How can I create a pivot table in Pandas with multiple aggfunc functions?

Authored by

stats writer

June 26, 2024

Creating a pivot table in Pandas with multiple aggfunc functions allows for efficient and customizable data analysis. By using the pivot_table() function in Pandas, users can summarize and group data based on specific variables and apply multiple aggregation functions to those variables simultaneously. This not only saves time and effort in data analysis, but also provides a comprehensive and insightful view of the data. This feature is particularly useful in data exploration and decision making processes.

Pandas: Create Pivot Table with Multiple aggfunc

You can use the following syntax to create a pivot table in pandas and provide multiple values to the aggfunc argument:

```
df.pivot_table(index='col1', values='col2', aggfunc=('sum', 'mean'))
```

This particular example creates a pivot table that displays the sum and the mean of values in col2, grouped by col1.

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

Example: Create Pandas Pivot Table with Multiple aggfunc

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': })

#view DataFrame
print(df)
```

team points assists
0 A 4 2
1 A 4 2
2 A 2 5
3 A 8 5
4 B 9 4
5 B 5 7
6 B 5 5
7 B 7 3
8 C 8 9
9 C 8 8
10 C 4 4
11 C 3 4

We can use the following code to create a pivot table

that summarizes both the sum and the mean number of points scored by each team:

```
#create pivot table to summarize sum and mean of points by team
df.pivot_table(index='team', values='points', aggfunc=('sum', 'mean'))
```

	mean	sum
team		
A	4.50	18
B	6.50	26
C	5.75	23

The resulting pivot table summarizes the mean and the sum of the points scored by each team.

For example, we can see:

Players on team A had a mean points value of 4.50 and a sum points value of 18.Players on team B had a mean points value of 6.50 and a sum points value of 26.Players on team C had a mean points value of 5.75 and a sum points value of 23.

Note that we aggregated using the sum and the mean in this example, but we could also aggregate by other metrics such as:

countminmaxmedianstd (standard deviation)

The following example shows how to aggregate the values in the points column by these metrics for each team:

#create pivot table to summarize several metrics for points by team
df.pivot_table(index='team', values='points',
aggfunc=('count', 'min', 'max', 'median', 'std'))

count max median min std
team
A 4 8 4.0 2 2.516611
B 4 9 6.0 5 1.914854
C 4 8 6.0 3 2.629956

Note: You can find the complete documentation for the pandas pivot_table() function .

The following tutorials explain how to perform other

common tasks in pandas: