

PANDAS FILLNA

SHARP SIGHT

WHAT YOU'LL LEARN

- How to use the Pandas `fillna()` method
- How to fill in missing data in a Pandas DataFrame

PANDAS FILLNA OVERVIEW

FILLNA FILLS IN MISSING VALUES IN A PANDAS DATAFRAME

- Data cleaning is one of the most important tasks of data science
- Handling missing values is one of the most common data cleaning issues
- The `fillna()` method enables you fill in missing values in Pandas
 - fill in missing all values
 - or, fill in missing values for a particular variable

EXAMPLE: YOU HAVE A PANDAS DATAFRAME

name	region	sales	expenses
William	East	50000	42000
Emma	NaN	52000	43000
Sofia	East	90000	NaN
Markus	South	NaN	44000
Edward	West	42000	38000
Thomas	West	72000	39000
Ethan	South	49000	42000
Olivia	West	NaN	NaN
Arun	West	67000	39000
Anika	East	65000	44000
Paulo	South	67000	45000

THE DATAFRAME HAS MISSING VALUES IN SEVERAL DIFFERENT COLUMNS

name	region	sales	expenses
William	East	50000	42000
Emma	NaN	52000	43000
Sofia	East	90000	NaN
Markus	South	NaN	44000
Edward	West	42000	38000
Thomas	West	72000	39000
Ethan	South	49000	42000
Olivia	West	NaN	NaN
Arun	West	67000	39000
Anika	East	65000	44000
Paulo	South	67000	45000

YOU CAN USE `FILLNA()` TO FILL IN THE MISSING DATA

```
salespeople.fillna(value = {'region': 'North',  
                           'sales': 0,  
                           'expenses': 0  
})
```

name	region	sales	expenses
William	East	50000	42000
Emma	NaN	52000	43000
Sofia	East	90000	NaN
Markus	South	NaN	44000
Edward	West	42000	38000
Thomas	West	72000	39000
Ethan	South	49000	42000
Olivia	West	NaN	NaN
Arun	West	67000	39000
Anika	East	65000	44000
Paulo	South	67000	45000

YOU CAN USE `FILLNA()` TO FILL IN THE MISSING DATA

```
salespeople.fillna(value = {'region': 'North',  
                           'sales': 0,  
                           'expenses': 0  
})
```

name	region	sales	expenses
William	East	50000	42000
Emma	North	52000	43000
Sofia	East	90000	0
Markus	South	0	44000
Edward	West	42000	38000
Thomas	West	72000	39000
Ethan	South	49000	42000
Olivia	West	0	0
Arun	West	67000	39000
Anika	East	65000	44000
Paulo	South	67000	45000

PANDAS FILLNA SYNTAX

SYNTAX: PANDAS FILLNA

The name of the DataFrame you want to operate on

The name of the method (called with "dot" syntax)

```
myDataFrame.fillna(value=, inplace =)
```

The value or values you want to use to replace the missing values

This can be a single value, or a dictionary of values

A True or False value that specifies if you want to directly modify the original dataframe

PARAMETERS OF PANDAS FILLNA

THE PARAMETERS OF PANDAS FILLNA

Input	What it does	Format
<code>value=</code>	Specifies the value or values you want to use to replace the missing values	Either a single value, or a dictionary of values that specify particular values to use for particular columns
<code>inplace=</code>	Specifies whether you want to modify the original DataFrame directly	True or False

THE OUTPUT OF PANDAS FILLNA

- By default, the `fillna()` method outputs a new DataFrame
 - by default, `inplace` is set to `inplace = False`
- If you set `inplace = True`, `fillna()` will directly modify the original dataframe

RECAP

RECAP OF WHAT WE LEARNED

- How to use the `fillna()` method
- How to fill in missing values in a dataframe
- **Next Steps:** Watch the code walkthrough video for step-by-step examples of how to use the Pandas `fillna()` method