

```
In [1]: import pandas as pd
```

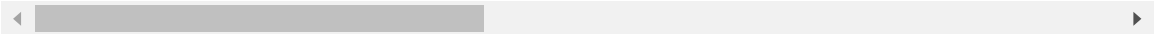
```
In [19]: store = pd.read_csv(r"C:\Users\arpan\Downloads\Sample - Superstore_Orders.csv")
```

```
In [21]: store
```

Out[21]:

	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 [23]: id(store)

Out[23]: 2511349397312

```
In [27]: len(store)
```

Out[27]: 10194

```
In [32]: store.columns
```

```
Out[32]: 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 [36]: store.shape
```

Out[36]: (10194, 19)

```
In [38]: len(store.columns)
```

Out[38]: 19

```
In [40]: store.dtypes
```

```
Out[40]: 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 [42]: 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

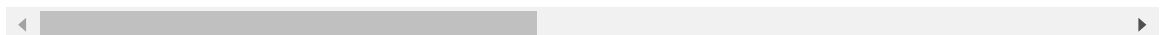
```

In [44]: `store.isnull()`

Out[44]:

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

```
Out[46]: 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 [50]: store.head()
```

```
Out[50]:
```

	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 [52]: store.tail()
```

Out[52]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Post Code
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 [54]:

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

Out[54]:

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

In [56]:

```
store.tail(3)
```

Out[56]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order
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

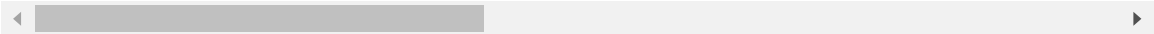
In [58]:

```
store[:]
```

Out[58]:

	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



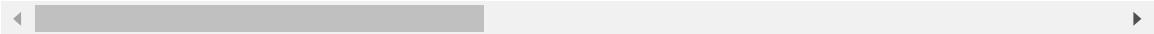
In [60]:

```
store[:: -1]
```


Out[60]:

	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



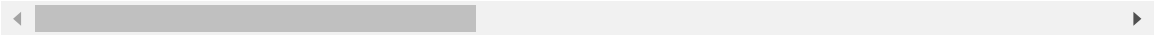
In [62]:

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

Out[62]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	1038
10	Office Supplies	Henderson	United States	Maria Etezadi	Southworth	06-01-2020	167
20	Furniture	Dover	United States	Seth Vernon	DAX	11-01-2020	1306
30	Office Supplies	San Francisco	United States	Brian Dahlen	Tennsco	13-01-2020	157
40	Office Supplies	Scottsdale	United States	Toby Swindell	GBC	19-01-2020	1465
...
10150	Office Supplies	Newark	United States	Dan Reichenbach	BIC	27-12-2023	1344
10160	Office Supplies	New York City	United States	Jennifer Ferguson	Storex	28-12-2023	1648
10170	Technology	New York City	United States	Jennifer Ferguson	Other	28-12-2023	1648
10180	Office Supplies	New York City	United States	Michael Chen	Other	29-12-2023	1026
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	1154

1020 rows × 19 columns



In [64]:

```
store[:10]
```

Out[64]:

	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	Po C
9	Office Supplies	Henderson	United States	Maria Etezadi	Alliance	06-01-2020	US-2020-167199	42

```
In [66]: store[10:]
```

Out[66]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	
10	Office Supplies	Henderson	United States	Maria Etezadi	Southworth	06-01-2020	1
11	Office Supplies	Los Angeles	United States	Lycoris Saunders	Xerox	06-01-2020	1
12	Technology	Henderson	United States	Maria Etezadi	Other	06-01-2020	1
13	Technology	Henderson	United States	Maria Etezadi	GE	06-01-2020	1
14	Furniture	Huntsville	United States	Vivek Sundaresam	Howard Miller	07-01-2020	1
...
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	1
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	1
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	1
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	1
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	1

10184 rows × 19 columns

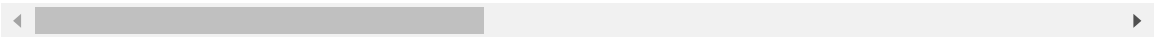


In [72]: store

Out[72]:

	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 [74]: store.dtypes
```

```
Out[74]: 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 [76]: store.columns
```

```
Out[76]: 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 [78]: store['Category']
```

```
Out[78]: 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 [80]: store['City']
```

```
Out[80]: 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 [86]: store[['Category', 'City']]
```

Out[86]:

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

10194 rows × 2 columns

```
In [116... store_categorical_data = store[['Category', 'City', 'Country/Region', 'Customer
      'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',
      'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category',]]
store_categorical_data
```


Out[116...

	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 [100... store_numerical_data = store[['Discount', 'Profit', 'Quantity', 'Sales']]
store_numerical_data
```

```
Out[100...
```

	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 [98]: store_numerical_data.dtypes
```

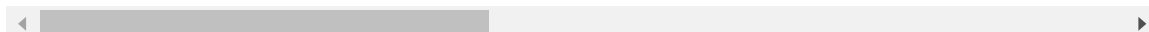
```
Out[98]: Discount    float64
Profit      float64
dtype: object
```

```
In [102... store
```

Out[102...

	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 [104...

store.describe()

Out[104...

	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 []: