How can the maximum value be found by group in Pandas?

Authored by

stats writer

April 23, 2024

The process of finding the maximum value by group in Pandas involves using the built-in groupby function to group data based on a specific column or set of columns. Then, the max() function can be applied to the grouped data to find the maximum value within each group. This allows for efficient and organized analysis of data, especially in large datasets. By grouping data and finding the maximum value within each group, insights can be gained on the distribution and patterns of the data. This approach is commonly used in data analysis and can be easily implemented in Pandas through its user-friendly syntax and functions.

Find the Max Value by Group in Pandas

Often you may be interested in finding the max value by group in a pandas DataFrame.

Fortunately this is easy to do using the groupby() and max() functions with the following syntax:

df.groupby('column_name').max()

This tutorial explains several examples of how to use this function in practice using the following pandas DataFrame:

import pandas as pd

#create pandas DataFrame
df = pd.DataFrame({'team': ,
'points':,
'rebounds': })

#display DataFrame
print(df)

team points rebounds
0 A 24 11
1 A 23 8
2 B 27 7
3 B 11 6
4 B 14 6
5 C 8 5
6 C 13 12

Example 1: Max Value of Multiple Columns Grouped by One Variable

The following code shows how to find the max value of multiple columns, grouped by one variable in a DataFrame:

#find max values of points and rebounds, grouped by team
df.groupby('team').max().reset_index()

team points rebounds
0 A 24 11
1 B 27 7

2 C 13 12

From the output we can see that:

Team A has a max *points* value of 24 and a max *rebounds* value of 11.Team B has a max *points* value of 27 and a max *rebounds* value of 7.Team C has a max *points* value of 13 and a max *rebounds* value of 12.

Note that we used the reset_index() function to ensure that the index matches the index in the original DataFrame.

Example 2: Max Value of a Single Column Grouped by One Variable

The following code shows how to find the max value of just one column, grouped on a single variable:

```
#find max value of points, grouped by team
df.groupby('team').max().reset_index()
```

```
team points
0 A 24
1 B 27
2 C 13
```

Example 3: Sort by Max Values

We can also use the sort_values() function to sort the max values.

We can specify ascending=False to sort from largest to smallest:

```
#find max value by team, sort descending
df.groupby('team').max().reset_index().sort_values(, ascending=False)
```

```
team points
1 B 27
0 A 24
2 C 13
```

Or we can specify ascending=True to sort from smallest to largest:

```
#find max value by team, sort ascending
df.groupby('team').max().reset_index().sort_values(, ascending=True)
```

```
team points
2 C 13
```

0 A 24

1 B 27

How to Calculate the Sum of Columns in Pandas

How to Calculate the Mean of Columns in Pandas

How to Find the Max Value of Columns in Pandas