key2').reset_index()
```

```
   key1  key2
0    a     2
1    b     1
2    c     0
```

Option 5

```
pd.get_dummies(df.key1).mul(
    df.key2.eq('one'), 0
).sum().rename_axis('key1').reset_index(name='key2')
```

```
1    b    1
2    c    0
```

answered Aug 18 '17 at 9:10

piRSquared
214k   31   273   442

2

You can do this with applying groupby() on both keys and unstack().

```
df = df.groupby(['key1', 'key2']).size().unstack()
```
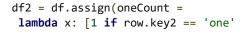
answered Aug 15 '18 at 16:47

Mehdi Golari
431   3   8

2

Maybe not the fastest solution, but you can create new data frame with column of ones if key2 is equal to 'one'.

```
df2 = df.assign(oneCount =
 lambda x: [1 if row.key2 == 'one' else 0 for index, row in x.iterrows()])
```

```
   key1 key2  oneCount
0    a   one         1
1    a   two         0
2    b   one         1
3    b   two         0
4    a   one         1
5    c   two         0
```

And then aggregate it.

```
df3 = df2.groupby('key1').agg({"oneCount":sum}).reset_index()
```

```
  key1   oneCount
0    a          2
1    b          1
2    c          0
```

answered Jun 9 '19 at 13:33

Igor Kołakowski
21   2

0

I need count 2 columns (lambda with two arguments) as the example:

`Pandas dataframe groupby func`, in the column `key2` like this:

edited Feb 24 at 13:24

shiva
2,585　4　11　34

answered Oct 10 '19 at 3:29

Andre Vieira de Lima
43　3