

How can I remove whitespace from columns in Pandas?

Authored by
stats writer

June 27, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I remove whitespace from columns in Pandas?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=154583>

Removing whitespace from columns in Pandas refers to the process of eliminating any spaces, tabs, or other blank characters in the column names or values within a Pandas DataFrame. This can be achieved by using built-in functions such as `strip()`, `replace()`, or `apply()`, which allow for the removal of leading, trailing, or all whitespace within a column. Removing whitespace is useful for data cleaning and manipulation, as it ensures uniformity in column names and values, making it easier to perform operations and analysis on the data. It is a simple yet crucial step in data preprocessing when working with Pandas.

Pandas: Strip Whitespace from Columns

You can use the following methods to strip whitespace from columns in a pandas DataFrame:

Method 1: Strip Whitespace from One Column

```
df = df.str.strip()
```

Method 2: Strip Whitespace from All String Columns

```
df = df.apply(lambda x: x.str.strip() if x.dtype == 'object' else x)
```

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': ,  
'position': ,  
'points': })
```

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

```
team position points  
0 Mavs Point Guard 11  
1 Heat Small Forward 8  
2 Nets Center 10  
3 Cavs Power Forward 6  
4 Hawks Point Guard 22  
5 Jazz Center 29
```

Example 1: Strip Whitespace from One Column

The following code shows how to strip whitespace from every string in the position column:

```
#strip whitespace from position column  
df = df.str.strip()  
  
#view updated DataFrame  
print(df)
```

```
team position points
0 Mavs Point Guard 11
1 Heat Small Forward 8
2 Nets Center 10
3 Cavs Power Forward 6
4 Hawks Point Guard 22
5 Jazz Center 29
```

Notice that all whitespace has been stripped from each string that had whitespace in the position column.

Example 2: Strip Whitespace from All String Columns

The following code shows how to strip whitespace from each string in all string columns of the DataFrame:

```
#strip whitespace from all string columns
df = df.apply(lambda x: x.str.strip() if x.dtype == 'object'
else x)

#view updated DataFrame
print(df)
```

```
team position points
0 Mavs Point Guard 11
1 Heat Small Forward 8
```

2 Nets Center 10

3 Cavs Power Forward 6

4 Hawks Point Guard 22

5 Jazz Center 29

Notice that all whitespace has been stripped from both the team and position columns, which are the two string columns in the DataFrame.

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