How can I represent the output of value_counts in Pandas as a percentage?

Authored by

stats writer

June 25, 2024

The value_counts function in Pandas allows for the counting and display of the unique values in a given column of a dataset. To represent this output as a percentage, the total count of all values must be calculated and then each individual count can be divided by the total and multiplied by 100 to get the percentage. This will provide a better understanding of the distribution of values in the dataset.

Pandas: Represent value_counts as Percentage

You can use the value_counts() function in pandas to count the occurrences of values in a given column of a DataFrame.

To represent the values as percentages, you can use one of the following methods:

Method 1: Represent Value Counts as Percentages (Formatted as Decimals)

df.my_col.value_counts(normalize=True)

Method 2: Represent Value Counts as Percentages (Formatted with Percent Symbols)

df.my_col.value_counts(normalize=True).mul(100).round(1).astype(str) + '%'

Method 3: Represent Value Counts as Percentages

(Along with Counts)

counts = df.my_col.value_counts()percs = df.my_col.value_counts(normalize=True)pd.concat(, axis=1, keys=)

The following examples show how to use each method in practice with the following pandas DataFrame:

import pandas as pd

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

#view DataFrame
print(df)

team points
0 A 15
1 A 12
2 B 18
3 B 20
4 B 22
5 B 28

6 B 35

7 C 40

Example 1: Represent Value Counts as Percentages (Formatted as Decimals)

The following code shows how to count the occurrence of each value in the team column and represent the occurrences as a percentage of the total, formatted as a decimal:

```
#count occurrence of each value in 'team' column as percentage of total
df.team.value_counts(normalize=True)
```

```
B 0.625
A 0.250
C 0.125
Name: team, dtype: float64
```

From the output we can see:

The value B represents 62.5% of the occurrences in the team column.The value A represents 25% of the occurrences in the team column.The value C represents

12.5% of the occurrences in the team column.

Notice that the percentages are formatted as decimals.

Example 2: Represent Value Counts as Percentages (Formatted with Percent Symbols)

```
#count occurrence of each value in 'team' column as percentage of total
df.team.value_counts(normalize=True).mul(100).round(1).astype(str) + '%'
```

```
B 62.5%
A 25.0%
C 12.5%
Name: team, dtype: object
```

Notice that the percentages are formatted as strings with percent symbols.

Example 3: Represent Value Counts as Percentages (Along with Counts)

The following code shows how to count the occurrence of each value in the team column and represent the occurrences as both counts and percentages:

```
#count occurrence of each value in 'team' column
counts = df.team.value_counts()

#count occurrence of each value in 'team' column as percentage of total
percs = df.team.value_counts(normalize=True)

#concatenate results into one DataFrame
pd.concat(, axis=1, keys=)
```

```
count percentage
B 5 0.625
A 2 0.250
C 1 0.125
```

Notice that the count column displays the count of each unique value in the team column while the percentage column displays each unique value as a percentage of the total occurrences.

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