How can I compare three columns in Pandas, and what is an example of how to do so?

Authored by

stats writer

June 27, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I compare three columns in Pandas, and what is an example of how to do so?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=155574

ARABPSYCHOLOGY.COM

Pandas is a popular library for data manipulation and analysis in Python. One common task in data analysis is comparing multiple columns within a dataset. To compare three columns in Pandas, the "compare" function can be used. This function allows for the comparison of multiple columns and returns a boolean value indicating if the values in the columns are equal or not. An example of using this function would be comparing the sales data of three different products in a company's quarterly report. By using the "compare" function, a data analyst can easily identify any discrepancies or trends among the three products' sales. Overall, the "compare" function in Pandas provides a quick and efficient way to compare multiple columns within a dataset.

Compare Three Columns in Pandas (With Example)

You can use the following basic syntax to compare the values in three columns in pandas:

df = df.apply(lambda x: x.col1 == x.col2 == x.col3, axis = 1)

This syntax creates a new column called all_matching that returns a value of True if all of the columns have matching values, otherwise it returns False.

The following example shows how to use this syntax in practice.

Example: Compare Three Columns in Pandas

Suppose we have the following pandas DataFrame with three columns:

ARABPSYCHOLOGY.COM

```
import pandas as pd

#create DataFrame
df = pd.DataFrame({'A': ,
'B': ,
'C': })

#view DataFrame
print(df)
```

A B C
0 4 4 4
1 0 2 0
2 3 3 3
3 3 5 5
4 6 6 5
5 8 4 10
6 7 7 7
7 9 7 9
8 12 12 12

We can use the following code to create a new column called all_matching that returns True if all three columns match in a given row and False if they do not:

ARABPSYCHOLOGY.COM

#create new column that displays whether or not all column values match
df = df.apply(lambda x: x.A == x.B == x.C, axis = 1)

#view updated DataFrame
print(df)

A B C all_matching
0 4 4 4 True
1 0 2 0 False
2 3 3 3 True
3 3 5 5 False
4 6 6 5 False
5 8 4 10 False
6 7 7 7 True
7 9 7 9 False
8 12 12 12 True

The new column called all_matching shows whether or not the values in all three columns match in a given row.

For example:

All three values match in the first row, so True is

ARABPSYCHOLOGY.COM

returned.Not every value matches in the second row, so False is returned.

And so on.

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

ARABPSYCHOLOGY.COM