

```
In [13]: import pandas as pd
```

```
In [17]: pd.__version__
```

```
Out[17]: '2.2.2'
```

```
In [31]: store = pd.read_csv(r'C:\Users\Vansh\OneDrive\Desktop\Sample - Superstore_Orders
```

```
In [33]: store
```

Out[33]:

	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 [37]: id(store)

Out[37]: 2445493562032

```
In [39]: len(store)
```

Out[39]: 10194

```
In [41]: store.shape
```

Out[41]: (10194, 19)

```
In [43]: store.columns
```

Out[43]: 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 [55]: store.isnull()
```

Out[55]:

	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 [57]: store.isnull().sum()
```

```
Out[57]: 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 [59]: store[:]
```

Out[59]:

	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 [61]: store[0:20:5]

Out[61]:

	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	
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 [63]:

store.head

```

Out[63]: <bound method NDFrame.head of
ion      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 [65]: store.head(3)
```

Out[65]:

	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	7701
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	US-2020-112326	6054
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	US-2020-112326	6054

```
In [67]: store.tail()
```


Out[67]:

	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 [69]: store.isna

Out[69]: <bound method DataFrame.isna of

	region	Customer Name \	City	Country/R	Category
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 [71]: `store.isna()`

Out[71]:

	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 x 19 columns



introduction to stastical concepts in pandas

In [35]: `store.describe()`

Out[35]:

	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 []: