

```
In [2]: !pip install panda
```

```
Collecting panda
  Downloading panda-0.3.1.tar.gz (5.8 kB)
  Preparing metadata (setup.py): started
  Preparing metadata (setup.py): finished with status 'done'
Requirement already satisfied: setuptools in c:\users\appala nithin\anaconda3\lib\site-packages (from panda) (69.5.1)
Requirement already satisfied: requests in c:\users\appala nithin\anaconda3\lib\site-packages (from panda) (2.32.2)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\appala nithin\anaconda3\lib\site-packages (from requests->panda) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in c:\users\appala nithin\anaconda3\lib\site-packages (from requests->panda) (3.7)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\appala nithin\anaconda3\lib\site-packages (from requests->panda) (2.2.2)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\appala nithin\anaconda3\lib\site-packages (from requests->panda) (2024.7.4)
Building wheels for collected packages: panda
  Building wheel for panda (setup.py): started
  Building wheel for panda (setup.py): finished with status 'done'
  Created wheel for panda: filename=panda-0.3.1-py3-none-any.whl size=7246 sha256=0c9260347f9e9132977e9e2ce495d3824deedb63c3f0145bd7170b67ff1c6ee8
  Stored in directory: c:\users\appala nithin\appdata\local\pip\cache\wheels\98\41\5b\6ca54e0b6a35e1b7248c12f56fcb753dfb7717fefaa0fb45f5
Successfully built panda
Installing collected packages: panda
Successfully installed panda-0.3.1
Note: you may need to restart the kernel to use updated packages.
```

```
In [4]: import pandas as pd
```

```
In [6]: pd.__version__
```

```
Out[6]: '2.2.2'
```

```
In [12]: store=pd.read_csv(r'C:\Users\Appala nithin\OneDrive\Pictures\Documents\NARESH-IT\21
```

```
In [14]: store
```

Out[14]:

	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 [18]: `id(store)#address of the location`

Out[18]: 1952924650768

In [20]: `len(store) #no.of rows from the data set`

Out[20]: 10194

In [22]: `store.columns #columns of the data set`

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 [24]: `len(store.columns)`

Out[24]: 19

In [26]: `store.shape`

Out[26]: (10194, 19)

In [28]: `store`

Out[28]:

	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

store.dtypes

In [30]: `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 [32]: `type(store)`

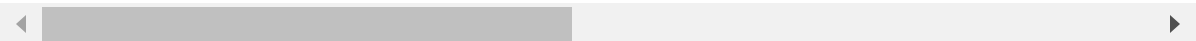
Out[32]: `pandas.core.frame.DataFrame`

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

Out[34]:

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

Out[38]:

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

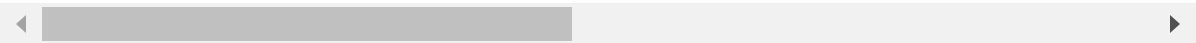
In [67]: `store_1=pd.read_csv(r'C:\Users\AppData\Local\Microsoft\OneDrive\Pictures\Documents\NARESH-IT\`

In [69]: `store_1.isnull()`

Out[69]:

	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 [71]: store\_1.isnull().sum()

Out[71]: 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 [73]: 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 [75]: store.columns
```

```

Out[75]: 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 [79]: #categorical & numerical data
        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 [81]: store.columns
```

```
Out[81]: 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 [83]: store['Category']
```

```
Out[83]: 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 [85]: store[['Customer Name', 'City']]
```

Out[85]:

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

```
In [91]: store_categorical
```

Out[91]:

	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 × 15 columns

In [93]: `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 [99]: `store_numerical=store[['Discount','Profit','Quantity','Sales']]`  
`store_numerical`

Out[99]:

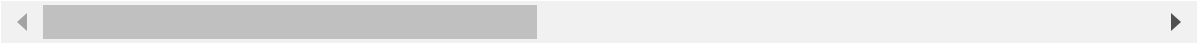
	Discount	Profit	Quantity	Sales
0	0.2	5.5512	2	16.448
1	0.8	-5.4870	2	3.540
2	0.2	4.2717	3	11.784
3	0.2	-64.7748	3	272.736
4	0.2	4.8840	3	19.536
...	...	...	...	...
10189	0.2	19.7910	3	52.776
10190	0.2	6.4750	2	20.720
10191	0.2	-0.6048	3	3.024
10192	0.0	2.7279	7	90.930
10193	0.2	-0.6048	3	3.024

10194 rows × 4 columns

In [101... `store.head()` #top 5 rows

Out[101...

	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	77015
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 [103...

```
store.tail() #bottom 5 records
```

Out[103...

	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 [105...

```
store[:]
```

Out[105...

	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 [107...

`store[0:10]`



Out[107...

	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	77015
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	42420

In [109...

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

Out[109...

	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	77015
2	Office Supplies	Naperville	United States	Phillina Ober	Avery	04-01-2020	US-2020-112326	60540
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	05-01-2020	US-2020-141817	19143
6	Office Supplies	Henderson	United States	Maria Etezadi	Rogers	06-01-2020	US-2020-167199	42420
8	Office Supplies	Henderson	United States	Maria Etezadi	Ibico	06-01-2020	US-2020-167199	42420

In [111...

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

Out[111...

	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 [113...]

store[6:7]

Out[113...]

	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 [123...]

store[6:7]

Out[123...]

	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 [ ]: