

```
In [68]: import pandas as pd
```

```
In [70]: pd.__version__
```

```
Out[70]: '2.2.2'
```

```
In [72]: store=pd.read_csv(r'C:\Users\Lenovo\Desktop\Sandya\PythonAI\FSDS-AI-830\Pandas\S
```

```
In [74]: store
```

Out[74]:

	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 [76]:

```
id(store)
```

Out[76]: 1315545333344

In [78]: `len(store)`

Out[78]: 10194

In [80]: `store.shape`

Out[80]: (10194, 19)

In [82]: `store.columns`

Out[82]: 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 [84]: `len(store.columns)`

Out[84]: 19

In [86]: `store.isnull()`

Out[86]:

	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 []:

In [89]: `store.isnull().sum()`

```
Out[89]: 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 [91]: store[:]
```

Out[91]:

	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 [93]:

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

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
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

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Pos
9	Office Supplies	Henderson	United States	Maria Etezadi	Alliance	06-01-2020	US-2020-167199	42

In [95]: store[0:20:5]

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Pos
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	
5	Furniture	Henderson	United States	Maria Etezadi	Global	06-01-2020	US-2020-167199	
10	Office Supplies	Henderson	United States	Maria Etezadi	Southworth	06-01-2020	US-2020-167199	
15	Office Supplies	Huntsville	United States	Vivek Sundaresam	Acco	07-01-2020	US-2020-105417	

In []:

In [98]: store.head()

Out[98]:

	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	7701
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
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	US-2020-141817	19101

In [100]:

```
store.head(3)
```

Out[100]:

	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	7701
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

In []:

```
store.tail()
```


Out[103...

	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 []:

In []:

In [107...

```
store.tail(1)
```

Out[107...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	201435



In [109...

```
store.isna()
```

Out[109...

	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 [111...

```
store.describe
```

```

Out[111... <bound method NDFrame.describe of
y/Region      Customer Name \
0      Office Supplies      Houston United States      Darren Powers
1      Office Supplies      Naperville United States      Phillina Ober
2      Office Supplies      Naperville United States      Phillina Ober
3      Office Supplies      Naperville United States      Phillina Ober
4      Office Supplies      Philadelphia United States      Mick Brown
...      ...      ...      ...      ...
10189 Office Supplies      New York City United States      Patrick O'Donnell
10190 Office Supplies      Fairfield United States      Erica Bern
10191 Office Supplies      Loveland United States      Jill Matthias
10192      Technology      New York City United States      Patrick O'Donnell
10193 Office Supplies      Charlottetown      Canada      Harry Olson

      Manufacturer      Order Date      Order ID      Postal Code \
0      Message Book      03-01-2020      US-2020-103800      77015
1      GBC      04-01-2020      US-2020-112326      60540
2      Avery      04-01-2020      US-2020-112326      60540
3      SAFCO      04-01-2020      US-2020-112326      60540
4      Avery      05-01-2020      US-2020-141817      19143
...      ...      ...      ...      ...
10189 Wilson Jones      30-12-2023      US-2023-143259      10009
10190      GBC      30-12-2023      US-2023-115427      94533
10191      Other      30-12-2023      US-2023-156720      80538
10192      Other      30-12-2023      US-2023-143259      10009
10193 Wilson Jones      30-12-2023      CA-2023-143500      C0A

      Product Name      Region \
0      Message Book, Wirebound, Four 5 1/2" X 4" Form...      Central
1      GBC Standard Plastic Binding Systems Combs      Central
2      Avery 508      Central
3      SAFCO Boltless Steel Shelving      Central
4      Avery Hi-Liter EverBold Pen Style Fluorescent ...      East
...      ...      ...
10189      Wilson Jones Legal Size Ring Binders      East
10190      GBC Binding covers      West
10191      Bagged Rubber Bands      West
10192      Gear Head AU3700S Headset      East
10193      Wilson Jones Impact Binders      East

      Segment      Ship Date      Ship Mode      State/Province \
0      Consumer      07-01-2020      Standard Class      Texas
1      Home Office      08-01-2020      Standard Class      Illinois
2      Home Office      08-01-2020      Standard Class      Illinois
3      Home Office      08-01-2020      Standard Class      Illinois
4      Consumer      12-01-2020      Standard Class      Pennsylvania
...      ...      ...      ...      ...
10189 Consumer      03-01-2024      Standard Class      New York
10190 Corporate      03-01-2024      Standard Class      California
10191 Consumer      03-01-2024      Standard Class      Colorado
10192 Consumer      03-01-2024      Standard Class      New York
10193 Consumer      03-01-2024      Standard Class      Prince Edward Island

      Sub-Category      Discount      Profit      Quantity      Sales
0      Paper      0.2      5.5512      2      16.448
1      Binders      0.8      -5.4870      2      3.540
2      Labels      0.2      4.2717      3      11.784
3      Storage      0.2      -64.7748      3      272.736
4      Art      0.2      4.8840      3      19.536
...      ...      ...      ...      ...

```

10189	Binders	0.2	19.7910	3	52.776
10190	Binders	0.2	6.4750	2	20.720
10191	Fasteners	0.2	-0.6048	3	3.024
10192	Phones	0.0	2.7279	7	90.930
10193	Binders	0.2	-0.6048	3	3.024

[10194 rows x 19 columns]>

In [113...

```
store.describe()
```

Out[113...

	Discount	Profit	Quantity	Sales
count	10194.000000	10194.000000	10194.000000	10194.000000
mean	0.155385	28.673417	3.791838	228.225854
std	0.206249	232.465115	2.228317	619.906839
min	0.000000	-6599.978000	1.000000	0.444000
25%	0.000000	1.760800	2.000000	17.220000
50%	0.200000	8.690000	3.000000	53.910000
75%	0.200000	29.297925	5.000000	209.500000
max	0.800000	8399.976000	14.000000	22638.480000

In [127...

```
store['Category']
```

Out[127...

```
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 [125...

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

Out[125...

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

10194 rows × 3 columns

In [131...

store.columns

Out[131...

```
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 [150...

```
store_cat=store[['Category', 'City', 'Country/Region', 'Customer Name', 'Manufac
                'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Regio
                'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category']]
```

In [145...

store_cat.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 [154... store_num=store[['Discount', 'Profit', 'Quantity', 'Sales']]
```

```
In [156... store_num.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10194 entries, 0 to 10193
Data columns (total 4 columns):
#   Column    Non-Null Count  Dtype
---  ---
0   Discount  10194 non-null  float64
1   Profit    10194 non-null  float64
2   Quantity  10194 non-null  int64
3   Sales     10194 non-null  float64
dtypes: float64(3), int64(1)
memory usage: 318.7 KB
```

```
In [160... print(len(store.columns))
print(len(store_cat.columns))
print(len(store_num.columns))
```

```
19
15
4
```

```
In [171... store_cat.describe()
```

Out[171...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Po C
count	10194	10194	10194	10194	10194	10194	10194	10
unique	3	542	2	800	183	1242	5111	
top	Office Supplies	New York City	United States	William Brown	Other	05-09-2022	US-2023-100111	10
freq	6128	915	9994	41	1940	38	14	

In [174...

```
store_num.describe()
```

Out[174...

	Discount	Profit	Quantity	Sales
count	10194.000000	10194.000000	10194.000000	10194.000000
mean	0.155385	28.673417	3.791838	228.225854
std	0.206249	232.465115	2.228317	619.906839
min	0.000000	-6599.978000	1.000000	0.444000
25%	0.000000	1.760800	2.000000	17.220000
50%	0.200000	8.690000	3.000000	53.910000
75%	0.200000	29.297925	5.000000	209.500000
max	0.800000	8399.976000	14.000000	22638.480000

In []:

In []:

In []:

In []:

In []: