

```
In [3]: import numpy as np  
import pandas as pd
```

```
In [5]: store = pd.read_csv(r"C:\Users\MANISHA\Downloads\Sample - Superstore_Orders.csv")
```

```
In [7]: store
```

Out[7]:

	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
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	US-2020-112326
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	US-2020-112326
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	US-2020-112326
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	US-2020-141817
...							
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	US-2023-143259
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	US-2023-115427
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	US-2023-156720
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	US-2023-143259
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	CA-2023-143500

10194 rows × 19 columns

```
In [9]: id(store)
```

```
Out[9]: 2310439289824
```

```
In [11]: len(store)
```

```
Out[11]: 10194
```

```
In [13]: store.columns
```

```
Out[13]: 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 [17]: store.shape #dataset contains 10194 rows * 19 columns
```

```
Out[17]: (10194, 19)
```

```
In [19]: len(store.columns)
```

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

```
In [21]: store.isnull()
```

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code	F
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 [25]: store.isna()
```

Out[25]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code	F
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 [29]: store.isnull().sum()
```

```
Out[29]: 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 [31]: store.dtypes
```

```
Out[31]: 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 [33]: 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 [35]: pd.__version__
```

```
Out[35]: '2.2.2'
```

```
In [37]: store.head()
```

Out[37]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	77095
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	19143

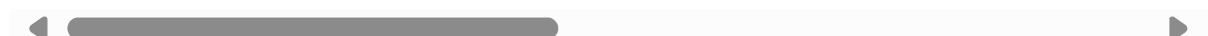


In [39]:

store.head(3)

Out[39]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	77095
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 [41]: `store.tail(3)`

Out[41]:

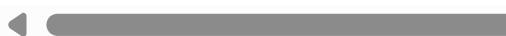
	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	US-2023-156720
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	US-2023-143259
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	CA-2023-143500



In [43]: `store.tail(3)`

Out[43]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	US-2023-156720
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	US-2023-143259
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	CA-2023-143500



In [45]: `store`

Out[45]:

	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
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	US-2020-112326
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	US-2020-112326
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	US-2020-112326
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	US-2020-141817
...							
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	US-2023-143259
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	US-2023-115427
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	US-2023-156720
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	US-2023-143259
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	CA-2023-143500

10194 rows × 19 columns

In [47]: `store[:]`

Out[47]:

	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
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	US-2020-112326
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	US-2020-112326
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	US-2020-112326
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	US-2020-141817
...							
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	US-2023-143259
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	US-2023-115427
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	US-2023-156720
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	US-2023-143259
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	CA-2023-143500

10194 rows × 19 columns

```
In [49]: store[0:50:10]
```

Out[49]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	77095
10	Office Supplies	Henderson	United States	Maria Etezadi	Southworth	06-01-2020	US-2020-167199	42420
20	Furniture	Dover	United States	Seth Vernon	DAX	11-01-2020	US-2020-130092	19901
30	Office Supplies	San Francisco	United States	Brian Dahlen	Tennsco	13-01-2020	US-2020-157147	94109
40	Office Supplies	Scottsdale	United States	Toby Swindell	GBC	19-01-2020	US-2020-146591	85254

```
In [51]: store.head(1)
```

Out[51]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800	V F X

```
In [55]: store['Category']
```

```
Out[55]: 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 [59]: store[['Category', 'City', 'Customer Name']]
```

```
Out[59]:
```

	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 [63]: store.columns
```

```
Out[63]: 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 [65]: store.dtypes
```

```
Out[65]: 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 [67]: store_num = store[['Discount', 'Profit', 'Quantity', 'Sales']]
store_num
```

```
Out[67]:
```

	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 [71]: store.columns
```

```
Out[71]: 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 [73]: store_cate = store[['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',  
    ]]
```

```
In [75]: store_cate.dtypes
```

```
Out[75]: 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  
dtype: object
```

```
In [79]: store['Profit'].mean()
```

```
Out[79]: 28.673417166960963
```

```
In [81]: store['Profit'].mode()
```

```
Out[81]: 0    0.0  
Name: Profit, dtype: float64
```

```
In [85]: store['Profit'].median()
```

```
Out[85]: 8.69
```

```
In [87]: store['Profit'].var()
```

```
Out[87]: 54040.02971828826
```

```
In [89]: store['Profit'].std()
```

```
Out[89]: 232.46511505662147
```

```
In [ ]:
```