

How can I select specific rows of a Pandas DataFrame based on a specific timestamp?

Authored by
stats writer

June 26, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I select specific rows of a Pandas DataFrame based on a specific timestamp?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=153947>

Pandas DataFrame is a powerful tool for data analysis and manipulation in Python. It allows users to filter and select specific rows based on certain conditions. When dealing with time series data, it is often necessary to select rows based on a specific timestamp. This can be achieved by using the built-in functions and methods provided by Pandas. By specifying the desired timestamp, the user can filter the DataFrame to only include rows that match the given timestamp. This provides a convenient and efficient way to extract relevant data from a large dataset. Overall, Pandas offers a user-friendly and versatile solution for selecting specific rows based on a specific timestamp in a Pandas DataFrame.

Pandas: Select Rows of DataFrame by Timestamp

You can use the following basic syntax to select rows between two timestamps in a pandas DataFrame:

```
df > '2022-10-25 04:30:00') & (df < '2022-10-27 11:00:00']
```

This syntax assumes that timestamp already has a dtype of datetime.

If it doesn't, you can use the following syntax to convert it to a datetime column:

```
df = pd.to_datetime(df)
```

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

Example: Select Rows of Pandas DataFrame by Timestamp

Suppose we have the following pandas DataFrame that contains information about sales at some retail store:

```
import pandas as pd
```

```
#create DataFrame
```

```
df = pd.DataFrame({'tstamp': ,  
'sales': })
```

```
#view DataFrame
```

```
print(df)
```

```
tstamp sales
```

```
0 2022-10-25 04:00:00 18
```

```
1 2022-10-25 11:55:12 22
```

```
2 2022-10-26 02:00:00 19
```

```
3 2022-10-27 10:30:00 14
```

```
4 2022-10-27 14:25:00 14
```

```
5 2022-10-28 01:15:27 11
```

Suppose we would like to select only the rows between the following two timestamps:

```
2022-10-25 04:30:00 2022-10-27 11:00:00
```

We can use the following syntax to do so:

```
#convert timestamp column to datetime dtype
```

```
df = pd.to_datetime(df)
```

```
#select rows between two timestamps
```

```
df > '2022-10-25 04:30:00') & (df < '2022-10-27 11:00:00')]
```

```
timestamp sales
```

```
1 2022-10-25 11:55:12 22
```

```
2 2022-10-26 02:00:00 19
```

```
3 2022-10-27 10:30:00 14
```

Notice that only the rows between the two timestamps that we specified are selected.

Also note that you can select rows by timestamp using only a date value.

For example, we could use the following code to select all rows where the timestamp is greater than 2022-10-27:

```
#convert timestamp column to datetime dtype
```

```
df = pd.to_datetime(df)
```

```
#select rows with timestamp after 2022-10-27
```

```
df > '2022-10-27']
```

```
tstamp sales
```

```
3 2022-10-27 10:30:00 14
```

```
4 2022-10-27 14:25:00 14
```

```
5 2022-10-28 01:15:27 11
```

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

ARABPSYCHOLOGY.COM