How can I check if multiple columns in a Pandas dataframe are equal?

Authored by

stats writer

June 26, 2024

The process of checking if multiple columns in a Pandas dataframe are equal involves comparing the values in each column and determining if they are the same. This can be done by using the "equals" function in Pandas, which returns a boolean value indicating if the columns are equal or not. This method allows for a quick and efficient way to verify the equality of multiple columns in a dataframe.

Pandas: Check if Multiple Columns are Equal

You can use the following methods to check if multiple columns are equal in pandas:

Method 1: Check if All Columns Are Equal

df = df.eq(df.iloc, axis=0).all(1)

Method 2: Check if Specific Columns Are Equal

df = df.apply(lambda x: x.col1 == x.col3 == x.col4, axis=1)

The following examples show how to use each method in practice with the following pandas DataFrame:

import pandas as pd

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

'B': ,
'C': ,
'D': })

#view DataFrame
print(df)

A B C D
0 4 4 4 4
1 0 2 0 0
2 3 3 3 3
3 3 5 3 3
4 6 6 5 3
5 8 4 10 8
6 7 7 7 7

Example 1: Check if All Columns Are Equal

We can use the following syntax to check if the value in every column in the DataFrame is equal for each row:

#create new column that checks if all columns match in each row
df = df.eq(df.iloc, axis=0).all(1)

#view updated DataFrame

print(df)

A B C D matching
0 4 4 4 4 True
1 0 2 0 0 False
2 3 3 3 3 True
3 3 5 3 3 False
4 6 6 5 3 False
5 8 4 10 8 False
6 7 7 7 7 True

If the value in each column is equal, then the matching column returns True.

Otherwise, it returns False.

Note that you can convert True and False values to 1 and 0 by using astype(int) as follows:

#create new column that checks if all columns match in each row
df = df.eq(df.iloc, axis=0).all(1).astype(int)

#view updated DataFrame
print(df)

A B C D matching
0 4 4 4 4 1
1 0 2 0 0 0
2 3 3 3 3 1
3 3 5 3 3 0
4 6 6 5 3 0
5 8 4 10 8 0
6 7 7 7 7 1

Example 2: Check if Specific Columns Are Equal

We can use the following syntax to check if the value in columns A, C, and D in the DataFrame are equal for each row:

```
#create new column that checks if values in columns A, C, and D are equal
df = df.apply(lambda x: x.A == x.C == x.D, axis=1)

#view updated DataFrame
print(df)
```

A B C D matching
0 4 4 4 4 True
1 0 2 0 0 True

2 3 3 3 3 True

3 3 5 3 3 True

4 6 6 5 3 False

5 8 4 10 8 False

6 7 7 7 7 True

Otherwise, it returns False.

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