## Python Pandas sort by Time and group by user ID

Asked 2 years, 8 months ago Active 2 years, 8 months ago Viewed 3k times



I am loading a CSV file with pandas. It has three columns: a column with date and time, a column with a user id, and another 'campaignID'. Example rows:





```
date user_id campaign_id 2018-01-10 0:21:09 151312395 GOOGLE 2018-01-10 0:21:19 151312395 GOOGLE 2018-01-10 0:21:32 151312395 GOOGLE
```



I want to group the data by the user id, and then for each user id group the rows by time and the campaign ID, it should look as follows.

```
      user_id
      date
      ad_campaign

      151312395
      2018-01-10 0:21:09
      GOOGLE

      2018-01-10 0:21:19
      GOOGLE

      2018-01-10 0:21:32
      GOOGLE
```

This is what I have made until now: import pandas as pd import numpy as np import datetime

```
def dateparse(time_in_secs):
    return datetime.datetime.fromtimestamp(float(time_in_secs))
columnnames = ['date','user_id', 'ad_campaign']
columnnames, sep='\t', usecols=[0,1,3],index_col = 'date')
df=pd.read_csv(r'C:\Users\L\Desktop\Data.csv',
    sep='\t',names = columnnames, usecols=[0,1,3],
    parse_dates=True,date_parser=dateparse)
df.date = pd.to_datetime(df.date)
df = df.sort_values(by = 'date')
g = df.groupby('user_id')['ad_campaign']
print(g)
```

This gives the following output:

```
<pandas.core.groupby.SeriesGroupBy object at 0x04EF26F0>
[Finished in 0.6s]
```

Why doesnt the print provide the sorted columns?

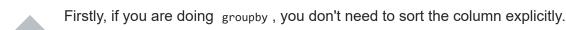
```
python pandas csv pandas-groupby Edit tags
```

asked Apr 26 '18 at 14:18

Laila Van Ments

125 1 12





3 You can do:



Method 1:



```
df.date = pd.to_datetime(df.date)
g = df.groupby(['user_id','date'])['ad_campaign']
print(g.first())
```

## Method 2:

```
df.set_index(['user_id','date']).sort_index()
```

answered Apr 26 '18 at 15:12



I don't know if I'm allowed to ask a follow up question here, but even though it works I am stuck at my following step.. – Laila Van Ments Apr 29 '18 at 13:44



@LailaVanMents I don't know either, but if it's the same question you can ask. What's that? - YOLO Apr 29 '18 at 13:51





You could try df.set\_index(['user\_id', 'date']).





edited Apr 26 '18 at 14:26

answered Apr 26 '18 at 14:21



Victor Valente **611** 9 19

