How can I create a pivot table in Pandas that displays percentages instead of counts?

Authored by

stats writer

June 25, 2024

A pivot table in Pandas is a data analysis tool that allows for summarizing and grouping data in a structured format. By default, the values in a pivot table are displayed as counts. However, it is possible to display percentages instead of counts by using the "aggfunc" parameter and setting it to "np.mean" or "np.sum". This will calculate the percentage of each value in relation to the total number of values in the pivot table. This can be useful in understanding the proportion of each value in the data set and identifying any trends or patterns.

Pandas: Create Pivot Table with Percentages

You can use the following syntax to add a column to a pivot table in pandas that shows the percentage of the total for a specific column:

my_table = (my_table/my_table.sum())*100

This particular syntax adds a new column called % points to a pivot table called my_table that displays the percentage of total values in the points column.

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

Example: Create Pandas Pivot Table With Percentages

Suppose we have the following pandas DataFrame that shows the number of points scored by various basketball players:

```
import pandas as pd

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

#view DataFrame
print(df)

team position points
0 A Guard 22
1 A Guard 30
2 A Forward 14
3 A Forward 15
4 B Guard 19
5 B Guard 30
6 B Forward 23
7 B Forward 20
```


We can use the pivot_table() function to create a pivot table that shows the sum of points by team and position:

```
#create pivot table to calculate sum of points by team
```

and position
my_table = pd.pivot_table(df, index=, aggfunc='sum')

#view pivot table
print(my_table)

points
team position
A Forward 29
Guard 52
B Forward 43
Guard 49

From the output we can see:

Forwards on team A scored a total of 29 points.Guards on team A scored a total of 52 points.Forwards on team B scored a total of 43 points.Guards on team B scored a total of 49 points.

We can then use the following syntax to add a new column called % points that displays the percentage of the total points for each row:

#add column that displays points as a percentage of

total points
my_table = (my_table/my_table.sum())*100

#view updated pivot table
print(my_table)

points % points
team position
A Forward 29 16.763006
Guard 52 30.057803
B Forward 43 24.855491
Guard 49 28.323699

The new % points column now displays the points values as a percentage of total points.

Also note that you can use the round() function to round the percent values to a certain number of decimal places.

#add column that displays points as a percentage of total points (rounded)
my_table = round((my_table/my_table.sum())*100, 2)#view updated pivot table
print(my_table)

points % points

team position

A Forward 29 16.76

Guard 52 30.06

B Forward 43 24.86

Guard 49 28.32

The percentage values are now rounded to two decimal places.

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