How can we fill NaN values in a Pandas dataframe using a dictionary?

Authored by

stats writer

June 25, 2024

The process of filling NaN (Not a Number) values in a Pandas dataframe using a dictionary involves using a predefined mapping of key-value pairs to replace the missing values. This method is useful for handling missing data and can be easily implemented in a Pandas dataframe using the "fillna()" function. By providing a dictionary as an argument to this function, the NaN values are replaced with the corresponding values from the dictionary. This allows for efficient and customizable handling of missing data in a Pandas dataframe.

Pandas: Fill NaN Values Using a Dictionary

You can use the fillna() function with a dictionary to replace NaN values in one column of a pandas DataFrame based on values in another column.

You can use the following basic syntax to do so:

```
#define dictionary
dict = {'A':5, 'B':10, 'C':15, 'D':20}

#replace values in col2 based on dictionary values in col1
df = df.fillna(df.map(dict))
```

The following example shows how to use this syntax in practice.

Example: Fill NaN Values in Pandas Using a Dictionary

Suppose we have the following pandas DataFrame that

contains information about the sales made at various retail stores:

```
import pandas as pd
import numpy as np

#create DataFrame
df = pd.DataFrame({'store': ,
'sales': })

#view DataFrame
print(df)

store sales
0 A 12.0
1 A NaN
2 B 30.0
3 C NaN
4 D 24.0
5 C NaN
6 B NaN
7 D 13.0
```

Notice that there are several NaN values in the sales column.

Suppose we would like to fill these NaNs in the sales column using values that correspond to specific values in the store column.

We can use the following syntax to do so:

```
#define dictionary
dict = {'A':5, 'B':10, 'C':15, 'D':20}

#replace values in sales column based on dictionary values in store column
df = df.fillna(df.map(dict))

#view updated DataFrameprint(df)
```

```
store sales
0 A 12.0
1 A 5.0
2 B 30.0
3 C 15.0
4 D 24.0
5 C 15.0
6 B 10.0
7 D 13.0
```

We used a dictionary to make the following replacements in the sales column:

If store is A, replace NaN in sales with the value 5.If store is B, replace NaN in sales with the value 10.If store is C, replace NaN in sales with the value 15.If store is D, replace NaN in sales with the value 20.

You can find the complete online documentation for the fillna() function .

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