Pandas: Convert JSON to Pandas Dataframe

Ask Question

Asked 9 months ago
Active 9 months ago
Viewed 869 times

2

Experts,

I am having issues parsing Json to Pandas and then save it in CSV format.

```
data2 = {"date":"2018-01-02","data":{"AAPL":{"open":"170.16","close":"172.26","high":"172.30","low":"169.26","volume":"25555934"},"MS
```

If I try :

```
df = pd.DataFrame.from_dict(json_normalize(data2), orient='columns')
print(df)
```

Everything gets printed in a single line :

```
                data.AAPL.close  data.AAPL.high  data.AAPL.low  data.AAPL.open   ...    data.MSFT.low  data.MSFT.open  data.MSFT.volume        date
0                      172.26          172.30         169.26          170.16   ...            85.50           86.13          22483797  2018-01-02
```

If I do :

```
jdata = json.loads(data2)
df = pd.DataFrame(jdata)
print(df.T)
```

I get an error : TypeError: the JSON object must be str, bytes or bytearray, not 'dict'

I am looking to print it in the following table format so I can then save it as csv :

```
Date        Data      Open    Close   High    Low     Volume
2018-01-02  AAPL      170.16  172.26  172.30  169.26  25555934
2018-01-02  MSFT      86.13   85.95   86.31   85.50   22483797
```

What is the correct way to achieve my goal ?

Thanks !!

`json` `python-3.x` `pandas` `dataframe`

share improve this question

asked Dec 13 '18 at 2:46

Eric Arambula
90 ● 9

add a comment

2 Answers

active   oldest   **votes**

2

✔

You can use `apply` to turn the dict keys into pandas `Series`

```
df = pd.DataFrame.from_dict(data2)
df = pd.concat([df['date'],df['data'].apply(pd.Series)], axis=1)
print(df)

           date      open    close    high     low     volume
AAPL  2018-01-02  170.16  172.26  172.30  169.26  25555934
MSFT  2018-01-02   86.13   85.95   86.31   85.50  22483797
```

share improve this answer

answered Dec 13 '18 at 3:04

zrelova
758 ● 1 ● 5 ● 26

@W-B, let me pick your brain, what is the difference vs the method you proposed ? What are the downsides ? – Eric Arambula Dec 13 '18 at 3:22

@EricArambula speed, apply is timing cost function , since you have stock data , when the size is increase , you will find apply + pd.Serise slow – WeNYoBen Dec 13 '18 at 3:22

add a comment

▲

1

▼

I will using your original output and modify it

```
s=pd.DataFrame(data2)
pd.concat([s.drop('data',1),pd.DataFrame(s.data.tolist(),index=s.index)],1)
          date   close   high    low    open    volume
AAPL  2018-01-02  172.26  172.30  169.26  170.16  25555934
MSFT  2018-01-02   85.95   86.31   85.50   86.13  22483797
```

share improve this answer

answered Dec 13 '18 at 3:01

WeNYoBen
159k ●8 ●55 ●86

add a comment

Your Answer

Post Your Answer

By clicking "Post Your Answer", you agree to our terms of service, privacy policy and cookie policy.

Not the answer you're looking for? Browse other questions tagged json python-3.x pandas dataframe or ask your own question.

Related

6895

Can comments be used in JSON?

2840

How can I pretty-print JSON in a shell script?

9861

What is the correct JSON content type?

3900

Why does Google prepend while(1); to their JSON responses?

874

Selecting multiple columns in a pandas dataframe

`1521`

Renaming columns in pandas

`1069`

Delete column from pandas DataFrame

`1404`

How to iterate over rows in a DataFrame in Pandas?

`1528`

Select rows from a DataFrame based on values in a column in pandas

`870`

Get list from pandas DataFrame column headers

Hot Network Questions

- Plausibility and performance of a composite longbow
- Can a business put whatever they want into a contract?
- Exam design: give maximum score per question or not?
- What is this WWII four-engine plane on skis?
- Is Zack Morris's 'time stop' ability in "Saved By the Bell" a supernatural ability?

more hot questions

Question feed