Can a variable be used in the query() function in Pandas?

Authored by

stats writer

June 24, 2024

The query() function in Pandas is a powerful tool for filtering and selecting data from a dataframe. It allows for complex filtering conditions to be applied to the data. However, one question that may arise is whether a variable can be used in the query() function. The answer is yes, a variable can be used in the query() function as long as it is properly defined and formatted. This allows for more flexibility and dynamic filtering options in data analysis using Pandas.

Pandas: Use Variable in query() Function

You can use the following syntax to use the query() function in pandas and reference a variable name:

df.query('team == @team_name')

This particular query searches for rows in a pandas DataFrame where the team column is equal to the value saved in the variable called team_name.

The following example shows how to use this syntax in practice.

Example: How to Use Variable in Pandas Query

Suppose we have the following pandas DataFrame that contains information about various basketball players:

import pandas as pd

#create DataFrame

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

#view DataFrame
print(df)
```

```
team position points
0 A G 22
1 A G 25
2 A F 24
3 B G 39
4 B F 34
5 B F 20
6 B F 18
7 C G 17
8 C G 20
9 C F 19
10 C F 22
```


Now suppose that we would like to query for the rows where the value in the team column is equal to C.

We can use the following syntax to create a variable called team_name that is equal to 'C' and then reference

that variable in the query() function:

```
#specify team name to search for
team_name = 'C'

#query for rows where team is equal to team_name
df.query('team == @team_name')

team position points
7 C G 17
8 C G 20
9 C F 19
10 C F 22
```

Notice that the query() function returns all rows where the value in the team column is equal to C.

Also note that we can reference multiple variables in the query() function if we'd like.

For example, the following code shows how to use the query() function to return all rows where the value in the team column is equal to the value of a variable called team_A *or* a variable called team_C:

```
#create two variables
```

```
team_A = 'A'
team_C = 'C'

#query for rows where team is equal to either of the two variables
df.query('team == @team_A | team == @team_C')
```

team position points
0 A G 22
1 A G 25
2 A F 24
7 C G 17
8 C G 20
9 C F 19
10 C F 22

The query returns all of the rows in the DataFrame where team is equal to the values stored in one of the two variables that we specified.

Note: You can find the complete documentation for the pandas query() function .