How can I calculate the sum of columns in Pandas?

Authored by

stats writer

April 17, 2024

Calculating the sum of columns in Pandas is a simple and efficient way to obtain the total values of each column in a given dataset. This can be achieved by using the built-in "sum()" function in Pandas, which allows for quick and accurate calculation of the sum of all values in a selected column. By specifying the desired column within the parentheses, the function will automatically add up all the values in that column and return the sum as a single value. This feature in Pandas is particularly useful for data analysis and manipulation, as it provides a convenient way to obtain important numerical information from a dataset.

Calculate the Sum of Columns in Pandas

Often you may be interested in calculating the sum of one or more columns in a pandas DataFrame. Fortunately you can do this easily in pandas using the function.

This tutorial shows several examples of how to use this function.

Example 1: Find the Sum of a Single Column

Suppose we have the following pandas DataFrame:

```
import pandas as pd
import numpy as np

#create DataFrame
df = pd.DataFrame({'rating': ,
'points': ,
'assists': ,
```

'rebounds': })

#view DataFrame

df

rating points assists rebounds

0 90 25 5 NaN

1 85 20 7 8

2 82 14 7 10

3 88 16 8 6

4 94 27 5 6

5 90 20 7 9

6 76 12 6 6

7 75 15 9 10

8 87 14 9 10

9 86 19 5 7

We can find the sum of the column titled "points" by using the following syntax:

df.sum()

182

The sum() function will also exclude NA's by default.

For example, if we find the sum of the "rebounds" column, the first value of "NaN" will simply be excluded from the calculation:

```
df.sum()

72.0
```

Example 2: Find the Sum of Multiple Columns

We can find the sum of multiple columns by using the following syntax:

```
#find sum of points and rebounds columns
df].sum()

rebounds 72.0
points 182.0
dtype: float64
```

Example 3: Find the Sum of All Columns

We can find also find the sum of all columns by using the following syntax:

```
#find sum of all columns in DataFrame
```

df.sum()

rating 853.0
points 182.0
assists 68.0
rebounds 72.0
dtype: float64

For columns that are not numeric, the sum() function will simply not calculate the sum of those columns.

You can find the complete documentation for the sum() function .