

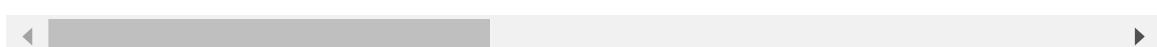
```
In [1]: import numpy as np  
import pandas as pd
```

```
In [2]: store=pd.read_csv(r'C:\Users\user\Downloads\14th- Introduction to pandas\storeor  
store
```

Out[2]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	O
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 [3]: `id(store)`

Out[3]: 2082278956816

In [4]: len(store)

Out[4]: 10194

In [5]: store.columns

Out[5]: 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 [6]: store.shape #dataset contains 10194 rows \* 19 column

Out[6]: (10194, 19)

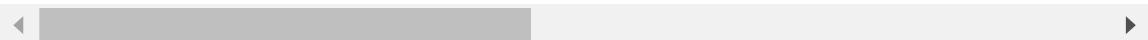
In [7]: len(store.columns)

Out[7]: 19

In [8]: store.isnull()

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

10194 rows × 19 columns



In [9]: store.isnull().sum()

```
Out[9]: 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 [10]: store.dtypes
```

```
Out[10]: 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 [11]: 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 [12]: pd.__version__
```

```
Out[12]: '2.2.2'
```

```
In [13]: store.head()
```

Out[13]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Pa 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 [14]: `store.tail(3)`

Out[14]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	O
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	21156
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	21143
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	21143

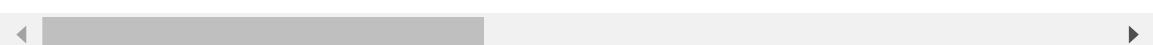
◀ ▶

In [15]: `store`

Out[15]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	O
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

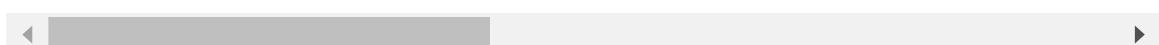


In [16]: store[:]

Out[16]:

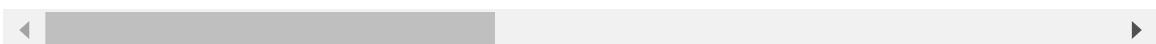
	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	O
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	20103
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	20112
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	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 [17]: `store[0:50:10] #it takes values from 0 to 50 by 10 count`

Out[17]:

	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	77
10	Office Supplies	Henderson	United States	Maria Etezadi	Southworth	06-01-2020	US-2020-167199	42
20	Furniture	Dover	United States	Seth Vernon	DAX	11-01-2020	US-2020-130092	19
30	Office Supplies	San Francisco	United States	Brian Dahlen	Tennsco	13-01-2020	US-2020-157147	94
40	Office Supplies	Scottsdale	United States	Toby Swindell	GBC	19-01-2020	US-2020-146591	85

In [19]: `store.head(1)`

Out[19]:

	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

In [20]: `store['Category']`

Out[20]:

```

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

Out[21]:

	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 [22]: `store.columns`

Out[22]: `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 [23]: `store.dtypes`

Out[23]:

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 [24]: `store_num = store[['Discount', 'Profit', 'Quantity', 'Sales']]`

Out[24]:

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

10194 rows × 4 columns

In [25]: `store.columns`

```
Out[25]: 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 [27]: `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',
 ]]`  
`store_cate`

Out[27]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	O
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 [28]: `store_cate.dtypes`

```
Out[28]: 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 [29]: store['Profit'].mean()
```

```
Out[29]: 28.673417166960963
```

```
In [30]: store['Profit'].median()
```

```
Out[30]: 8.69
```

```
In [31]: store['Profit'].mode()
```

```
Out[31]: 0    0.0
         Name: Profit, dtype: float64
```

```
In [32]: store['Profit'].var()
```

```
Out[32]: 54040.02971828826
```

```
In [33]: store['Profit'].std()
```

```
Out[33]: 232.46511505662147
```

```
In [34]: store.describe()
```

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

```
In [ ]:
```

