

How can I select specific rows from a Pandas DataFrame using a boolean series?

Authored by
stats writer

June 26, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I select specific rows from a Pandas DataFrame using a boolean series?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=153666>

Pandas DataFrame is a popular data analysis library that allows users to manipulate and analyze large datasets easily. One of the common tasks in data analysis is selecting specific rows from a DataFrame based on certain conditions. This can be achieved by using a boolean series, which is a series of True and False values. By creating a boolean series that matches the desired rows, users can filter the DataFrame and select only the rows that meet the specified criteria. This feature of Pandas DataFrame provides a flexible and efficient way to extract relevant data for further analysis.

Pandas: Select Rows from DataFrame Using Boolean Series

You can use the following basic syntax to select rows in a pandas DataFrame based on values in a boolean series:

```
#define boolean series  
bools = pd.Series()
```

```
#select rows in DataFrame based on values in boolean series  
df
```

This allows you to select each of the rows in the pandas DataFrame where the corresponding value in the boolean series is True.

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

Example: Select Rows from Pandas DataFrame Using Boolean Series

Suppose we have the following pandas DataFrame that contains information about various basketball players:

```
import pandas as pd
```

```
#create DataFrame
```

```
df = pd.DataFrame({'team': ,  
'points': ,  
'assists': ,  
'rebounds': })
```

```
#view DataFrame
```

```
print(df)
```

```
team points assists rebounds
```

```
0 A 18 5 11
```

```
1 B 22 7 8
```

```
2 C 19 7 10
```

```
3 D 14 9 6
```

```
4 E 14 12 6
```

```
5 F 11 9 5
```

```
6 G 20 9 9
```

7 H 28 4 12

We can use the following syntax to select all rows in the DataFrame where the corresponding value in a boolean series is True:

#define boolean series

bools = pd.Series()

#select rows in DataFrame based on values in boolean series

df

team points assists rebounds

0 A 18 5 11

2 C 19 7 10

3 D 14 9 6

7 H 28 4 12

Notice that the only rows returned are the ones where the corresponding value in the boolean series is True.

Also note that you can use the following syntax to only select the rows in the "points" column of the DataFrame where the corresponding value in the boolean series is

True.

```
#define boolean series
```

```
bools = pd.Series()
```

```
#select rows in points column based on values in  
boolean series
```

```
df
```

```
0 18
```

```
2 19
```

```
3 14
```

```
7 28
```

```
Name: points, dtype: int64
```

Notice that the only rows returned from the "points" column are the ones where the corresponding value in the boolean series is True.

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