How can I use Pandas to add a string to each value in a specific column of my dataset?

Authored by

stats writer

June 26, 2024

Pandas is a popular library in Python used for data manipulation and analysis. One useful function in Pandas is the ability to add a string to each value in a specific column of a dataset. This can be achieved by using the "apply" function along with a lambda function to iterate through each value in the column and add the desired string. This process can be helpful for tasks such as formatting or labeling data in a dataset. With the flexibility and efficiency of Pandas, adding a string to a specific column in a dataset can be easily accomplished.

Pandas: Add String to Each Value in Column

You can use the following methods to add a string to each value in a column of a pandas DataFrame:

Method 1: Add String to Each Value in Column

df = 'some_string' + df.astype(str)

Method 2: Add String to Each Value in Column Based on Condition

#define condition
mask = (df == 'A')

#add string to values in column equal to 'A'
df.loc = 'some_string' + df.astype(str)

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': ,
'points': ,
'assists': ,
'rebounds': })

#view DataFrame
print(df)

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

Example 1: Add String to Each Value in Column

The following code shows how to add the string 'team_' to each value in the team column:

```
#add string 'team_' to each value in team column
df = 'team_' + df.astype(str)

#view updated DataFrame
print(df)

team points assists rebounds
0 team_A 18 5 11
1 team_B 22 7 8
2 team_C 19 7 10
3 team_D 14 9 6
4 team_E 14 12 6
5 team_F 11 9 5
6 team_G 20 9 9
7 team_H 28 4 12
```

Notice that the prefix 'team_' has been added to each value in the team column.

You can also use the following syntax to instead add '_team' as a suffix to each value in the team column:

```
#add suffix 'team_' to each value in team column
df = df.astype(str) + '_team'
```

```
#view updated DataFrame
print(df)

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

Example 2: Add String to Each Value in Column Based on Condition

The following code shows how to add the prefix 'team_' to each value in the team column where the value is equal to 'A':

```
#define condition
mask = (df == 'A')

#add string 'team_' to values that meet the condition
df.loc = 'team_' + df.astype(str)
```

#view updated DataFrame
print(df)

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

Notice that the prefix 'team_' has only been added to the values in the team column whose value was equal to 'A'.

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