How can I use Pandas to retrieve the day of the year from a given date?

Authored by

stats writer

June 26, 2024

Pandas is a popular library for data analysis in Python. It provides various functionalities for manipulating and analyzing data, including the ability to retrieve the day of the year from a given date. This can be achieved by using the "dayofyear" method, which converts a given date into its corresponding day of the year. This feature can be particularly useful for tasks involving time series analysis and data visualization. By utilizing Pandas, users can easily retrieve the day of the year from a given date, saving time and effort in data processing.

Pandas: Get Day of Year from Date

You can use the following basic syntax to get the day of year from a date column in a pandas DataFrame:

df = df.dt.dayofyear

This particular example creates a new column called day_of_year that contains the day of the year of the value in the date column.

Note that the values for day_of_year will range from 1 (January 1st) to 365 (December 31st).

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

Example: Get Day of Year from Date in Pandas

Suppose we have the following pandas DataFrame that contains information about the total sales made at

some store on various dates:

```
import pandas as pd

#create DataFrame
df = pd.DataFrame({'date': pd.date_range(start='1/1/2022', freq='M', periods=10), 'sales': })

#view DataFrame
print(df)
```

date sales
0 2022-01-31 6
1 2022-02-28 8
2 2022-03-31 10
3 2022-04-30 5
4 2022-05-31 4
5 2022-06-30 8
6 2022-07-31 8
7 2022-08-31 3
8 2022-09-30 5
9 2022-10-31 14

Related:

We can use the following code to create a new column called day_of_year that contains the day of the year from the date column:

#create new column that contains day of year in 'date' column
df = df.dt.dayofyear
#view updated DataFrame
print(df)

date sales day_of_year
0 2022-01-31 6 31
1 2022-02-28 8 59
2 2022-03-31 10 90
3 2022-04-30 5 120
4 2022-05-31 4 151
5 2022-06-30 8 181
6 2022-07-31 8 212
7 2022-08-31 3 243
8 2022-09-30 5 273
9 2022-10-31 14 304

The new column called day_of_year contains the day of year from the date column.

It's worth noting that if you're working with a leap year, this function will automatically extend the range of possible values from 365 to 366.

Also note that if the column you're working with is a string column, you must first use pd.to_datetime() to convert the strings to recognizable dates:

```
#convert string column to datetime and calculate day of year
df = pd.to_datetime(df).dt.dayofyear
```

Note: You can find the complete documentation for the pandas dayofyear function .