

How can I filter a Pandas DataFrame by the values in a specific column?

Authored by
stats writer

April 26, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I filter a Pandas DataFrame by the values in a specific column?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=139842>

Filtering a Pandas DataFrame by the values in a specific column is a process of selecting and displaying only the rows of data that meet certain criteria in the chosen column. This can be achieved by using the "filter" function in Pandas, which allows the user to specify the column and the desired values to be filtered. This process is useful for data analysis and manipulation, as it allows for a more focused and targeted examination of specific data points within a larger dataset. By filtering a DataFrame, the user can gain valuable insights and make informed decisions based on the desired values in the chosen column.

Filter a Pandas DataFrame by Column Values

The simplest way to filter a pandas DataFrame by column values is to use the function.

This tutorial provides several examples of how to use this function in practice with the following pandas DataFrame:

```
import pandas as pd

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

#view DataFrame
df
```

team points assists rebounds

0 A 25 5 11

1 A 12 7 8

2 B 15 7 10

3 B 14 9 6

4 C 19 12 6

Example 1: Filter Based on One Column

The following code shows how to filter the rows of the DataFrame based on a single value in the "points" column:

```
df.query('points == 15')
```

team points assists rebounds

2 B 15 7 10

Example 2: Filter Based on Multiple Columns

The following code shows how to filter the rows of the DataFrame based on several values in different columns:

```
#return rows where points is equal to 15 or 14
```

```
df.query('points == 15 | points == 14')
```

team points assists rebounds

2 B 15 7 10

3 B 14 9 6

**#return rows where points is greater than 13 and
rebounds is greater than 6**

df.query('points > 13 & points > 6')

team points assists rebounds

0 A 25 5 11

2 B 15 7 10

Example 3: Filter Based on Values in a List

**The following code shows how to filter the rows of the
DataFrame based on values in a list**

#define list of values

**value_list = #return rows where points is in the list of
values**

df.query('points in @value_list')

team points assists rebounds

0 A 25 5 11

1 A 12 7 8

4 C 19 12 6

#return rows where points is *not* in the list of values
df.query('points not in @value_list')

team points assists rebounds

2 B 15 7 10

3 B 14 9 6

ARABPSYCHOLOGY.COM