How can I update column values in one Pandas DataFrame based on values from another DataFrame?

Authored by

stats writer

June 26, 2024

The process of updating column values in a Pandas DataFrame based on values from another DataFrame involves using the merge function to combine two DataFrames based on a common column, and then using the update function to overwrite the values in one DataFrame with the values from the other DataFrame. This allows for efficient and accurate updating of data within a DataFrame using information from another DataFrame.

Pandas: Update Column Values Based on Another DataFrame

Often you may want to update the values in one column of a pandas DataFrame using values from another DataFrame.

Fortunately this is easy to do using the merge() function in pandas.

The following example shows how to do so.

Example: Update Column Values in Pandas DataFrame Based on Another DataFrame

Suppose we have the following pandas DataFrame that contains information about various basketball players:

import pandas as pd

#create DataFrame
df1 = pd.DataFrame({'team': ,

'points': ,
'assists': })

#view DataFrame
print(df1)

team points assists
0 A 18 0
1 B 22 0
2 C 19 0
3 D 14 1
4 E 14 0
5 F 11 0
6 G 20 0
7 H 28 1

Now suppose the values in the assists column are not updated in this DataFrame.

However, suppose we have the following second DataFrame that does have updated values for the assists column:

#create second DataFrame
df2 = pd.DataFrame({'team': ,

'points': ,
'assists': })

#view second DataFrame
print(df2)

team points assists
0 A 18 8
1 B 22 7
2 C 19 7
3 D 14 4
4 E 14 9
5 F 11 12
6 G 20 3
7 H 28 5

We can use the following syntax to update the values in the assists column of the first DataFrame using the values in the assists column of the second DataFrame:

#merge two DataFrames
df1 = df1.merge(df2, on='team', how='left')

#drop original DataFrame columns
df1.drop(, inplace=True, axis=1)

```
#rename columns
df1.rename(columns={'points_y':'points','assists_y':'assists'}, inplace=True)

#view updated DataFrame
print(df1)

team points assists
0 A 18 8
1 B 22 7
2 C 19 7
3 D 14 4
4 E 14 9
5 F 11 12
6 G 20 3
7 H 28 5
```

Notice that the values in the assists column of the first DataFrame have been updated using the values from the assists column in the second DataFrame.

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