

How can I find the median value by group in Pandas?

Authored by
stats writer

July 2, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I find the median value by group in Pandas?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=165904>

Finding the median value by group in Pandas refers to the process of determining the middle value in a group of data points, organized into groups by a specific category or feature, using the Pandas library in Python. This can be achieved by grouping the data based on the desired feature and then calculating the median for each group. This method allows for efficient analysis and comparison of data within different groups.

Find the Median Value by Group in Pandas

You can use the following basic syntax to calculate the median value by group in pandas:

```
df.groupby().median().reset_index()
```

You can also use the following syntax to calculate the median value, grouped by several columns:

```
df.groupby().median().reset_index()
```

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

Example 1: Find Median Value by One Group

Suppose we have the following pandas DataFrames:

```
import pandas as pd
```

```
#create DataFrame
```

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

```
#view DataFrame
```

```
df
```

```
team position points rebounds
```

```
0 A G 5 11
```

```
1 A G 7 8
```

```
2 A F 7 10
```

```
3 A F 9 6
```

```
4 B G 12 6
```

```
5 B G 9 5
```

```
6 B F 9 9
```

```
7 B F 4 12
```

We can use the following code to find the median value of the 'points' column, grouped by team:

```
#calculate median points by team
```

```
df.groupby().median().reset_index()
```

```
team points
```

```
0 A 7.0
```

```
1 B 9.0
```

From the output we can see:

The median points scored by players on team A is 7. The median points scored by players on team B is 9.

Note that we can also find the median value of two variables at once:

```
#calculate median points and median rebounds by team  
df.groupby().median()
```

```
team points rebounds
```

```
0 A 7.0 9.0
```

```
1 B 9.0 7.5
```

Example 2: Find Median Value by Multiple Groups

The following code shows how to find the median value of the 'points' column, grouped by team and position:

```
#calculate median points by team  
df.groupby().median().reset_index()
```

team position points

0 A F 8.0

1 A G 6.0

2 B F 6.5

3 B G 10.5

The median points scored by players in the 'F' position on team A is 8. The median points scored by players in the 'G' position on team A is 6. The median points scored by players in the 'F' position on team B is 6.5. The median points scored by players in the 'G' position on team B is 10.5.

Additional Resources

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