

How do I set a column as the index in Pandas?

Authored by
stats writer

May 3, 2024

RECOMMENDED CITATION

stats writer (2024). *How do I set a column as the index in Pandas?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=142384>

Setting a column as the index in Pandas is a simple process that allows for easier and more efficient data manipulation. To do so, the user can use the "set_index()" function and specify the desired column as the parameter. This will replace the default numerical index with the chosen column, making it the primary identifier for each row in the data. This can be particularly useful when working with large datasets or when performing specific operations that require referencing specific rows. Additionally, the index can also be reset back to the default numerical index if needed.

Pandas: Set Column as Index

You can use the following syntax to set a column in a pandas DataFrame as the index:

```
#set one column as index  
df.set_index('col1')
```

```
#set multiple columns as multi index  
df.set_index()
```

The following examples show how to use this syntax in practice with the following DataFrame:

```
import pandas as pd  
  
#create DataFrame  
df = pd.DataFrame({'points': ,  
'assists': ,  
'team': ,
```

```
'conference': })
```

```
#view DataFrame
```

```
df
```

points	assists	team	conference
0	5	11	A
1	7	8	B
2	7	10	C
3	9	6	D
4	12	6	E
5	9	5	F

Example 1: Set One Column as Index

The following code shows how to set one column of the pandas DataFrame as the index:

```
df.set_index('team')
```

```
points
```

assists	conference
5	11
7	8
7	10
9	6
12	6
9	5

F 9 5 6

Example 2: Set Multiple Columns as Index

The following code shows how to set multiple columns of the pandas DataFrame as a Multi-Index::

```
df.set_index()
```

```
points assists  
team conference
```

```
A 1 5 11
```

```
B 2 7 8
```

```
C 3 7 10
```

```
D 4 9 6
```

```
E 5 12 6
```

```
F 6 9 5
```