How can I use the isin function in Pandas to check for multiple columns?

Authored by

stats writer

June 24, 2024

The isin function in Pandas is a useful tool for checking if values in a column or series are contained within a given list or array. This function can also be used to check for multiple columns by passing a list of column names as the argument to the isin function. By doing so, the function will return a boolean series indicating which rows have values that match any of the given columns. This allows for efficient and quick identification of specific data within a dataset. Additionally, the isin function can be combined with other Pandas functions to filter, sort, or manipulate the data further. Overall, the isin function is a valuable tool for data analysis and manipulation in Pandas.

Pandas: Use isin for Multiple Columns

You can use the following methods with the pandas isin() function to filter based on multiple columns in a pandas DataFrame:

Method 1: Filter where Multiple Columns Are Equal to Specific Values

df = df].isin().all(axis=1)]

This particular example filters the DataFrame for rows where the team column is equal to 'A' and the position column is equal to 'Guard.'

Method 2: Filter where At Least One Column is Equal to Specific Value

df = df].isin().any(axis=1)]

This particular example filters the DataFrame for rows where the team column is equal to 'A' or the position column is equal to 'Guard.'

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({'team': ,
'position': ,
'points': })

#view DataFrame
print(df)
```

```
team position points
0 A Guard 11
1 A Guard 18
2 A Forward 10
3 A Forward 22
4 B Guard 26
5 B Guard 35
6 B Forward 19
```

7 B Forward 12

Example 1: Filter where Multiple Columns Are Equal to Specific Values

We can use the following syntax to filter the DataFrame to only contain rows where the team column is equal to 'A' and the position column is equal to 'Guard.'

#filter rows where team column is 'A' and position column is 'Guard'
df = df].isin().all(axis=1)]

#view filtered DataFrameprint(df)

team position points
0 A Guard 11
1 A Guard 18

Notice that only the rows where the team column is equal to 'A' and the position column is equal to 'Guard' remain in the filtered DataFrame.

Example 2: Filter where At Least One Column is Equal to Specific Value

We can use the following syntax to filter the DataFrame to only contain rows where the team column is equal to 'A' or the position column is equal to 'Guard.'

```
#filter rows where team column is 'A' or position column is 'Guard'
df = df].isin().any(axis=1)]

#view filtered DataFrameprint(df)
```

```
team position points
0 A Guard 11
1 A Guard 18
2 A Forward 10
3 A Forward 22
4 B Guard 26
5 B Guard 35
```

Notice that only the rows where the team column is equal to 'A' or the position column is equal to 'Guard' remain in the filtered DataFrame.

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