

**** 21st ****

PANDAS -- framework to handle data frame

excel sheet if any values are missing (you plot -- graph dont has meaning) pandas will handle missing data

```
In [161]: import pandas as pd
```

```
In [165]: pd.__version__
```

```
Out[165]: '2.2.2'
```

```
In [167]: store = pd.read_csv(r'C:\DataScience Classnotes\21st August class\Sample - Supersto
```

```
In [169]: store.shape
```

```
Out[169]: (10194, 19)
```

```
In [171]: store.columns
```

```
Out[171]: 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]: len(store)
```

```
Out[17]: 10194
```

```
In [21]: len(store.columns)
```

```
Out[21]: 19
```

```
In [95]: type(store)
```

```
Out[95]: pandas.core.frame.DataFrame
```

```
In [33]: store.isnull().sum()
```

```
Out[33]: 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 [25]: store.isnull().sum()
```

```
Out[25]: 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 [27]: store.dtypes
```

```
Out[27]: 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 [35]: 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 [37]: id(store)
```

```
Out[37]: 1930650187968
```

```
In [39]: store.columns
```

```
Out[39]: 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 [41]: store['City']
```

```
Out[41]: 0          Houston
          1      Naperville
          2      Naperville
          3      Naperville
          4      Philadelphia
          ...
10189    New York City
10190      Fairfield
10191      Loveland
10192    New York City
10193    Charlottetown
Name: City, Length: 10194, dtype: object
```

```
In [43]: store[['Category', 'City', 'Country/Region']]
```

```
Out[43]:
```

	Category	City	Country/Region
0	Office Supplies	Houston	United States
1	Office Supplies	Naperville	United States
2	Office Supplies	Naperville	United States
3	Office Supplies	Naperville	United States
4	Office Supplies	Philadelphia	United States
...
10189	Office Supplies	New York City	United States
10190	Office Supplies	Fairfield	United States
10191	Office Supplies	Loveland	United States
10192	Technology	New York City	United States
10193	Office Supplies	Charlottetown	Canada

10194 rows × 3 columns

```
In [45]: store[['Category', 'City', 'Country/Region', 'Customer Name', 'Manufacturer']]
```

Out[45]:

	Category	City	Country/Region	Customer Name	Manufacturer
0	Office Supplies	Houston	United States	Darren Powers	Message Book
1	Office Supplies	Naperville	United States	Phillina Ober	GBC
2	Office Supplies	Naperville	United States	Phillina Ober	Avery
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery
...
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC
10191	Office Supplies	Loveland	United States	Jill Matthias	Other
10192	Technology	New York City	United States	Patrick O'Donnell	Other
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones

10194 rows × 5 columns

```
In [47]: store_text = store[['Category', 'City', 'Country/Region', 'Customer Name', 'Manufac
      'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',
      'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category']]
```

```
In [49]: store_text
```

Out[49]:

	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 [51]: `store.columns`

```
Out[51]: 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 [53]: `store_text.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 [55]: `store_number = store[['Discount', 'Profit', 'Quantity', 'Sales']]`In [57]: `store_number`

Out[57]:

	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 [59]: `store.columns`

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

Out[61]: 19

In [63]: `store_text.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'], dtype='object')

In [65]: `len(store_text.columns)`

Out[65]: 15

In [71]: `store_number.columns`

Out[71]: Index(['Discount', 'Profit', 'Quantity', 'Sales'], dtype='object')

In [73]: `len(store_number.columns)`

Out[73]: 4

```
In [75]: print(len(store.columns))
print(len(store_text.columns))
print(len(store_number.columns))

19
15
4
```

```
In [77]: store.head()
```

Out[77]:

	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 [79]: store.tail()
```

Out[79]:

	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 [81]:

```
store[:]
```

Out[81]:

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

Out[85]:

	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 [87]: `store[0:100:10]`

Out[87]:

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Product ID
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
50	Office Supplies	Jonesboro	United States	Hunter Lopez	Newell	20-01-2020	US-2020-147627	72
60	Furniture	Miami	United States	Tom Boeckenhauer	Ultra	21-01-2020	US-2020-110422	33
70	Furniture	Alexandria	United States	Shirley Daniels	GE	26-01-2020	US-2020-155502	22
80	Office Supplies	New York City	United States	Jasper Cacioppo	Pressboard	28-01-2020	US-2020-100328	10
90	Office Supplies	Escondido	United States	Mark Haberlin	GBC	04-02-2020	US-2020-104808	92

In [89]: store[10:11]

Out[89]:

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

In [91]: store[:4]

Out[91]:

	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

In [93]: store_text[:4:-1]

Out[93]:

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

10189 rows × 15 columns

```
In [117... store2 = pd.read_csv(r'C:\DataScience Classnotes\21st August class\Sample - Superst
```

```
In [119... store2
```

Out[119...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order
0	NaN	Houston	United States	Darren Powers	Message Book	3/1/2020	201038
1	NaN	Naperville	United States	Phillina Ober	GBC	4/1/2020	201123
2	NaN	Naperville	United States	Phillina Ober	Avery	4/1/2020	201123
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	4/1/2020	201123
4	Office Supplies	Philadelphia	United States	Mick Brown	Avery	5/1/2020	201418
...
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	201432
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	201154
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	201567
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	201432
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	201432

10194 rows × 19 columns

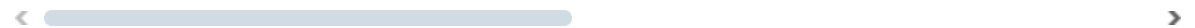
In [121...

store2.isnull()

Out[121...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code	F
0	True	False	False	False	False	False	False	False	
1	True	False	False	False	False	False	False	False	
2	True	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 [123...

store2.isnull().sum()

Out[123...

Category	3
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 [127... store2 = pd.read_csv(r'C:\DataScience Classnotes\21st August class\Sample - Superst
```

```
In [129... store2
```

Out[129...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID
0	NaN	Houston	United States	Darren Powers	Message Book	3/1/2020	201038
1	NaN	Naperville	United States	Phillina Ober	GBC	4/1/2020	201123
2	NaN	Naperville	NaN	Phillina Ober	Avery	4/1/2020	201123
3	Office Supplies	Naperville	United States	Phillina Ober	SAFCO	NaN	201123
4	Office Supplies	Philadelphia	United States	NaN	Avery	NaN	201418
...
10189	Office Supplies	New York City	United States	Patrick O'Donnell	Wilson Jones	30-12-2023	201432
10190	Office Supplies	Fairfield	United States	Erica Bern	GBC	30-12-2023	201154
10191	Office Supplies	Loveland	United States	Jill Matthias	Other	30-12-2023	201567
10192	Technology	New York City	United States	Patrick O'Donnell	Other	30-12-2023	201432
10193	Office Supplies	Charlottetown	Canada	Harry Olson	Wilson Jones	30-12-2023	201432

10194 rows × 19 columns

In [131...

store2.isnull()

Out[131...

	Category	City	Country/Region	Customer Name	Manufacturer	Order Date	Order ID	Postal Code	F
0	True	False	False	False	False	False	False	False	
1	True	False	False	False	False	False	False	False	
2	True	False	True	False	False	False	False	False	
3	False	False	False	False	False	True	False	False	
4	False	False	False	True	False	True	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 [135...

store2.isnull().sum()

Out[135...

Category	3
City	0
Country/Region	3
Customer Name	1
Manufacturer	0
Order Date	3
Order ID	1
Postal Code	0
Product Name	2
Region	0
Segment	0
Ship Date	4
Ship Mode	0
State/Province	3
Sub-Category	0
Discount	0
Profit	0
Quantity	0
Sales	0
dtype:	int64

In [139...

store2.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10194 entries, 0 to 10193
Data columns (total 19 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Category              10191 non-null  object
1   City                  10194 non-null  object
2   Country/Region        10191 non-null  object
3   Customer Name         10193 non-null  object
4   Manufacturer           10194 non-null  object
5   Order Date            10191 non-null  object
6   Order ID              10193 non-null  object
7   Postal Code           10194 non-null  object
8   Product Name          10192 non-null  object
9   Region                10194 non-null  object
10  Segment               10194 non-null  object
11  Ship Date             10190 non-null  object
12  Ship Mode             10194 non-null  object
13  State/Province        10191 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 [141...

store.columns

Out[141...

```
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 [143...

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 [157... store_catagorical = 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']]
store_catagorical
```

Out[157...

	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 [147...

store_catagorical.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 [149...

store_numerical = store[['Discount', 'Profit', 'Quantity', 'Sales']]
store_numerical

Out[149...

	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 [153...

```
store_numerical.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10194 entries, 0 to 10193
Data columns (total 4 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Discount    10194 non-null  float64
1   Profit      10194 non-null  float64
2   Quantity    10194 non-null  int64
3   Sales       10194 non-null  float64
dtypes: float64(3), int64(1)
memory usage: 318.7 KB
```

In [173...

```
store.head()
```

Out[173...

	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 [175...

```
store.tail()
```

Out[175...

	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 [179...

```
store[::1]
```

Out[179...

	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 [181...

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

Out[181...

	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	42420

In [185...

store[:10:2]

Out[185...

	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
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 []: