

# Pandas

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

```
In [2]: pd.__version__
```

```
Out[2]: '2.2.2'
```

```
In [17]: store=pd.read_csv(r'D:\New folder\Sample - Superstore_Orders.csv')
```

```
In [4]: store
```

Out[4]:

	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-2020-10380C
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-2023-15672C
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-2023-14350C

10194 rows × 19 columns

```
In [5]: id(store)
```

```
Out[5]: 2705173118512
```

```
In [6]: len(store)
```

```
Out[6]: 10194
```

```
In [7]: store.columns
```

```
Out[7]: 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 [9]: len(store.columns)
```

```
Out[9]: 19
```

```
In [10]: store.shape
```

```
Out[10]: (10194, 19)
```

```
In [11]: store
```

Out[11]:

	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-2020-10380C
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-2023-15672C
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-2023-14350C

10194 rows × 19 columns

In [12]: `store.dtypes`

```
Out[12]: 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 [13]: `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 [14]: `type(store)`

Out[14]: pandas.core.frame.DataFrame

In [15]: store.isnull()

Out[15]:

	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 [16]: store.isnull().sum()

```
Out[16]: 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 [18]: store\_1=pd.read\_csv(r'D:\New folder\Sample - Superstore\_Orders.csv')

In [19]: `store_1.isnull()`

Out[19]:

	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 [21]: `store_1.isnull().sum()`

Out[21]:

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 [22]: `store_1.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 [23]: `store.columns`

```
Out[23]: 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 [24]: `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 [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['Category']`

Out[27]:

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 [29]: `store[['Customer Name', 'City']]`

Out[29]:

	Customer Name	City
0	Darren Powers	Houston
1	Phillina Ober	Naperville
2	Phillina Ober	Naperville
3	Phillina Ober	Naperville
4	Mick Brown	Philadelphia
...	...	...
10189	Patrick O'Donnell	New York City
10190	Erica Bern	Fairfield
10191	Jill Matthias	Loveland
10192	Patrick O'Donnell	New York City
10193	Harry Olson	Charlottetown

10194 rows × 2 columns

In [30]:

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

In [31]:

```
store_categorical
```

Out[31]:

	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-2020-10380C
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-2023-15672C
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-2023-14350C

10194 rows × 15 columns

In [33]: `store_categorical.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10194 entries, 0 to 10193
Data columns (total 15 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  
dtypes: object(15)
memory usage: 1.2+ MB
```

In [34]: `store_numerical = store[['Discount', 'Profit', 'Quantity', 'Sales']]  
store_numerical`

Out[34]:

	<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 [35]: `store.head()`

Out[35]:

	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 [36]: `store.tail()`

Out[36]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	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



In [37]: store[:]

Out[37]:

	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-2020-10380C
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-2023-15672C
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-2023-14350C

10194 rows × 19 columns

In [38]: `store[0:10]`

	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
5	Furniture	Henderson	United States	Maria Etezadi	Global	06-01-2020	US-2020-167199	42420
6	Office Supplies	Henderson	United States	Maria Etezadi	Rogers	06-01-2020	US-2020-167199	42420
7	Office Supplies	Athens	United States	Jack O'Briant	Dixon	06-01-2020	US-2020-106054	30605
8	Office Supplies	Henderson	United States	Maria Etezadi	Ibico	06-01-2020	US-2020-167199	42420

Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
9	Office Supplies	Henderson	United States	Maria Etezadi	Alliance	06-01-2020	US-2020-167199

In [39]: `store[0:10:2]`

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
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	US-2020-112326
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	US-2020-141817
6	Office Supplies	Henderson	United States	Maria Etezadi	Rogers	06-01-2020	US-2020-167199
8	Office Supplies	Henderson	United States	Maria Etezadi	Ibico	06-01-2020	US-2020-167199



In [40]: `store[:::-1]`

Out[40]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	CA-2023-143500
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	US-2023-143259
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	US-2023-156720
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	US-2023-115427
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	US-2023-143259
...	...	...	...	...	...	...	...
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	US-2020-141817
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	04-01-2020	US-2020-112326
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	US-2020-112326
1	Office Supplies	Naperville	United States	Phillina Ober	GBC	04-01-2020	US-2020-112326
0	Office Supplies	Houston	United States	Darren Powers	Message Book	03-01-2020	US-2020-103800

10194 rows × 19 columns

In [41]: store[6:7]

Out[41]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code
6	Office Supplies	Henderson	United States	Maria Etezadi	Rogers	06-01-2020	US-2020-167199	42420



In [ ]: