

```
In [53]: #pip install pandas
```

```
In [54]: import pandas as pd
```

```
In [55]: Store=pd.read_csv(r"D:\NIT\NOV\17nov (pandas)\17th\json,csv,excel read\Sample - Sup
```

```
In [56]: Store
```

Out[56]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2019	US-2019-103800
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2019	US-2019-112326
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2019	US-2019-112326
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2019	US-2019-112326
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2019	US-2019-141817
...							
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2022	US-2022-143259
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2022	US-2022-115427
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2022	US-2022-156720
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2022	US-2022-143259
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2022	CA-2022-143500

10194 rows × 19 columns

no. of rows in the dataset

```
In [57]: len(Store)
```

```
Out[57]: 10194
```

no of column in the dataset

```
In [58]: store_col=Store.columns
```

```
In [59]: store_col
```

```
Out[59]: 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 [60]: len(store_col)
```

```
Out[60]: 19
```

dimension of the dataframe (row & column)

```
In [61]: Store.shape
```

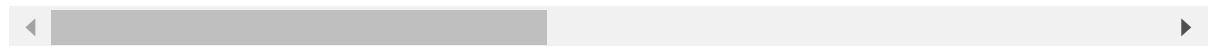
```
Out[61]: (10194, 19)
```

top 5 rows

```
In [62]: Store.head()
```

Out[62]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2019	US-2019-103800	77095
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2019	US-2019-112326	60540
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2019	US-2019-112326	60540
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2019	US-2019-112326	60540
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2019	US-2019-141817	19143

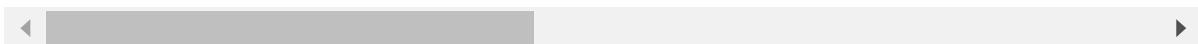


bottom 5 rows

In [63]: `Store.tail()`

Out[63]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2022	US-2022-143259
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2022	US-2022-115427
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2022	US-2022-156720
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2022	US-2022-143259
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2022	CA-2022-143500

In [64]: `store_reverse_df = Store[::-1]`In [65]: `store_reverse_df`

Out[65]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2022	CA-2022-143500
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2022	US-2022-143259
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2022	US-2022-156720
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2022	US-2022-115427
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2022	US-2022-143259
...
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2019	US-2019-141817
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2019	US-2019-112326
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2019	US-2019-112326
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2019	US-2019-112326
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2019	US-2019-103800

10194 rows × 19 columns

Give you the missing value

In [66]: `Store.isnull()`

Out[66]:

	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

Give you the count of missing value

In [67]: `Store.isnull().sum()`

```
Out[67]: 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
```

information about data

```
In [68]: 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 [69]: Store
```

Out[69]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2019	US-2019-103800
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2019	US-2019-112326
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2019	US-2019-112326
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2019	US-2019-112326
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2019	US-2019-141817
...							
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2022	US-2022-143259
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2022	US-2022-115427
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2022	US-2022-156720
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2022	US-2022-143259
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2022	CA-2022-143500

10194 rows × 19 columns

In [70]: `Store[:10]`

Out[70]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2019	US-2019-103800	77095
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2019	US-2019-112326	60540
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2019	US-2019-112326	60540
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2019	US-2019-112326	60540
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2019	US-2019-141817	19143
5	Furniture	Henderson	United States	Maria Etezadi	Global	06-01-2019	US-2019-167199	42420
6	Office Supplies	Henderson	United States	Maria Etezadi	Rogers	06-01-2019	US-2019-167199	42420
7	Office Supplies	Athens	United States	Jack O'Briant	Dixon	06-01-2019	US-2019-106054	30605
8	Office Supplies	Henderson	United States	Maria Etezadi	Ibico	06-01-2019	US-2019-167199	42420

Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
9	Office Supplies	Henderson	United States	Maria Etezadi	Alliance	06-01-2019	US-2019-167199

display the data types of each column

In [71]: `Store.dtypes`

```
Out[71]: 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 [72]: `Store['Discount']`

```
Out[72]: 0      0.2
1      0.8
2      0.2
3      0.2
4      0.2
...
10189 0.2
10190 0.2
10191 0.2
10192 0.0
10193 0.2
Name: Discount, Length: 10194, dtype: float64
```

In [73]: `store_numerical = Store[['Discount', 'Profit', 'Quantity', 'Sales']]`

```
In [74]: store_numerical
```

```
Out[74]:
```

	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 [75]: store_numerical.head(2)
```

```
Out[75]:
```

	Discount	Profit	Quantity	Sales
0	0.2	5.5512	2	16.448
1	0.8	-5.4870	2	3.540

```
In [76]: store_category = Store[['Category', 'City', 'Country/Region', 'Customer Name', 'Manager Name', 'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region', 'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category']]
```

```
In [77]: store_category
```

Out[77]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2019	US-2019-103800
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2019	US-2019-112326
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2019	US-2019-112326
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2019	US-2019-112326
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2019	US-2019-141817
...							
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2022	US-2022-143259
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2022	US-2022-115427
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2022	US-2022-156720
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2022	US-2022-143259
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2022	CA-2022-143500

10194 rows × 15 columns

In [78]: `store_category.head(2)`

Out[78]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2019	US-2019-103800	77095
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2019	US-2019-112326	60540

◀ ▶

In [79]: `Store.head(1)`

Out[79]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2019	US-2019-103800	V F X

◀ ▶

In [80]: `store_reverse_df.head(1)`

Out[80]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2022	CA-2022-143500

◀ ▶

In [81]: `store_numerical.head(1)`

Out[81]:

	Discount	Profit	Quantity	Sales
0	0.2	5.5512	2	16.448

```
In [82]: store_category.head(1)
```

```
Out[82]:
```

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2019	US-2019-103800	77095 VFX



```
In [ ]:
```