

How can I use the first column of a Pandas dataframe as the index?

Authored by
stats writer

June 27, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I use the first column of a Pandas dataframe as the index?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=154670>

Using the first column of a Pandas dataframe as the index is a simple and efficient way to organize and access data. This feature allows you to use the values in the first column of your dataframe as the row labels, making it easier to locate and manipulate specific rows of data. By setting the index, you can easily filter, sort, and perform operations on your data based on the values in the first column. This method is particularly useful when working with large datasets, as it can improve the speed and accuracy of data analysis. To use the first column as the index, simply specify the column name when creating or loading the dataframe, or use the "set_index()" function to change the index after the dataframe has been created.

Pandas: Use First Column as Index

You can use the following methods to use the first column as the index column in a pandas DataFrame:

Method 1: Use First Column as Index When Importing DataFrame

```
df = pd.read_csv('my_data.csv', index_col=0)
```

Method 2: Use First Column as Index with Existing DataFrame

```
df = df.set_index()
```

The following examples show how to use each method in practice.

Example 1: Use First Column as Index When Importing DataFrame

Suppose we have the following CSV file called my_data.csv:

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

If we import the CSV file without specifying an index column, pandas will simply create an index column with numerical values starting at 0:

```
#import CSV file without specifying index column  
df = pd.read_csv('my_data.csv')
```

```
#view DataFrameprint(df)
```

```
team points assists
```

```
0 A 18 5
1 B 22 7
2 C 19 7
3 D 14 9
4 E 14 12
5 F 11 9
6 G 20 9
7 H 28 4
```

However, we can use the `index_col` argument to specify that the first column in the CSV file should be used as the index column:

```
#import CSV file and specify index column
df = pd.read_csv('my_data.csv', index_col=0)
```

```
#view DataFrameprint(df)
```

```
points assists
```

```
team
```

```
A 18 5
```

```
B 22 7
```

```
C 19 7
```

```
D 14 9
```

```
E 14 12
```

F 11 9

G 20 9

H 28 4

Notice that the team column is now used as the index column.

Example 2: Use First Column as Index with Existing DataFrame

Suppose we have the following existing pandas DataFrame:

```
import pandas as pd

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

#view DataFrame
df
team points assists
0 A 18 5
1 B 22 7
2 C 19 7
3 D 14 9
4 E 14 12
```

5 F 11 9

6 G 20 9

7 H 28 4

#set 'team' column as index column

```
df = df.set_index()
```

#view updated DataFrame

```
print(df)
```

points assists

team

A 18 5

B 22 7

C 19 7

D 14 9

E 14 12

F 11 9

G 20 9

H 28 4

Notice that the team column is now used as the index column.

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

ARABPSYCHOLOGY.COM