How can I filter a Pandas DataFrame based on a specific index value?

Authored by

stats writer

June 27, 2024

ARABPSYCHOLOGY.COM

RECOMMENDED CITATION

stats writer (2024). *How can I filter a Pandas DataFrame based on a specific index value?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=155641

To filter a Pandas DataFrame based on a specific index value, one can use the .loc function which allows for label-based indexing. This function takes in the specific index value as a parameter and returns a new DataFrame with only the rows that match the given index value. This is a useful tool for selecting and manipulating specific data within a DataFrame, making it easier to analyze and work with the data.

Pandas: Filter by Index Value

You can use the following basic syntax to filter the rows of a pandas DataFrame based on index values:

df_filtered = df

This will filter the pandas DataFrame to only include the rows whose index values are contained in some_list.

The following examples show how to use this syntax in practice.

Example 1: Filter by Numeric Index Values

Suppose we have the following pandas DataFrame:

import pandas as pd

#create DataFrame
df = pd.DataFrame({'points': ,
'assists': ,

ARABPSYCHOLOGY.COM

'rebounds': })

#view DataFrame
print(df)

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

Notice that the index values are numeric.

Suppose we would like to filter for rows where the index value is equal to 1, 5, 6, or 7.

We can use the following syntax to do so:

#define list of index values
some_list =

#filter for rows in list

ARABPSYCHOLOGY.COM

```
df_filtered = df

#view filtered DataFrame
print(df_filtered)

points assists rebounds
1 22 7 8
5 11 9 5
6 20 9 9
7 28 4 12
```

Notice that the only rows returned are those whose index value is equal to 1, 5, 6, or 7.

Example 2: Filter by Non-Numeric Index Values

Suppose we have the following pandas DataFrame:

```
import pandas as pd

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

ARABPSYCHOLOGY.COM

#view DataFrame
print(df)

points assists rebounds
A 18 5 11
B 22 7 8
C 19 7 10
D 14 9 6
E 14 12 6
F 11 9 5
G 20 9 9
H 28 4 12

Notice that the index values are character values.

Suppose we would like to filter for rows where the index value is equal to A, C, F, or G.

#define list of index values
some_list =

#filter for rows in list
df_filtered = df

#view filtered DataFrame

ARABPSYCHOLOGY.COM

print(df_filtered)

points assists rebounds

A 18 5 11

C 19 7 10

F 11 9 5

G 20 9 9

Notice that the only rows returned are those whose index value is equal to A, C, F, or G.

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

ARABPSYCHOLOGY.COM