How can I filter rows in Pandas based on the length of a string?

Authored by

stats writer

June 27, 2024

Pandas is a popular data analysis library in Python that offers various functionalities for manipulating data. One of its useful features is the ability to filter rows in a dataframe based on the length of a string. This can be achieved by using the "str.len()" function, which calculates the length of each string in a specified column. By setting a condition on the length of the string, we can select only the rows that meet the desired criteria. This allows for efficient data filtering and manipulation, making it a valuable tool for data analysts and researchers.

Pandas: Filter Rows Based on String Length

You can use the following methods to filter for rows that contain a string with a specific length in a pandas DataFrame:

Method 1: Filter Rows Based on String Length in One Column

#filter rows where col1 has a string length of 5
df.loc.str.len() == 5]

Method 2: Filter Rows Based on String Length of Multiple Columns

#filter rows where col1 has string length of 5 and col2 has string length of 7
df.loc.str.len() == 5) & (df.str.len() == 7)]

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({'conf': ,
'pos': ,
'points': })

#view DataFrame
print(df)
```

```
conf pos points
0 East Guard 5
1 East Guard 7
2 North Forward 7
3 West Center 9
4 North Center 12
5 South Forward 9
```


Example 1: Filter Rows Based on String Length in One Column

The following code shows how to filter for rows in the DataFrame that have a string length of 5 in the conf column:

#filter rows where conf has a string length of 5
df.loc.str.len() == 5]

conf pos points
2 North Forward 7
4 North Center 12
5 South Forward 9

Only the rows where the conf column has a string length of 5 are returned.

We can see that two different strings met this criteria in the conf column:

"North""South"

Both strings have a length of 5.

Example 2: Filter Rows Based on String Length of Multiple Columns

The following code shows how to filter for rows in the DataFrame that have a string length of 5 in the conf column and a string length of 7 in the pos column:

#filter rows where conf has string length of 5 and pos has string length of 7

df.loc.str.len() == 5) & (df.str.len() == 7)]

conf pos points
2 North Forward 7
5 South Forward 9

Note: You can find the complete documentation for the str.len() function in pandas .

Additional Resources

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