

How can I sort a Pandas DataFrame by both the index and a specific column?

Authored by
stats writer

April 29, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I sort a Pandas DataFrame by both the index and a specific column?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=141100>

To sort a Pandas DataFrame by both the index and a specific column, use the `sort_values()` function and specify the parameters `by` and `axis` to indicate the column and index, respectively. This will rearrange the rows of the DataFrame in ascending or descending order based on the values in the specified column and index. This method allows for efficient organization of data in a DataFrame, providing a useful tool for data analysis and manipulation.

Pandas: Sort DataFrame by Both Index and Column

You can use the following syntax to sort a pandas DataFrame by both index and column:

```
df = df.sort_values(by = , ascending = )
```

The following examples show how to use this syntax in practice.

Examples: Sort DataFrame by Both Index and Column

The following code shows how to sort a pandas DataFrame by the column named points and then by the index column:

```
import pandas as pd

#create DataFrame
df = pd.DataFrame({'id': ,
'points': ,
'assists': ,
```

```
'rebounds': }).set_index('id')
```

```
#view first few rows
```

```
df.head()
```

```
points assists rebounds
```

```
id
```

```
1 25 5 11
```

```
2 15 7 8
```

```
3 15 7 10
```

```
4 14 9 6
```

```
5 20 12 6
```

```
#sort by points and then by index
```

```
df.sort_values(by = , ascending = )
```

```
points assists rebounds
```

```
id
```

```
8 29 4 12
```

```
1 25 5 11
```

```
7 25 9 9
```

```
5 20 12 6
```

```
6 20 9 5
```

```
2 15 7 8
```

```
3 15 7 10
```

4 14 9 6

The resulting DataFrame is sorted by points in descending order and then by the index in ascending order (if there happen to be two players who score the same number of points).

Note that if we don't use the ascending argument, then each column will use ascending as the default sorting method:

```
#sort by points and then by index  
df.sort_values(by = )
```

```
points assists rebounds
```

```
id
```

```
4 14 9 6
```

```
2 15 7 8
```

```
3 15 7 10
```

```
5 20 12 6
```

```
6 20 9 5
```

```
1 25 5 11
```

```
7 25 9 9
```

```
8 29 4 12
```

If the index column is currently not named, you can rename it and then sort accordingly:

#sort by points and then by index

```
df.rename_axis('index').sort_values(by = )
```

points assists rebounds

id

4 14 9 6

2 15 7 8

3 15 7 10

5 20 12 6

6 20 9 5

1 25 5 11

7 25 9 9

8 29 4 12