

What is the purpose and how do you use the Pandas value_counts() function? Provide examples.

Authored by
stats writer

May 5, 2024

RECOMMENDED CITATION

stats writer (2024). *What is the purpose and how do you use the Pandas value_counts() function? Provide examples..* PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=142945>

The purpose of the Pandas value_counts() function is to count the unique values in a column or series of data and present them in a descending order. This function can be used to get an overview of the distribution of values in a dataset, identify the most frequent values, and detect any missing or unexpected values.

To use the value_counts() function, first import the Pandas library and load the dataset. Then, call the function on the desired column or series, specifying any additional parameters such as sorting or including missing values. The function will return a series object with the unique values as the index and their corresponding counts as the values.

For example, if we have a dataset of customer ratings for a product and want to see the distribution of ratings, we can use the value_counts() function on the "rating" column. This will give us a series with the ratings (e.g. 1 star, 2 stars) as the index and the number of occurrences as the values. We can then use this information to visualize the distribution or further analyze the data.

Use Pandas value_counts() Function (With Examples)

You can use the value_counts() function to count the frequency of unique values in a pandas Series.

This function uses the following basic syntax:

```
my_series.value_counts()
```

The following examples show how to use this syntax in practice.

Example 1: Count Frequency of Unique Values

The following code shows how to count the occurrences of unique values in a pandas Series:

```
import pandas as pd
```

```
#create pandas Series
```

```
my_series = pd.Series()
```

```
#count occurrences of unique values in Series
```

```
my_series.value_counts()
```

```
3 4
```

```
4 2
```

```
7 2
```

```
8 1
```

```
9 1
```

```
dtype: int64
```

This tells us:

The value 3 occurs 4 times. The value 4 occurs 2 times. The value 7 occurs 2 times.

And so on.

Example 2: Count Frequency of Unique Values (Including NaNs)

By default, the value_counts() function does not show the frequency of NaN values.

However, you can use the dropna argument to display the frequency of NaN values:

```
import pandas as pd
import numpy as np

#create pandas Series with some NaN values
my_series = pd.Series()

#count occurrences of unique values in Series,
including NaNs
my_series.value_counts(dropna=False)

3.0 4
4.0 2
7.0 2
NaN 2
8.0 1
9.0 1
dtype: int64
```

Example 3: Count Relative Frequency of Unique Values

The following code shows how to use the normalize argument to count the relative frequency of unique values in a pandas Series:

```
import pandas as pd
```

```
#create pandas Series
```

```
my_series = pd.Series()
```

```
#count occurrences of unique values in Series
```

```
my_series.value_counts(normalize=True)
```

```
3 0.4
```

```
4 0.2
```

```
7 0.2
```

```
8 0.1
```

```
9 0.1
```

```
dtype: float64
```

The value 3 represents 40% of all values in the Series. The value 4 represents 20% of all values in the Series. The value 7 represents 20% of all values in the Series.

And so on.

Example 4: Count Frequency in Bins

The following code shows how to use the bins argument to count the frequency of values in a pandas

Series that fall into equal-sized bins:

```
import pandas as pd
```

```
#create pandas Series
```

```
my_series = pd.Series()
```

```
#count occurrences of unique values in Series
```

```
my_series.value_counts(bins=3)
```

```
(3.0, 5.0] 6
```

```
(5.0, 7.0] 2
```

```
(7.0, 9.0] 2
```

```
dtype: int64
```

This tells us:

There are 6 values that fall in the range 3 to 5. There are 2 values that fall in the range 5 to 7. There are 2 values that fall in the range 7 to 9.

Example 5: Count Frequency of Values in Pandas DataFrame

We can also use the `value_counts()` function to calculate the frequency of unique values in a specific column of a pandas DataFrame:

```
import pandas as pd
```

```
#create DataFrame
```

```
df = pd.DataFrame({'points': ,  
'assists': ,  
'rebounds': })
```

```
#count occurrences of unique values in 'points' column  
df.value_counts()
```

```
9 3
```

```
10 2
```

```
13 1
```

```
15 1
```

```
22 1
```

```
Name: points, dtype: int64
```

The following tutorials explain how to use other common functions in pandas: