

How can I subtract two columns in a Pandas DataFrame?

Authored by
stats writer

May 3, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I subtract two columns in a Pandas DataFrame?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=142409>

The process of subtracting two columns in a Pandas DataFrame involves using the built-in mathematical functions and methods provided by the Pandas library. This allows for efficient and accurate manipulation of data within a DataFrame. By selecting the two columns to be subtracted and using the appropriate function, the resulting column will contain the difference between the two original columns. This can be useful for performing various data analysis and calculations within a DataFrame.

Subtract Two Columns in Pandas DataFrame

You can use the following syntax to subtract one column from another in a pandas DataFrame:

```
#subtract column 'B' from column 'A'  
df = df.A- df.B
```

The following examples show how to use this syntax in practice.

Example 1: Subtract Two Columns in Pandas

The following code shows how to subtract one column from another in a pandas DataFrame and assign the result to a new column:

```
import pandas as pd  
  
#create DataFrame  
df = pd.DataFrame({'A': ,
```

```
'B': ,  
'C': })
```

#subtract column B from column A

```
df = df.A - df.B#view DataFrame
```

```
df
```

```
A B C A-B  
0 25 5 11 20  
1 12 7 8 5  
2 15 8 10 7  
3 14 9 6 5  
4 19 12 6 7  
5 23 9 5 14  
6 25 12 9 13  
7 29 4 12 25
```

The new column called 'A-B' displays the results of subtracting the values in column B from the values in column A.

Example 2: Subtract Two Columns with Missing Values

If we subtract one column from another in a pandas DataFrame and there happen to be missing values in

one of the columns, the result of the subtraction will always be a missing value:

```
import pandas as pd
```

```
import numpy as np
```

```
#create DataFrame with some missing valuesdf =  
pd.DataFrame({'A': ,
```

```
'B': ,
```

```
'C': })
```

```
#subtract column B from column A
```

```
df = df.A - df.B
```

```
#view DataFrame
```

```
df
```

```
A B C A-B
```

```
0 25 5.0 NaN 20.0
```

```
1 12 7.0 8.0 5.0
```

```
2 15 NaN 10.0 NaN
```

```
3 14 9.0 6.0 5.0
```

```
4 19 12.0 6.0 7.0
```

```
5 23 NaN 5.0 NaN
```

```
6 25 12.0 9.0 13.0
```

```
7 29 4.0 12.0 25.0
```

If you'd like, you can replace all of the missing values in the DataFrame with zeros using the `df.fillna(0)` function before subtracting one column from another:

```
import pandas as pd
import numpy as np

#create DataFrame with some missing valuesdf =
pd.DataFrame({'A': ,
'B': ,
'C': })
#replace all missing values with zeros
df = df.fillna(0)

#subtract column B from column A
df = df.A - df.B#view DataFrame
df
A B C A-B
0 25 5.0 0.0 20.0
1 12 7.0 8.0 5.0
2 15 0.0 10.0 15.0
3 14 9.0 6.0 5.0
4 19 12.0 6.0 7.0
5 23 0.0 5.0 23.0
6 25 12.0 9.0 13.0
```

7 29 4.0 12.0 25.0

ARABPSYCHOLOGY.COM