

How can I sum specific columns in Pandas?

Authored by
stats writer

July 1, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I sum specific columns in Pandas?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=165327>

Pandas is a popular open-source library in Python used for data analysis and manipulation. It offers a variety of functions to easily sum specific columns in a dataset. These functions allow users to calculate the sum of numerical values within a specific column or across multiple columns. For example, the "sum()" function can be used to sum all the values in a single column, while the "agg()" function can be used to sum multiple columns at once by specifying the desired columns as parameters. Other useful functions include "groupby()" which can sum columns based on a certain category or criteria, and "apply()" which enables custom summing operations. These functions make it convenient for users to efficiently calculate and analyze data in a Pandas DataFrame.

Sum Specific Columns in Pandas (With Examples)

You can use the following methods to find the sum of a specific set of columns in a pandas DataFrame:

Method 1: Find Sum of All Columns

```
#find sum of all columns
```

```
df = df.sum(axis=1)
```

Method 2: Find Sum of Specific Columns

```
#specify the columns to sum
```

```
cols =
```

```
#find sum of columns specified
```

```
df = df.sum(axis=1)
```

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({'points': ,
'assists': ,
'rebounds': })
```

```
#view DataFrame
print(df)
```

```
points assists rebounds
0 18 5 11
1 22 7 8
2 19 7 10
3 14 9 6
4 14 12 6
5 11 9 5
6 20 9 9
7 28 4 12
```

Example 1: Find Sum of All Columns

The following code shows how to sum the values of the rows across all columns in the DataFrame:

```
#define new column that contains sum of all columns
```

```
df = df.sum(axis=1)
```

```
#view updated DataFrame
```

```
df
```

```
points assists rebounds sum_stats
```

```
0 18 5 11 34
```

```
1 22 7 8 37
```

```
2 19 7 10 36
```

```
3 14 9 6 29
```

```
4 14 12 6 32
```

```
5 11 9 5 25
```

```
6 20 9 9 38
```

```
7 28 4 12 44
```

The `sum_stats` column contains the sum of the row values across all columns.

For example, here's how the values were calculated:

Sum of row 0: $18 + 5 + 11 = 34$

Sum of row 1: $22 + 7 + 8 = 37$

Sum of row 2: $19 + 7 + 10 = 36$

And so on.

Example 2: Find Sum of Specific Columns

The following code shows how to sum the values of the rows across all columns in the DataFrame:

```
#specify the columns to sum
```

```
cols =
```

```
#define new column that contains sum of specific columns
```

```
df = df.sum(axis=1)
```

```
#view updated DataFrame
```

```
df
```

```
points assists rebounds sum_stats
```

```
0 18 5 11 23
```

```
1 22 7 8 29
```

```
2 19 7 10 26
```

```
3 14 9 6 23
```

```
4 14 12 6 26
```

```
5 11 9 5 20
```

```
6 20 9 9 29
```

```
7 28 4 12 32
```

For example, here's how the values were calculated:

Sum of row 0: $18 + 5 + 11 = 23$

Sum of row 1: $22 + 7 = 29$

Sum of row 2: $19 + 7 = 26$

And so on.

Additional Resources

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

ARABPSYCHOLOGY.COM