# How can I find unique values in multiple columns in Pandas?

Authored by

**stats writer**

April 18, 2024

Pandas is a popular data analysis library in Python that offers various functions for data manipulation and exploration. One of the common tasks in data analysis is to identify unique values within a dataset. In Pandas, this can be achieved by using the "unique" function, which returns an array of unique values from a specified column. However, if the goal is to find unique values across multiple columns, additional steps are required. This can be achieved by combining the "unique" function with the "drop_duplicates" function, which removes duplicate rows from a dataset. By using these two functions together, one can easily find unique values in multiple columns in Pandas. This approach can be particularly useful in identifying patterns and outliers within a dataset, providing valuable insights for data analysis.

## Find Unique Values in Multiple Columns in Pandas

**Often you may be interested in finding all of the unique values across multiple columns in a pandas DataFrame. Fortunately this is easy to do using the pandas unique() function combined with the ravel() function:**

**unique(): Returns unique values in order of appearance.ravel(): Returns a flattened data series.**

**For example, suppose we have the following pandas DataFrame:**

```
import pandas as pd

#create DataFrame
df = pd.DataFrame({'col1': ,
'col2': ,
'col3': })
```

#view DataFrame
print(df)

col1 col2 col3
0 a a 11
1 b c 8
2 c e 10
3 d f 6
4 e g 6

### Return Array of Unique Values

The following code shows how to find the unique values in col1 and col2:

pd.unique(df].values.ravel())

array(, dtype=object)

From the output we can see that there are 7 unique values across these two columns: a, b, c, d, e, f, g.

### Return DataFrame of Unique Values

If you'd like to return these values as a DataFrame instead of an array, you can use the following code:

uniques = pd.unique(df].values.ravel())

pd.DataFrame(uniques)

0

0 a

1 b

2 c

3 e

4 d

5 f

6 g

### Return Number of Unique Values

If you simply want to know the number of unique values across multiple columns, you can use the following code:

uniques = pd.unique(df].values.ravel())

len(uniques)

7

This tell us that there are 7 unique values across these two columns.

## How to Merge Pandas DataFrames on Multiple Columns

## How to Filter a Pandas DataFrame on Multiple Conditions