

pandas

It will create dataframe

```
In [4]: import pandas as pd
```

```
In [6]: pd.__version__
```

```
Out[6]: '2.2.2'
```

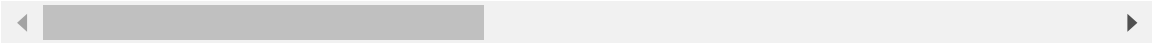
```
In [8]: store = pd.read_csv(r'C:\Users\samik\Downloads\Sample - Superstore_Orders.csv')
```

```
In [10]: store
```

Out[10]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	20103
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	20112
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	20112
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	20112
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	20141
...
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20143
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20115
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20143

10194 rows × 19 columns



In [12]:

```
store.shape
```

Out[12]: (10194, 19)

In [14]: `store.columns`

Out[14]: Index(['Category', 'City', 'Country/Region', 'Customer Name', 'Manufacturer', 'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region', 'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category', 'Discount', 'Profit', 'Quantity', 'Sales'], dtype='object')

In [16]: `len(store)`

Out[16]: 10194

In [18]: `len(store.columns)`

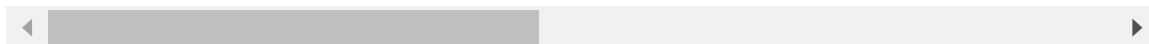
Out[18]: 19

In [20]: `store.isnull()`

Out[20]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
0	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False
...
10189	False	False	False	False	False	False	False	False
10190	False	False	False	False	False	False	False	False
10191	False	False	False	False	False	False	False	False
10192	False	False	False	False	False	False	False	False
10193	False	False	False	False	False	False	False	False

10194 rows × 19 columns



In [22]: `store.isnull().sum()`

```
Out[22]: Category      0
          City          0
          Country/Region 0
          Customer Name  0
          Manufacturer   0
          Order Date     0
          Order ID       0
          Postal Code     0
          Product Name    0
          Region         0
          Segment        0
          Ship Date      0
          Ship Mode       0
          State/Province  0
          Sub-Category    0
          Discount       0
          Profit         0
          Quantity       0
          Sales          0
          dtype: int64
```

```
In [24]: store.dtypes
```

```
Out[24]: Category      object
          City          object
          Country/Region object
          Customer Name  object
          Manufacturer   object
          Order Date     object
          Order ID       object
          Postal Code     object
          Product Name    object
          Region         object
          Segment        object
          Ship Date      object
          Ship Mode       object
          State/Province  object
          Sub-Category    object
          Discount       float64
          Profit         float64
          Quantity       int64
          Sales          float64
          dtype: object
```

```
In [30]: store.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10194 entries, 0 to 10193
Data columns (total 19 columns):
 #   Column                Non-Null Count  Dtype  
---  -
 0   Category              10194 non-null object  
 1   City                  10194 non-null object  
 2   Country/Region        10194 non-null object  
 3   Customer Name         10194 non-null object  
 4   Manufacturer           10194 non-null object  
 5   Order Date            10194 non-null object  
 6   Order ID              10194 non-null object  
 7   Postal Code           10194 non-null object  
 8   Product Name          10194 non-null object  
 9   Region                10194 non-null object  
10   Segment               10194 non-null object  
11   Ship Date             10194 non-null object  
12   Ship Mode             10194 non-null object  
13   State/Province        10194 non-null object  
14   Sub-Category          10194 non-null object  
15   Discount              10194 non-null float64 
16   Profit                10194 non-null float64 
17   Quantity              10194 non-null int64  
18   Sales                 10194 non-null float64 
dtypes: float64(3), int64(1), object(15)
memory usage: 1.5+ MB

```

In [55]: `store.head(2)`

Out[55]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Post Coc
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	7705
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	US-2020-112326	6054

In [32]: `id(store)`

Out[32]: 2270900517392

In [34]: `store.columns`

Out[34]: Index(['Category', 'City', 'Country/Region', 'Customer Name', 'Manufacturer', 'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region', 'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category', 'Discount', 'Profit', 'Quantity', 'Sales'], dtype='object')

```
In [38]: store[['City']]
```

```
Out[38]: 0          Houston
1          Naperville
2          Naperville
3          Naperville
4          Philadelphia
...
10189     New York City
10190          Fairfield
10191          Loveland
10192     New York City
10193     Charlottetown
Name: City, Length: 10194, dtype: object
```

```
In [40]: store[['City', 'Category']]
```

```
Out[40]:
```

	City	Category
0	Houston	Office Supplies
1	Naperville	Office Supplies
2	Naperville	Office Supplies
3	Naperville	Office Supplies
4	Philadelphia	Office Supplies
...
10189	New York City	Office Supplies
10190	Fairfield	Office Supplies
10191	Loveland	Office Supplies
10192	New York City	Technology
10193	Charlottetown	Office Supplies

10194 rows × 2 columns

```
In [42]: store[['City', 'Category']]
```

Out[42]:

	City	Category
0	Houston	Office Supplies
1	Naperville	Office Supplies
2	Naperville	Office Supplies
3	Naperville	Office Supplies
4	Philadelphia	Office Supplies
...
10189	New York City	Office Supplies
10190	Fairfield	Office Supplies
10191	Loveland	Office Supplies
10192	New York City	Technology
10193	Charlottetown	Office Supplies

10194 rows × 2 columns

In [44]:

```
store[['City', 'Category']]
```

Out[44]:

	City	Category
0	Houston	Office Supplies
1	Naperville	Office Supplies
2	Naperville	Office Supplies
3	Naperville	Office Supplies
4	Philadelphia	Office Supplies
...
10189	New York City	Office Supplies
10190	Fairfield	Office Supplies
10191	Loveland	Office Supplies
10192	New York City	Technology
10193	Charlottetown	Office Supplies

10194 rows × 2 columns

In [50]:

```
store[['City', 'Category', 'Order ID']]
```

Out[50]:

	City	Category	Order ID
0	Houston	Office Supplies	US-2020-103800
1	Naperville	Office Supplies	US-2020-112326
2	Naperville	Office Supplies	US-2020-112326
3	Naperville	Office Supplies	US-2020-112326
4	Philadelphia	Office Supplies	US-2020-141817
...
10189	New York City	Office Supplies	US-2023-143259
10190	Fairfield	Office Supplies	US-2023-115427
10191	Loveland	Office Supplies	US-2023-156720
10192	New York City	Technology	US-2023-143259
10193	Charlottetown	Office Supplies	CA-2023-143500

10194 rows × 3 columns

```
In [61]: store_text = store[[ 'Category', 'City', 'Country/Region', 'Customer Name', 'Man
    'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',
    'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category' ]]
```

```
In [65]: store_text
```


Out[65]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	20103
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	20112
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	20112
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	20112
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	20141
...
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20143
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20115
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20143

10194 rows × 15 columns



In [67]: store.columns

```
Out[67]: Index(['Category', 'City', 'Country/Region', 'Customer Name', 'Manufacturer',
              'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',
              'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category',
              'Discount', 'Profit', 'Quantity', 'Sales'],
              dtype='object')
```

```
In [69]: store_text.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10194 entries, 0 to 10193
Data columns (total 15 columns):
 #   Column                Non-Null Count  Dtype
---  -
 0   Category              10194 non-null  object
 1   City                  10194 non-null  object
 2   Country/Region        10194 non-null  object
 3   Customer Name          10194 non-null  object
 4   Manufacturer           10194 non-null  object
 5   Order Date             10194 non-null  object
 6   Order ID               10194 non-null  object
 7   Postal Code            10194 non-null  object
 8   Product Name           10194 non-null  object
 9   Region                 10194 non-null  object
10   Segment                10194 non-null  object
11   Ship Date              10194 non-null  object
12   Ship Mode              10194 non-null  object
13   State/Province         10194 non-null  object
14   Sub-Category           10194 non-null  object
dtypes: object(15)
memory usage: 1.2+ MB
```

```
In [77]: store_number = store[['Discount', 'Profit', 'Quantity', 'Sales']]
store_number
```

```
Out[77]:
```

	Discount	Profit	Quantity	Sales
0	0.2	5.5512	2	16.448
1	0.8	-5.4870	2	3.540
2	0.2	4.2717	3	11.784
3	0.2	-64.7748	3	272.736
4	0.2	4.8840	3	19.536
...
10189	0.2	19.7910	3	52.776
10190	0.2	6.4750	2	20.720
10191	0.2	-0.6048	3	3.024
10192	0.0	2.7279	7	90.930
10193	0.2	-0.6048	3	3.024

10194 rows × 4 columns

```
In [79]: store.columns
```

```
Out[79]: Index(['Category', 'City', 'Country/Region', 'Customer Name', 'Manufacturer',  
              'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',  
              'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category',  
              'Discount', 'Profit', 'Quantity', 'Sales'],  
             dtype='object')
```

```
In [81]: len(store.columns)
```

```
Out[81]: 19
```

```
In [83]: store_text.columns
```

```
Out[83]: Index(['Category', 'City', 'Country/Region', 'Customer Name', 'Manufacturer',  
              'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',  
              'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category'],  
             dtype='object')
```

```
In [85]: len(store_text.columns)
```

```
Out[85]: 15
```

```
In [87]: store_number.columns
```

```
Out[87]: Index(['Discount', 'Profit', 'Quantity', 'Sales'], dtype='object')
```

```
In [89]: len(store_number.columns)
```

```
Out[89]: 4
```

```
In [91]: print(len(store.columns))  
         print(len(store_text.columns))  
         print(len(store_number.columns))
```

```
19
```

```
15
```

```
4
```

```
In [93]: store.head()
```

Out[93]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Po C
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	77
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	US-2020-112326	60
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	US-2020-112326	60
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	US-2020-112326	60
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	US-2020-141817	19

In [95]:

```
store.tail()
```

Out[95]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20143
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20115
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20143

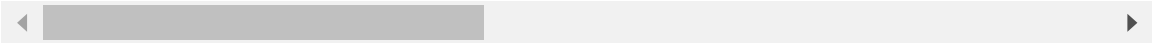
In [97]:

```
store[:]
```

Out[97]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	20103
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	20112
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	20112
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	20112
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	20141
...
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20143
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20115
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20143

10194 rows × 19 columns

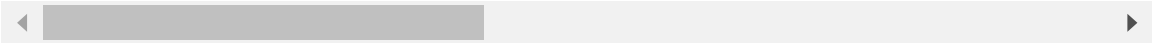


In [99]: store[:, -1]

Out[99]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20143
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20143
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20156
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20115
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20143
...
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	20141
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	20112
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	20112
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	20112
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	20103

10194 rows × 19 columns

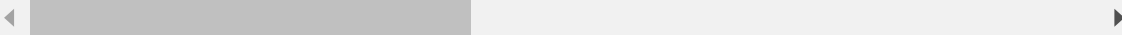


In [101...]

```
store[0:100:10]
```

Out[101...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800
10	Office Supplies	Henderson	United States	Maria Etezadi	Southworth	06-01-2020	US-2020-167199
20	Furniture	Dover	United States	Seth Vernon	DAX	11-01-2020	US-2020-130092
30	Office Supplies	San Francisco	United States	Brian Dahlen	Tennsco	13-01-2020	US-2020-157147
40	Office Supplies	Scottsdale	United States	Toby Swindell	GBC	19-01-2020	US-2020-146591
50	Office Supplies	Jonesboro	United States	Hunter Lopez	Newell	20-01-2020	US-2020-147627
60	Furniture	Miami	United States	Tom Boeckenhauer	Ultra	21-01-2020	US-2020-110422
70	Furniture	Alexandria	United States	Shirley Daniels	GE	26-01-2020	US-2020-155502
80	Office Supplies	New York City	United States	Jasper Cacioppo	Pressboard	28-01-2020	US-2020-100328
90	Office Supplies	Escondido	United States	Mark Haberlin	GBC	04-02-2020	US-2020-104808



In [105...

store[10:11]

Out[105...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Post Code
10	Office Supplies	Henderson	United States	Maria Etezadi	Southworth	06-01-2020	US-2020-167199	42



In [107...

```
store_text[:4]
```

Out[107...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Post Code
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	77058
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	US-2020-112326	60540
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	US-2020-112326	60540
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	US-2020-112326	60540

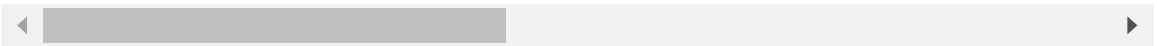


In [109...

```
store[1:10:3]
```

Out[109...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
1	Office Supplies	Naperville	United States	Phyllina Ober	GBC	04-01-2020	US-2020-112326	60
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	US-2020-141817	19
7	Office Supplies	Athens	United States	Jack O'Briant	Dixon	06-01-2020	US-2020-106054	30



In [111...

```
store.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10194 entries, 0 to 10193
Data columns (total 19 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Category              10194 non-null object
1   City                  10194 non-null object
2   Country/Region        10194 non-null object
3   Customer Name         10194 non-null object
4   Manufacturer           10194 non-null object
5   Order Date            10194 non-null object
6   Order ID              10194 non-null object
7   Postal Code           10194 non-null object
8   Product Name          10194 non-null object
9   Region                10194 non-null object
10  Segment               10194 non-null object
11  Ship Date             10194 non-null object
12  Ship Mode             10194 non-null object
13  State/Province        10194 non-null object
14  Sub-Category          10194 non-null object
15  Discount              10194 non-null float64
16  Profit                10194 non-null float64
17  Quantity              10194 non-null int64
18  Sales                 10194 non-null float64
dtypes: float64(3), int64(1), object(15)
memory usage: 1.5+ MB
```

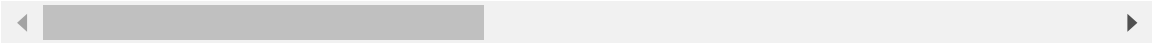
In [113...

```
store
```

Out[113...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	20103
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	20112
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	20112
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	20112
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	20141
...
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20143
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20115
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20143

10194 rows × 19 columns



In [115...

```
store[['Category', 'City', 'Manufacturer']]
```

Out[115...

	Category	City	Manufacturer
0	Office Supplies	Houston	Message Book
1	Office Supplies	Naperville	GBC
2	Office Supplies	Naperville	Avery
3	Office Supplies	Naperville	SAFCO
4	Office Supplies	Philadelphia	Avery
...
10189	Office Supplies	New York City	Wilson Jones
10190	Office Supplies	Fairfield	GBC
10191	Office Supplies	Loveland	Other
10192	Technology	New York City	Other
10193	Office Supplies	Charlottetown	Wilson Jones

10194 rows × 3 columns

In [118...

!pip install panda

Collecting panda

Downloading panda-0.3.1.tar.gz (5.8 kB)

Preparing metadata (setup.py): started

Preparing metadata (setup.py): finished with status 'done'

Requirement already satisfied: setuptools in c:\users\samik\anaconda3\lib\site-packages (from panda) (69.5.1)

Requirement already satisfied: requests in c:\users\samik\anaconda3\lib\site-packages (from panda) (2.32.2)

Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\samik\anaconda3\lib\site-packages (from requests->panda) (2.0.4)

Requirement already satisfied: idna<4,>=2.5 in c:\users\samik\anaconda3\lib\site-packages (from requests->panda) (3.7)

Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\samik\anaconda3\lib\site-packages (from requests->panda) (2.2.2)

Requirement already satisfied: certifi>=2017.4.17 in c:\users\samik\anaconda3\lib\site-packages (from requests->panda) (2024.7.4)

Building wheels for collected packages: panda

Building wheel for panda (setup.py): started

Building wheel for panda (setup.py): finished with status 'done'

Created wheel for panda: filename=panda-0.3.1-py3-none-any.whl size=7246 sha256=24f942ddb63c0ee75338e4b0908c7a557a630ebf38e264fd8863e47be90806e1

Stored in directory: c:\users\samik\appdata\local\pip\cache\wheels\98\41\5b\6ca54e0b6a35e1b7248c12f56fcb753dfb7717fefaa0fb45f5

Successfully built panda

Installing collected packages: panda

Successfully installed panda-0.3.1

In [120...

type(store)

Out[120...

pandas.core.frame.DataFrame

In [122...

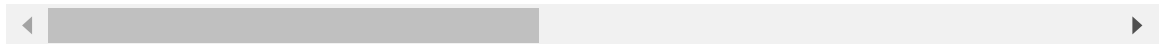
store.isnull()

The history saving thread hit an unexpected error (OperationalError('database or disk is full')).History will not be written to the database.

Out[122...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
0	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False
...
10189	False	False	False	False	False	False	False	False
10190	False	False	False	False	False	False	False	False
10191	False	False	False	False	False	False	False	False
10192	False	False	False	False	False	False	False	False
10193	False	False	False	False	False	False	False	False

10194 rows × 9 columns



In [126...

```
store.isnull().sum()
```

Out[126...

```
Category      0
City          0
Country/Region 0
Customer Name 0
Manufacturer   0
Order Date    0
Order ID      0
Postal Code    0
Product Name   0
Region        0
Segment       0
Ship Date     0
Ship Mode     0
State/Province 0
Sub-Category   0
Discount      0
Profit        0
Quantity      0
Sales         0
dtype: int64
```

In [128...

```
store_1 = pd.read_csv(r'C:\Users\samik\Downloads\Sample - Superstore_Orders.csv')
```

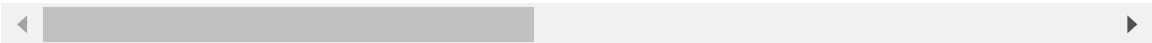
In [130...

```
store_1.isnull()
```

Out[130...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
0	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False
...
10189	False	False	False	False	False	False	False	False
10190	False	False	False	False	False	False	False	False
10191	False	False	False	False	False	False	False	False
10192	False	False	False	False	False	False	False	False
10193	False	False	False	False	False	False	False	False

10194 rows × 19 columns



In [132...

```
store_1.isnull().sum()
```

Out[132...

```
Category      0
City          0
Country/Region 0
Customer Name 0
Manufacturer   0
Order Date    0
Order ID      0
Postal Code   0
Product Name  0
Region        0
Segment       0
Ship Date     0
Ship Mode     0
State/Province 0
Sub-Category  0
Discount      0
Profit        0
Quantity      0
Sales         0
dtype: int64
```

In [134...

```
store_1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10194 entries, 0 to 10193
Data columns (total 19 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Category              10194 non-null  object
1   City                  10194 non-null  object
2   Country/Region        10194 non-null  object
3   Customer Name         10194 non-null  object
4   Manufacturer          10194 non-null  object
5   Order Date            10194 non-null  object
6   Order ID              10194 non-null  object
7   Postal Code           10194 non-null  object
8   Product Name          10194 non-null  object
9   Region                10194 non-null  object
10  Segment               10194 non-null  object
11  Ship Date             10194 non-null  object
12  Ship Mode             10194 non-null  object
13  State/Province        10194 non-null  object
14  Sub-Category          10194 non-null  object
15  Discount              10194 non-null  float64
16  Profit                10194 non-null  float64
17  Quantity              10194 non-null  int64
18  Sales                 10194 non-null  float64
dtypes: float64(3), int64(1), object(15)
memory usage: 1.5+ MB
```

```
In [172... store_1.columns
```

```
Out[172... Index(['Category', 'City', 'Country/Region', 'Customer Name', 'Manufacturer',
      'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',
      'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category',
      'Discount', 'Profit', 'Quantity', 'Sales'],
      dtype='object')
```

```
In [174... store_1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10194 entries, 0 to 10193
Data columns (total 19 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Category              10194 non-null  object
1   City                  10194 non-null  object
2   Country/Region        10194 non-null  object
3   Customer Name         10194 non-null  object
4   Manufacturer           10194 non-null  object
5   Order Date            10194 non-null  object
6   Order ID              10194 non-null  object
7   Postal Code           10194 non-null  object
8   Product Name          10194 non-null  object
9   Region                10194 non-null  object
10  Segment               10194 non-null  object
11  Ship Date             10194 non-null  object
12  Ship Mode             10194 non-null  object
13  State/Province        10194 non-null  object
14  Sub-Category          10194 non-null  object
15  Discount              10194 non-null  float64
16  Profit                10194 non-null  float64
17  Quantity              10194 non-null  int64
18  Sales                 10194 non-null  float64
dtypes: float64(3), int64(1), object(15)
memory usage: 1.5+ MB
```

```
In [176... store_1.columns
```

```
Out[176... Index(['Category', 'City', 'Country/Region', 'Customer Name', 'Manufacturer',
      'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',
      'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category',
      'Discount', 'Profit', 'Quantity', 'Sales'],
      dtype='object')
```

```
In [178... store_1['Category']
```

```
Out[178... 0      Office Supplies
1      Office Supplies
2      Office Supplies
3      Office Supplies
4      Office Supplies
...
10189   Office Supplies
10190   Office Supplies
10191   Office Supplies
10192      Technology
10193   Office Supplies
Name: Category, Length: 10194, dtype: object
```

```
In [180... store_1[['Customer Name', 'City']]
```


Out[180...

	Customer Name	City
0	Darren Powers	Houston
1	Phillina Ober	Naperville
2	Phillina Ober	Naperville
3	Phillina Ober	Naperville
4	Mick Brown	Philadelphia
...
10189	Patrick O'Donnell	New York City
10190	Erica Bern	Fairfield
10191	Jill Matthias	Loveland
10192	Patrick O'Donnell	New York City
10193	Harry Olson	Charlottetown

10194 rows × 2 columns

In [182...

```
store_1_categorical = store_1[['Category', 'City', 'Country/Region', 'Customer N',
                               'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',
                               'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category']]
```

In [150...

```
store_1_categorical
```

Out[150...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	20103
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	20112
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	20112
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	20112
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	20141
...
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20143
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20115
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20143

10194 rows × 15 columns



In [184...

```
store_1_categorical.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10194 entries, 0 to 10193
Data columns (total 15 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Category              10194 non-null  object
1   City                  10194 non-null  object
2   Country/Region        10194 non-null  object
3   Customer Name         10194 non-null  object
4   Manufacturer          10194 non-null  object
5   Order Date            10194 non-null  object
6   Order ID              10194 non-null  object
7   Postal Code           10194 non-null  object
8   Product Name          10194 non-null  object
9   Region                10194 non-null  object
10  Segment               10194 non-null  object
11  Ship Date             10194 non-null  object
12  Ship Mode             10194 non-null  object
13  State/Province        10194 non-null  object
14  Sub-Category          10194 non-null  object
dtypes: object(15)
memory usage: 1.2+ MB
```

```
In [186... store_1_categorical = store_1[['Discount','Profit','Quantity','Sales']]
store_1_categorical
```

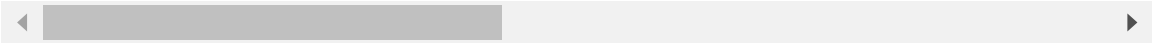
```
Out[186...
      Discount  Profit  Quantity  Sales
0          0.2    5.5512         2   16.448
1          0.8   -5.4870         2    3.540
2          0.2    4.2717         3   11.784
3          0.2  -64.7748         3  272.736
4          0.2    4.8840         3   19.536
...         ...      ...      ...    ...
10189        0.2   19.7910         3   52.776
10190        0.2    6.4750         2   20.720
10191        0.2   -0.6048         3    3.024
10192        0.0    2.7279         7   90.930
10193        0.2   -0.6048         3    3.024
```

10194 rows × 4 columns

```
In [188... store_1.head()
```

Out[188...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Po C
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	77
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	US-2020-112326	60
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	US-2020-112326	60
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	US-2020-112326	60
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	US-2020-141817	19



In [190...

```
store_1.tail()
```

Out[190...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20143
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20115
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20143

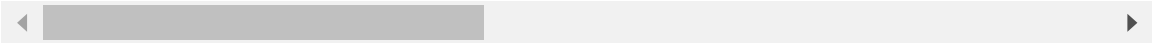
In [192...

```
store_1[:]
```

Out[192...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	20103
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	20112
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	20112
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	20112
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	20141
...
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20143
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20115
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20143

10194 rows × 19 columns



In [194...

```
store_1[0:10]
```

Out[194...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Po C
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	77
1	Office Supplies	Naperville	United States	Phyllina Ober	GBC	04-01-2020	US-2020-112326	60
2	Office Supplies	Naperville	United States	Phyllina Ober	Avery	04-01-2020	US-2020-112326	60
3	Office Supplies	Naperville	United States	Phyllina Ober	SAFCO	04-01-2020	US-2020-112326	60
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	US-2020-141817	19
5	Furniture	Henderson	United States	Maria Etezadi	Global	06-01-2020	US-2020-167199	42
6	Office Supplies	Henderson	United States	Maria Etezadi	Rogers	06-01-2020	US-2020-167199	42
7	Office Supplies	Athens	United States	Jack O'Briant	Dixon	06-01-2020	US-2020-106054	30
8	Office Supplies	Henderson	United States	Maria Etezadi	Ibico	06-01-2020	US-2020-167199	42
9	Office Supplies	Henderson	United States	Maria Etezadi	Alliance	06-01-2020	US-2020-167199	42

Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Po C

In [196...

```
store_1[0:10:2]
```

Out[196...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Po C
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	77
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	US-2020-112326	60
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	US-2020-141817	19
6	Office Supplies	Henderson	United States	Maria Etezadi	Rogers	06-01-2020	US-2020-167199	42
8	Office Supplies	Henderson	United States	Maria Etezadi	Ibico	06-01-2020	US-2020-167199	42

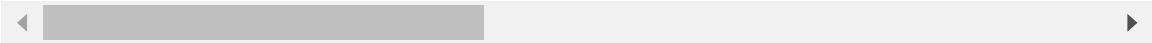
In [198...

```
store_1[:, -1]
```


Out[198...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	20143
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	20143
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	20156
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	20115
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	20143
...
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	20141
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	20112
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	20112
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	20112
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	20103

10194 rows × 19 columns



In [202...

```
store_1[6:7]
```

Out[202...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Pos Co
6	Office Supplies	Henderson	United States	Maria Etezadi	Rogers	06-01-2020	US-2020-167199	424



In []: