How can I count the unique values in a Pandas dataframe, and what are some examples of how to do so?

Authored by

stats writer

May 11, 2024

Counting the unique values in a Pandas dataframe is a useful tool for understanding the data and identifying any patterns or discrepancies. To count the unique values, the Pandas library offers the "nunique()" function, which returns the number of distinct values in a column or the entire dataframe. This function allows for quick and efficient counting of unique values without having to manually sort and count them.

Some examples of using the "nunique()" function include counting the number of unique cities in a "City" column, the number of unique customers in a "Customer ID" column, or the number of unique products in a "Product Name" column. Additionally, this function can be applied to the entire dataframe to obtain the total number of unique values across all columns. Overall, the "nunique()" function is a valuable tool for data analysis and provides insights into the diversity and distribution of data within a Pandas dataframe.

Count Unique Values in Pandas (With Examples)

You can use the function to count the number of unique values in a pandas DataFrame.

This function uses the following basic syntax:

#count unique values in each column
df.nunique()

#count unique values in each row
df.nunique(axis=1)

The following examples show how to use this function in practice with the following pandas DataFrame:

import pandas as pd

#create DataFrame
df = pd.DataFrame({'team': ,
'points': ,
'assists': ,
'rebounds': })

#view DataFrame
df

team points assists rebounds
0 A 8 5 11
1 A 8 8 8
2 A 13 7 11
3 A 13 9 6
4 B 22 12 6
5 B 22 9 5
6 B 25 9 9
7 B 29 4 12

Example 1: Count Unique Values in Each Column

The following code shows how to count the number of unique values in each column of a DataFrame:

#count unique values in each column

```
df.nunique()

team 2
points 5
assists 5
rebounds 6
dtype: int64
```

From the output we can see:

The 'team' column has 2 unique valuesThe 'points' column has 5 unique valuesThe 'assists' column has 5 unique valuesThe 'rebounds' column has 6 unique values

Example 2: Count Unique Values in Each Row

The following code shows how to count the number of unique values in each row of a DataFrame:

```
#count unique values in each row
df.nunique(axis=1)

0 4
1 2
2 4
```

3 4

4 4

5 4

6 3

7 4

dtype: int64

From the output we can see:

The first row has 4 unique valuesThe second row has 2 unique valuesThe third row has 4 unique values

And so on.

Example 3: Count Unique Values by Group

```
#count unique 'points' values, grouped by team
df.groupby('team').nunique()
```

team

A 2

B 3

Name: points, dtype: int64

From the output we can see:

Team 'A' has 2 unique 'points' valuesTeam 'B' has 3 unique 'points' values

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