Pandas group-by and sum

Ask Question

83

19

I am using this data frame:

Fruit	Date	Name	Number
Apples	10/6/2016	Bob	7
Apples	10/6/2016	Bob	8
Apples	10/6/2016	Mike	9
Apples	10/7/2016	Steve	10
Apples	10/7/2016	Bob	1
Oranges	10/7/2016	Bob	2
Oranges	10/6/2016	Tom	15
Oranges	10/6/2016	Mike	57
Oranges	10/6/2016	Bob	65
Oranges	10/7/2016	Tony	1
Grapes	10/7/2016	Bob	1
Grapes	10/7/2016	Tom	87
Grapes	10/7/2016	Bob	22
Grapes	10/7/2016	Bob	12
Grapes	10/7/2016	Tony	15

I want to aggregate this by name and then by fruit to get a total number of fruit per name.

```
Bob,Apples,16 ( for example )
```

I tried grouping by Name and Fruit but how do I get the total number of fruit.

python pandas dataframe group-by aggregate

edited Oct 4 '18 at 19:26

mit

6,153 6 34 60

asked Oct 7 '16 at 17:36

Trying_hard

1,921 14 43 68

8 Answers

94

use the sum() method

```
df.groupby(['Fruit','Name']).sum(
```

```
Out[31]:
```

```
          Steve      10
Grapes    Bob        35
          Tom        87
          Tony       15
Oranges   Bob        67
          Mike       57
          Tom        15
          Tony        1
```

edited Oct 2 '17 at 16:20

answered Oct 7 '16 at 17:37

Steven G
6,326 2 23 42

23 How can pandas knows that I want to sum the col named `Number` ? – Kingname Oct 23 '17 at 12:32

3 @Kingname it's the last column left if you take out NAME and FRUIT. if you add 2 columns left, it would sum both columns – Steven G Oct 23 '17 at 16:51

4 @StevenG the `Date` column is also left... – Daniyal Shahrokhian May 20 '18 at 6:11

4 I find this solution a little hackish compared to the others – matanster Jul 24 '18 at 17:20

`Date` is not summed because it has `dtype = string` yes? – Wassadamo Sep 1 '18 at 2:28

80

Also you can use agg function,

```
df.groupby(['Name', 'Fruit'])['Num
```

answered Oct 8 '16 at 11:40

Saurabh
2,628 2 21 30

11 This should be the accepted answer, it is more explicit. – shahar_m Oct 4 '18 at 10:02

24

If you want to keep the original columns `Fruit` and `Name` , use `reset_index()` . Otherwise `Fruit`

```
df.groupby(['Fruit','Name'])['Numb
```

Fruit	Name	Number
Apples	Bob	16
Apples	Mike	9
Apples	Steve	10
Grapes	Bob	35
Grapes	Tom	87
Grapes	Tony	15
Oranges	Bob	67
Oranges	Mike	57
Oranges	Tom	15
Oranges	Tony	1

As seen in the other answers:

```
df.groupby(['Fruit','Name'])['Numb
```

		Number
Fruit	Name	
Apples	Bob	16
	Mike	9
	Steve	10
Grapes	Bob	35
	Tom	87
	Tony	15
Oranges	Bob	67
	Mike	57
	Tom	15
	Tony	1

answered Jul 2 '18 at 10:01

Gazala Muhamed

340 3 4

22

Both the other answers accomplish what you want.

You can use the `pivot` functionality to arrange the data in a nice table

```
df.groupby(['Fruit','Name'],as_ind
False).sum().pivot('Fruit','Name')
```

Name	Bob	Mike	Steve	To
Fruit				
Apples	16.0	9.0	10.0	0.
Grapes	35.0	0.0	0.0	87
Oranges	67.0	57.0	0.0	15

answered Oct 7 '16 at 18:35

Demetri Pananos

2,132 14 34

10

```
df.groupby(['Fruit','Name'])['Numb
```

You can select different columns to sum numbers.

6

You can use `groupby` and `sum`:

```
df.groupby(['Name', 'Fruit']).sum(
```

```
              Number
Name  Fruit
Bob   Apples      16
      Grapes      35
      Oranges     67
Mike  Apples       9
      Oranges     57
Steve Apples      10
Tom   Grapes      87
      Oranges     15
Tony  Grapes      15
      Oranges      1
```

answered Oct 7 '16 at 17:44

Batsu
2,012 15 29

1

This answer only to understand how groupby and sum works.

I am using data-set "Rainfall in India since 1900 to 2015"

My Dataset includes columns like "subdivision" and "annual".

So, here i would like to calculate sum of annual rainfall for each subdivision.

```
Total = Data.groupby('SUBDIVISION'
print (Total)
```

this is how subdivision will get grouped and we will get sum of annual rainfall per subdivision.

answered Aug 3 '18 at 10:06

vaibhav pawar
59 5

1

You can set the `groupby` column to `index` then using `sum` with `level`

```
df.set_index(['Fruit','Name']).sum
Out[175]:
```

```
              Number
Fruit   Name
Apples  Bob       16
        Mike       9
        Steve     10
Oranges Bob       67
        Tom       15
        Mike      57
        Tony       1
```

answered Nov 21 '18 at 3:01

Wen-Ben
124k 8 36 71