

How can I use the LIKE operator inside a query() function in Pandas?

Authored by
stats writer

June 27, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I use the LIKE operator inside a query() function in Pandas?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=154519>

The LIKE operator is a powerful tool in SQL for pattern matching within strings. In Pandas, the query() function allows for the execution of SQL-like queries on a DataFrame object. By using the LIKE operator within the query() function, users can easily filter their data based on specific patterns within string values. This can be useful for performing tasks such as searching for specific words or characters within a column or identifying data that follows a certain format. Overall, the combination of the LIKE operator and the query() function in Pandas provides a flexible and efficient way to manipulate and analyze data.

Pandas: Use LIKE inside query()

You can use the following methods to use LIKE (similar to SQL) inside a pandas query() function to find rows that contain a particular pattern:

Method 1: Find Rows that Contain One Pattern

```
df.query('my_column.str.contains("pattern1")')
```

Method 2: Find Rows that Contain One of Several Patterns

```
df.query('my_column.str.contains("pattern1|pattern2")')
```

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': ,  
'rebounds': })
```

#view DataFrame

```
print(df)
```

```
team points rebounds
```

```
0 Cavs 3 15
```

```
1 Heat 3 14
```

```
2 Mavs 4 14
```

```
3 Mavs 5 10
```

```
4 Nets 4 8
```

```
5 Heat 7 14
```

```
6 Cavs 8 13
```

```
7 Jazz 7 9
```

```
8 Jazz 12 5
```

```
9 Hawks 14 4
```

Example 1: Find Rows that Contain One Pattern

The following code shows how to use the query() function to find all rows in the DataFrame that contain "avs" in the team column:

```
df.query('team.str.contains("avs")')
```

```
team points rebounds
```

```
0 Cavs 3 15
```

```
2 Mavs 4 14
```

```
3 Mavs 5 10
```

```
6 Cavs 8 13
```

Each row that is returned contains "avs" somewhere in the team column.

Also note that this syntax is case-sensitive.

Thus, if we used "AVS" instead then we would not receive any results because no row contains uppercase "AVS" in the team column.

Example 2: Find Rows that Contain One of Several Patterns

The following code shows how to use the query() function to find all rows in the DataFrame that contain "avs" or "eat" in the team column:

```
df.query('team.str.contains("avs|eat")')
```

```
team points rebounds
```

0 Cavs 3 15

1 Heat 3 14

2 Mavs 4 14

3 Mavs 5 10

5 Heat 7 14

6 Cavs 8 13

Each row that is returned contains either "avs" or "eat" somewhere in the team column.

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