

How do I use the `idxmax()` function in Pandas?

Authored by
stats writer

May 1, 2024

RECOMMENDED CITATION

stats writer (2024). *How do I use the `idxmax()` function in Pandas?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=141563>

The `idxmax()` function in Pandas is used to find the index of the maximum value in a given column or row of a dataframe. This function takes in a parameter, `axis`, which specifies whether to find the maximum value along the column (`axis=0`) or row (`axis=1`). Once the maximum value is determined, the function returns the index of that value. This can be useful for identifying the location of the highest value in a dataset or for further analysis and manipulation. To use the `idxmax()` function, one must import the Pandas library and apply it to the desired column or row in the dataframe.

Use `idxmax()` Function in Pandas (With Examples)

You can use the `pandas.DataFrame.idxmax()` function to return the index of the maximum value across a specified axis in a pandas DataFrame.

This function uses the following syntax:

```
DataFrame.idxmax(axis=0, skipna=True)
```

where:

axis: The axis to use (0 = rows, 1 = columns). Default is 0.
skipna: Whether or not to exclude NA or null values. Default is True.

The following examples show how to use this function in practice with the following pandas DataFrame:

```
import pandas as pd
```

#create DataFrame

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

#view DataFrame

df

points assists rebounds

Andy 25 5 11

Bob 12 7 8

Chad 15 7 11

Dan 8 9 6

Eric 9 12 6

Frank 23 9 5

Example 1: Find Index that has Max Value for Each Column

The following code shows how to find the index that has the maximum value for each column:

#find index that has max value for each column

```
df.idxmax(axis=0)
```

points Andy

assists Eric
rebounds Andy
dtype: object

From the output we can see:

The player with the highest value in the points column is Andy. The player with the highest value in the assists column is Eric. The player with the highest value in the rebounds column is Andy.

It's important to note that the idxmax() function will return the first occurrence of the maximum value.

For example, notice that Andy and Chad both had 11 rebounds. Since Andy appears first in the DataFrame his name is returned.

Example 2: Find Column that has Max Value for Each Row

The following code shows how to find the column that has the maximum value for each row:

```
#find column that has max value for each row  
df.idxmax(axis=1)
```

Andy points

Bob points

Chad points

Dan assists

Eric assists

Frank points

dtype: object

The highest value in the row labelled "Andy" can be found in the points column. The highest value in the row labelled "Bob" can be found in the points column. The highest value in the row labelled "Chad" can be found in the points column. The highest value in the row labelled "Dan" can be found in the assists column. The highest value in the row labelled "Eric" can be found in the assists column. The highest value in the row labelled "Frank" can be found in the points column.

Refer to the for a complete explanation of the idxmax() function.