How can we convert a date column in a Pandas dataframe to the YYYYMMDD format?

Authored by

stats writer

June 26, 2024

To convert a date column in a Pandas dataframe to the YYYYMMDD format, we can use the strftime() function. This function allows us to specify the desired date format, in this case "%Y%m%d", which represents the year, month, and day in a numerical format without any separators. This will convert the date column into a string format in the desired format. Additionally, we can use the to_datetime() function to convert the string dates back to datetime objects if needed. This method is useful for organizing and sorting date data in a Pandas dataframe.

Pandas: Convert Date to YYYYMMDD Format

You can use the following syntax to convert a date column in a pandas DataFrame to a YYYYMMDD format:

```
#convert date column to datetime
df = pd.to_datetime(df)

#convert date to YYYYMMDD format
df = df.dt.strftime('%Y%m%d').astype(int)
```

The following example shows how to use this syntax in practice.

Example: Convert Date to YYYYMMDD Format in Pandas

Suppose we have the following pandas DataFrame that shows the sales made by some company on various dates:

```
import pandas as pd
```

```
#create DataFrame
df = pd.DataFrame({'date': pd.date_range(start='1/1/2022', freq='MS', periods=8), 'sales': })

#view DataFrame
print(df)
```

date sales
0 2022-01-01 18
1 2022-02-01 22
2 2022-03-01 19
3 2022-04-01 14
4 2022-05-01 14
5 2022-06-01 11
6 2022-07-01 20
7 2022-08-01 28

Now suppose that we would like to format the values in the date column as YYYYMMDD.

We can use the following syntax to do so:

```
#convert date column to datetime
df = pd.to_datetime(df)
```

```
#convert date to YYYYMMDD format
df = df.dt.strftime('%Y%m%d').astype(int)

#view updated DataFrame
print(df)

date sales
0 20220101 18
1 20220201 22
2 20220301 19
3 20220401 14
4 20220501 14
5 20220601 11
6 20220701 20
7 20220801 28
```

Notice that the values in the date column are now formatted in a YYYYMMDD format.

Note that in this example, the date column already had a class of datetime.

However, we can use the to_datetime() function anyway to ensure that a given column has a class of datetime before applying a YYYYMMDD format.

The following tutorials explain how to perform other common operations in pandas: