

How can I split a string column in Pandas into multiple columns?

Authored by
stats writer

May 3, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I split a string column in Pandas into multiple columns?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=142415>

The process of splitting a string column in Pandas into multiple columns involves using the "str.split()" function to divide the string values in the column based on a provided delimiter. This creates a new column for each segment of the split string, effectively separating the original column into multiple columns. This is useful for organizing and manipulating data within a Pandas DataFrame.

Split String Column in Pandas into Multiple Columns

You can use the following basic syntax to split a string column in a pandas DataFrame into multiple columns:

```
#split column A into two columns: column A and  
column B  
df] = df.str.split(',', 1, expand=True)
```

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

Example 1: Split Column by Comma

The following code shows how to split a column in a pandas DataFrame, based on a comma, into two separate columns:

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

```
'points': })
```

```
#view DataFrame
```

```
df
```

```
team points
```

```
0 Mavs, West 112
```

```
1 Spurs, West 104
```

```
2 Nets, East 127
```

```
#split team column into two columns
```

```
df] = df.str.split(',', 1, expand=True)
```

```
#view updated DataFrame
```

```
df
```

```
team points conference
```

```
0 Mavs 112 West
```

```
1 Spurs 104 West
```

```
2 Nets 127 East
```

Note that you can also reorder the columns after performing the split if you'd like:

```
#reorder columns
```

```
df = df]
```

```
#view DataFrame
```

```
df
```

```
team conference points
```

```
0 Mavs West 112
```

```
1 Spurs West 104
```

```
2 Nets East 127
```

```
Example 2: Split Column by Other Delimiters
```

We can use the same syntax to split a column by other delimiters.

For example, we can split a column by a space:

```
import pandas as pd
```

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

```
#split team column into two columns
```

```
df] = df.str.split(' ', 1, expand=True)
```

```
#view updated DataFrame
```

```
df
```

team conference points

0 Mavs West 112

1 Spurs West 104

2 Nets East 127

We can also split a column by a slash:

```
import pandas as pd
```

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

```
#split team column into two columns
```

```
df] = df.str.split('/', 1, expand=True)
```

```
#view updated DataFrame
```

```
df
```

team conference points

0 Mavs West 112

1 Spurs West 104

2 Nets East 127

Using this syntax, we can split a column by any delimiter we'd like.