How can I merge Pandas DataFrames using multiple columns?

Authored by

stats writer

April 18, 2024

To merge Pandas DataFrames using multiple columns, one can use the merge function in Pandas library. This function allows for the combination of two or more DataFrames based on one or more common columns. By specifying the columns to merge on, the function will join the DataFrames and create a new DataFrame with the combined data. This method is useful for combining datasets that have related information in multiple columns, providing a comprehensive view of the data.

Merge Pandas DataFrames on Multiple Columns

Often you may want to merge two pandas DataFrames on multiple columns. Fortunately this is easy to do using the pandas merge() function, which uses the following syntax:

```
pd.merge(df1, df2, left_on=, right_on = )
```

This tutorial explains how to use this function in practice.

Example 1: Merge on Multiple Columns with Different Names

Suppose we have the following two pandas DataFrames:

```
import pandas as pd

#create and view first DataFrame
df1 = pd.DataFrame({'a1': ,
```

```
'b': ,
'c': })

print(df1)

a1 b c
0 0 0 11
1 0 0 8
2 1 1 10
3 1 1 6
4 2 1 6

#create and view second DataFrame
df2 = pd.DataFrame({'a2': ,
'b': ,
'd': })

print(df2)

a2 b d
0 0 0 22
1 1 0 24
2 1 0 25
3 1 1 33
4 3 1 37
```

The following code shows how to perform a left join using multiple columns from both DataFrames:

pd.merge(df1, df2, how='left', left_on=, right_on =)

a1 b c a2 d
0 0 0 11 0.0 22.0
1 0 0 8 0.0 22.0
2 1 1 10 1.0 33.0
3 1 1 6 1.0 33.0
4 2 1 6 NaN NaN

Example 2: Merge on Multiple Columns with Same Names

Suppose we have the following two pandas DataFrames with the same column names:

import pandas as pd

#create DataFrames
df1 = pd.DataFrame({'a': ,
'b': ,
'c': })

df2 = pd.DataFrame({'a': ,
'b': ,

'd': })

In this case we can simplify use on = since the column names are the same in both DataFrames:

pd.merge(df1, df2, how='left', on=)

a b c d
0 0 0 11 22.0
1 0 0 8 22.0
2 1 1 10 33.0
3 1 1 6 33.0
4 2 1 6 NaN

How to Merge Two Pandas DataFrames on Index
How to Stack Multiple Pandas DataFrames