

How can I count the missing values in a Pandas DataFrame?

Authored by
stats writer

April 18, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I count the missing values in a Pandas DataFrame?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=136805>

Counting the missing values in a Pandas DataFrame involves using the built-in functions provided by the Pandas library. These functions allow for a quick and efficient way to identify and count the number of missing values in a DataFrame. By using these functions, users can easily determine the extent to which their data is incomplete and make informed decisions on how to handle missing values in their dataset. This process is essential in data analysis and can be easily implemented in Python programming language.

Count Missing Values in a Pandas DataFrame

Often you may be interested in counting the number of missing values in a pandas DataFrame.

This tutorial shows several examples of how to count missing values using the following DataFrame:

```
import pandas as pd
import numpy as np

#create DataFrame with some missing values
df = pd.DataFrame({'a': ,
'b': ,
'c': })

#view DataFrameprint(df)
```

```
a b c
```

```
0 4.0 NaN 11.0
```

```
1 NaN 6.0 8.0
```

2 NaN 8.0 10.0
3 7.0 14.0 6.0
4 8.0 29.0 6.0
5 12.0 NaN NaN

Count the Total Missing Values in Entire DataFrame

The following code shows how to calculate the total number of missing values in the entire DataFrame:

```
df.isnull().sum().sum()
```

5

This tells us that there are 5 total missing values.

Count the Total Missing Values per Column

The following code shows how to calculate the total number of missing values in each column of the DataFrame:

```
df.isnull().sum()
```

a 2

b 2

c 1

This tells us:

Column 'a' has 2 missing values. Column 'b' has 2 missing values. Column 'c' has 1 missing value.

You can also display the number of missing values as a percentage of the entire column:

```
df.isnull().sum()/len(df)*100
```

```
a 33.333333
```

```
b 33.333333
```

```
c 16.666667
```

This tells us:

33.33% of values in Column 'a' are missing. 33.33% of values in Column 'b' are missing. 16.67% of values in Column 'c' are missing.

Count the Total Missing Values per Row

```
df.isnull().sum(axis=1)
```

```
0 1
```

```
1 1
```

```
2 1
```

3 0

4 0

5 2

This tells us:

Row 1 has 1 missing value. Row 2 has 1 missing value. Row 3 has 1 missing value. Row 4 has 0 missing values. Row 5 has 0 missing values. Row 6 has 2 missing values.

How to Find Unique Values in Multiple Columns in Pandas

How to Create a New Column Based on a Condition in Pandas