### s3eeeznvu

#### March 20, 2025

```
[1]:
     import pandas as pd
[2]: pd. version
[2]: '2.2.2'
     store = pd.read_csv(r"C:\Users\Dhanwantari Devre\Sample_orders.csv")
[4]:
    store
[4]:
                   Category
                                       City Country/Region
                                                                  Customer Name
            Office Supplies
                                             United States
                                                                  Darren Powers
     0
                                    Houston
     1
            Office Supplies
                                 Naperville
                                              United States
                                                                  Phillina Ober
     2
            Office Supplies
                                 Naperville
                                             United States
                                                                 Phillina Ober
     3
                                                                 Phillina Ober
            Office Supplies
                                 Naperville
                                             United States
     4
            Office Supplies
                               Philadelphia
                                              United States
                                                                     Mick Brown
                                                             Patrick O'Donnell
            Office Supplies
                              New York City
                                              United States
     10189
            Office Supplies
                                  Fairfield
                                                                     Erica Bern
     10190
                                             United States
     10191
            Office Supplies
                                   Loveland
                                             United States
                                                                  Jill Matthias
                 Technology
     10192
                              New York City
                                             United States
                                                             Patrick O'Donnell
     10193
            Office Supplies
                              Charlottetown
                                                     Canada
                                                                    Harry Olson
            Manufacturer
                                              Order ID Postal Code
                           Order Date
     0
            Message Book
                           03-01-2020
                                       US-2020-103800
                                                             77015
     1
                     GBC
                           04-01-2020
                                       US-2020-112326
                                                              60540
     2
                   Avery
                           04-01-2020
                                       US-2020-112326
                                                              60540
     3
                   SAFCO
                           04-01-2020
                                       US-2020-112326
                                                              60540
     4
                   Avery
                           05-01-2020
                                       US-2020-141817
                                                             19143
            Wilson Jones
     10189
                           30-12-2023
                                       US-2023-143259
                                                             10009
                     GBC
     10190
                           30-12-2023
                                       US-2023-115427
                                                             94533
     10191
                           30-12-2023
                   Other
                                       US-2023-156720
                                                             80538
     10192
                   Other
                           30-12-2023
                                       US-2023-143259
                                                              10009
     10193 Wilson Jones
                           30-12-2023
                                       CA-2023-143500
                                                               COA
```

Product Name

Region \

```
0
       Message Book, Wirebound, Four 5 1/2" X 4" Form... Central
1
              GBC Standard Plastic Binding Systems Combs
                                                             Central
2
                                                  Avery 508
                                                              Central
3
                            SAFCO Boltless Steel Shelving
                                                              Central
4
       Avery Hi-Liter EverBold Pen Style Fluorescent ...
                                                               East
10189
                     Wilson Jones Legal Size Ring Binders
                                                                 East
10190
                                        GBC Binding covers
                                                                 West
10191
                                       Bagged Rubber Bands
                                                                 West
                                 Gear Head AU3700S Headset
10192
                                                                 East
10193
                              Wilson Jones Impact Binders
                                                                 East
           Segment
                      Ship Date
                                                         State/Province
                                       Ship Mode
0
          Consumer
                     07-01-2020
                                  Standard Class
                                                                   Texas
1
       Home Office
                                  Standard Class
                     08-01-2020
                                                                Illinois
2
       Home Office
                     08-01-2020
                                  Standard Class
                                                                Illinois
3
       Home Office
                     08-01-2020
                                  Standard Class
                                                                Illinois
4
          Consumer
                     12-01-2020
                                  Standard Class
                                                           Pennsylvania
10189
          Consumer
                     03-01-2024
                                  Standard Class
                                                                New York
10190
         Corporate
                     03-01-2024
                                  Standard Class
                                                              California
          Consumer
                                  Standard Class
10191
                     03-01-2024
                                                                Colorado
10192
          Consumer
                     03-01-2024
                                  Standard Class
                                                                New York
10193
          Consumer
                     03-01-2024
                                  Standard Class
                                                  Prince Edward Island
      Sub-Category
                     Discount
                                Profit
                                         Quantity
                                                      Sales
0
             Paper
                          0.2
                                 5.5512
                                                 2
                                                     16.448
                                                      3.540
1
           Binders
                          0.8
                               -5.4870
                                                 2
2
            Labels
                          0.2
                                 4.2717
                                                 3
                                                     11.784
3
                          0.2 - 64.7748
                                                 3
                                                    272.736
           Storage
4
                Art
                          0.2
                                 4.8840
                                                 3
                                                     19.536
10189
           Binders
                          0.2 19.7910
                                                 3
                                                     52.776
10190
           Binders
                          0.2
                                6.4750
                                                 2
                                                     20.720
         Fasteners
                          0.2
                                                 3
                                                      3.024
10191
                               -0.6048
                                                 7
10192
            Phones
                          0.0
                                 2.7279
                                                     90.930
10193
           Binders
                          0.2
                               -0.6048
                                                 3
                                                      3.024
```

[10194 rows x 19 columns]

[5]: id(store)

[5]: 2385804921744

6]: len(store)

[6]: 10194

```
[7]: store.shape
 [7]: (10194, 19)
      store.columns
 [8]: 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')
      len(store.columns)
 [9]: 19
[10]:
      store.isnull()
[10]:
             Category
                         City
                               Country/Region Customer Name
                                                                Manufacturer \
      0
                False False
                                         False
                                                         False
                                                                       False
      1
                False
                       False
                                         False
                                                                       False
                                                         False
      2
                False False
                                         False
                                                         False
                                                                       False
      3
                False False
                                         False
                                                         False
                                                                       False
      4
                False False
                                         False
                                                         False
                                                                       False
                                                                       False
      10189
                False False
                                         False
                                                         False
      10190
                False False
                                         False
                                                         False
                                                                       False
      10191
                False False
                                         False
                                                         False
                                                                       False
      10192
                False False
                                         False
                                                         False
                                                                       False
      10193
                False False
                                         False
                                                         False
                                                                       False
                                                                 Region Segment
             Order Date
                          Order ID Postal Code
                                                 Product Name
      0
                   False
                             False
                                           False
                                                          False
                                                                  False
                                                                            False
      1
                   False
                             False
                                           False
                                                          False
                                                                  False
                                                                           False
      2
                   False
                             False
                                           False
                                                          False
                                                                  False
                                                                           False
      3
                   False
                             False
                                           False
                                                          False
                                                                  False
                                                                           False
      4
                   False
                             False
                                           False
                                                          False
                                                                  False
                                                                           False
      10189
                  False
                             False
                                           False
                                                          False
                                                                  False
                                                                           False
                  False
                             False
                                           False
                                                          False
                                                                  False
                                                                           False
      10190
      10191
                  False
                             False
                                           False
                                                          False
                                                                  False
                                                                           False
      10192
                   False
                             False
                                           False
                                                          False
                                                                  False
                                                                           False
      10193
                  False
                             False
                                           False
                                                          False
                                                                  False
                                                                           False
             Ship Date
                         Ship Mode
                                    State/Province
                                                     Sub-Category
                                                                    Discount
                                                                               Profit
      0
                 False
                             False
                                              False
                                                             False
                                                                       False
                                                                                False
      1
                  False
                             False
                                              False
                                                             False
                                                                       False
                                                                                False
```

2	False	False		False		False		False	False
3	False	False		False		False		False	False
4	False	False		False		False		False	False
•••	•••	•••	•••		•••	•••	•••		
10189	False	False		False		False		False	False
10190	False	False		False		False		False	False
10191	False	False		False		False		False	False
10192	False	False		False		False		False	False
10193	False	False		False		False		False	False
	Quantity Sa	ales							

	Quantity	Sales
0	False	False
1	False	False
2	False	False
3	False	False
4	False	False
•••		
 10189	 False	False
 10189 10190	 False False	False False
10190	False	False

[10194 rows x 19 columns]

## [11]: store.isnull().sum()

[11]: 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

## [12]: store[:]

[12]:		Catego	ory	City	Country	/Region	Customer Name	\
	0	Office Suppl:	ies H	louston	United	States	Darren Powers	
	1	Office Suppl:	ies Nape	rville	United	States	Phillina Ober	
	2	Office Suppl:	ies Nape	rville	United	States	Phillina Ober	
	3	Office Suppl:	ies Nape	rville	United	States	Phillina Ober	
	4	Office Suppl:	ies Philad	lelphia	United	States	Mick Brown	
	•••		••		•••		•••	
	10189	Office Suppl:	ies New Yor	k City	United	States	Patrick O'Donnell	
	10190	Office Suppl:	ies Fai	rfield	United	States	Erica Bern	
	10191	Office Suppl:	ies Lo	veland	United	States	Jill Matthias	
	10192	Technolo	ogy New Yor	k City	United	States	Patrick O'Donnell	
	10193	Office Suppl:	ies Charlot	tetown		Canada	Harry Olson	
		Manufacturer	Order Date	)	Order 3	ID Posta	l Code \	
	0	Message Book	03-01-2020	US-20	020-10380	00	77015	
	1	GBC	04-01-2020	US-20	020-11232	26	60540	
	2	Avery	04-01-2020	US-20	020-11232	26	60540	
	3	SAFCO	04-01-2020	US-20	020-11232	26	60540	
	4	Avery	05-01-2020	US-20	020-1418:	17	19143	
	•••	•••	•••			•••		
	10189	Wilson Jones	30-12-2023	US-20	023-1432	59	10009	
	10190	GBC	30-12-2023	US-20	023-11542	27	94533	
	10191	Other	30-12-2023	US-20	023-15672	20	80538	
	10192	Other	30-12-2023	US-20	023-1432	59	10009	
	10193	Wilson Jones	30-12-2023	CA-20	023-14350	00	COA	
					Pa	roduct Na	ame Region \	
	0	Message Book	, Wirebound,	Four !	5 1/2" X	4" Form.	Central	
	1	GBC St	andard Plas	tic Bi	nding Sys	stems Co	mbs Central	
	2					Avery	508 Central	
	3		SAFO	O Bolt	less Ste	el Shelv:	ing Central	
	4	Avery Hi-Lite	er EverBold	Pen St	yle Fluo	rescent .	East	
	•••					•••	•••	
	10189		Wilson Jone	s Legal	l Size R	ing Bind	ers East	
	10190				GBC Bind	ding cov	ers West	
	10191			]	Bagged Ru	ubber Ba	nds West	
	10192			Gear H	ead AU370	OOS Head	set East	
	10193		Wi	lson Jo	ones Impa	act Bind	ers East	
		${\tt Segment}$	Ship Date	S	Ship Mode	Э	State/Province \	
	0	Consumer	07-01-2020	Standa	ard Class	3	Texas	
	1	Home Office	08-01-2020		ard Class		Illinois	
	2	Home Office	08-01-2020	Standa	ard Class	3	Illinois	
	3	Home Office	08-01-2020	Standa	ard Class	3	Illinois	
	4	Consumer	12-01-2020	Standa	ard Class	3	Pennsylvania	

•••	***	***	•••			•••	
10189	Consumer	03-01-2024	Standaı	rd Class		Ne	w York
10190	Corporate	03-01-2024	Standaı	rd Class		Cali	fornia
10191	Consumer	03-01-2024	Standaı	rd Class		Co	lorado
10192	Consumer	03-01-2024	Standaı	rd Class		Ne	w York
10193	Consumer	03-01-2024	Standar	rd Class	Prince E	Edward	Island
	Sub-Category	Discount	Profit	Quantity	Sales	2	
^	0 0						
0	Paper	0.2		2	16.448		
1	Binders	0.8	-5.4870	2	3.540	)	
2	Labels	0.2	4.2717	3	11.784	1	
3	Storage	0.2 -	64.7748	3	272.736	3	
4	Art	0.2	4.8840	3	19.536	3	
•••	•••		•••				
10189	Binders	0.2	19.7910	3	52.776	3	
10190	Binders	0.2	6.4750	2	20.720	)	
10191	Fasteners	0.2	-0.6048	3	3.024	1	
10192	Phones	0.0	2.7279	7	90.930	)	
10193	Binders	0.2	-0.6048	3	3.024	1	

[10194 rows x 19 columns]

## [13]: store[0:10]

[13]:		Cate	gory	City	Country/	Region	Customer Name	Manufacturer	\
	0	Office Supp	-	Houston	United S	States	Darren Powers	Message Book	
	1	Office Supp	lies	Naperville	United S	States	Phillina Ober	GBC	
	2	Office Supp	lies	Naperville	United S	States	Phillina Ober	Avery	
	3	Office Supp	lies	Naperville	United S	States	Phillina Ober	SAFCO	
	4	Office Supp	lies	Philadelphia	United S	States	Mick Brown	Avery	
	5	Furni	ture	Henderson	United S	States	Maria Etezadi	Global	
	6	Office Supp	lies	Henderson	United S	States	Maria Etezadi	Rogers	
	7	Office Supp	lies	Athens	United S	States	Jack O'Briant	Dixon	
	8	Office Supp	lies	Henderson	United S	States	Maria Etezadi	Ibico	
	9	Office Supp	lies	Henderson	United S	States	Maria Etezadi	Alliance	
		Order Date		Order ID Pos	stal Code	\			
	0	03-01-2020	US-2	020-103800	77015				
	1	04-01-2020	US-2	020-112326	60540				
	2	04-01-2020	US-2	020-112326	60540				
	3	04-01-2020	US-2	020-112326	60540				
	4	05-01-2020	US-2	020-141817	19143				
	5	06-01-2020	US-2	020-167199	42420				
	6	06-01-2020	US-2	020-167199	42420				
	7	06-01-2020	US-2	020-106054	30605				
	8	06-01-2020	US-2	020-167199	42420				
	9	06-01-2020	US-2	020-167199	42420				

```
Message Book, Wirebound, Four 5 1/2" X 4" Form... Central
                                                                          Consumer
      1
                 GBC Standard Plastic Binding Systems Combs
                                                               Central
                                                                         Home Office
      2
                                                    Avery 508
                                                               Central
                                                                         Home Office
                              SAFCO Boltless Steel Shelving
      3
                                                               Central
                                                                         Home Office
      4
         Avery Hi-Liter EverBold Pen Style Fluorescent ...
                                                                East
                                                                          Consumer
      5
                    Global Deluxe High-Back Manager's Chair
                                                                 South
                                                                         Home Office
      6
                    Rogers Handheld Barrel Pencil Sharpener
                                                                         Home Office
                                                                 South
      7
         Dixon Prang Watercolor Pencils, 10-Color Set w...
                                                               South
                                                                         Corporate
                        Ibico Hi-Tech Manual Binding System
      8
                                                                 South
                                                                         Home Office
      9
                  Alliance Super-Size Bands, Assorted Sizes
                                                                 South
                                                                         Home Office
          Ship Date
                           Ship Mode State/Province Sub-Category
                                                                    Discount
                                                                                 Profit
         07-01-2020
                      Standard Class
                                               Texas
                                                             Paper
                                                                          0.2
                                                                                 5.5512
      0
      1
         08-01-2020
                      Standard Class
                                            Illinois
                                                           Binders
                                                                          0.8
                                                                                -5.4870
         08-01-2020
                      Standard Class
                                            Illinois
                                                                          0.2
      2
                                                            Labels
                                                                                 4.2717
      3
         08-01-2020
                      Standard Class
                                            Illinois
                                                           Storage
                                                                          0.2
                                                                               -64.7748
         12-01-2020
                      Standard Class
                                        Pennsylvania
                                                                          0.2
                                                                                 4.8840
      4
                                                               Art
                      Standard Class
      5
         10-01-2020
                                            Kentucky
                                                            Chairs
                                                                          0.0
                                                                               746.4078
      6
         10-01-2020
                      Standard Class
                                            Kentucky
                                                               Art
                                                                          0.0
                                                                                 1.4796
      7
         07-01-2020
                         First Class
                                             Georgia
                                                               Art
                                                                          0.0
                                                                                 5.2398
         10-01-2020
                      Standard Class
                                            Kentucky
                                                           Binders
                                                                          0.0
                                                                               274.4910
                      Standard Class
         10-01-2020
                                            Kentucky
                                                         Fasteners
                                                                          0.0
                                                                                 0.3112
         Quantity
                       Sales
      0
                 2
                      16.448
      1
                 2
                       3.540
      2
                 3
                      11.784
      3
                 3
                     272.736
      4
                 3
                      19.536
      5
                 9
                    2573.820
                 2
      6
                       5.480
      7
                 3
                      12.780
      8
                 2
                     609.980
      9
                 4
                      31.120
[14]: store[0:20:5]
[14]:
                                                            Customer Name \
                  Category
                                   City Country/Region
      0
          Office Supplies
                               Houston
                                         United States
                                                            Darren Powers
      5
                 Furniture
                             Henderson
                                         United States
                                                            Maria Etezadi
          Office Supplies
                             Henderson
                                         United States
                                                            Maria Etezadi
      10
      15
          Office Supplies
                            Huntsville
                                         United States
                                                         Vivek Sundaresam
          Manufacturer
                         Order Date
                                            Order ID Postal Code \
      0
          Message Book
                         03-01-2020
                                      US-2020-103800
                                                            77015
```

Product Name

Region

Segment \

```
10
            Southworth
                        06-01-2020
                                    US-2020-167199
                                                          42420
      15
                  Acco
                        07-01-2020
                                    US-2020-105417
                                                          77340
                                                Product Name
                                                                            Segment \
                                                               Region
          Message Book, Wirebound, Four 5 1/2" X 4" Form...
      0
                                                            Central
                                                                         Consumer
      5
                    Global Deluxe High-Back Manager's Chair
                                                                       Home Office
                                                                South
            Southworth 25% Cotton Granite Paper & Envelopes
      10
                                                                South
                                                                       Home Office
         Acco Four Pocket Poly Ring Binder with Label H... Central
      15
                                                                         Consumer
           Ship Date
                           Ship Mode State/Province Sub-Category
      0
          07-01-2020
                      Standard Class
                                               Texas
                                                            Paper
                                                                         0.2
      5
          10-01-2020
                      Standard Class
                                            Kentucky
                                                           Chairs
                                                                         0.0
      10
          10-01-2020
                      Standard Class
                                            Kentucky
                                                            Paper
                                                                         0.0
          12-01-2020 Standard Class
                                               Texas
                                                          Binders
                                                                         0.8
                    Quantity
            Profit
                                  Sales
      0
            5.5512
                           2
                                16.448
      5
          746.4078
                           9
                              2573.820
            3.0084
                                  6.540
      10
                           1
      15
         -18.2525
                           7
                                10.430
[15]: store.head()
[15]:
                Category
                                  City Country/Region
                                                        Customer Name
                                                                       Manufacturer \
      O 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
         Order Date
                           Order ID Postal Code
      0 03-01-2020
                    US-2020-103800
                                           77015
      1 04-01-2020
                     US-2020-112326
                                           60540
                                           60540
      2 04-01-2020
                     US-2020-112326
      3 04-01-2020
                     US-2020-112326
                                           60540
      4 05-01-2020 US-2020-141817
                                           19143
                                               Product Name
                                                                           Segment \
                                                              Region
         Message Book, Wirebound, Four 5 1/2" X 4" Form... Central
                                                                       Consumer
                GBC Standard Plastic Binding Systems Combs
                                                             Central
                                                                      Home Office
      1
      2
                                                  Avery 508
                                                             Central
                                                                      Home Office
      3
                             SAFCO Boltless Steel Shelving
                                                             Central
                                                                      Home Office
        Avery Hi-Liter EverBold Pen Style Fluorescent ...
                                                              East
                                                                       Consumer
                                                                  Discount
          Ship Date
                          Ship Mode State/Province Sub-Category
                                                                              Profit
      0 07-01-2020 Standard Class
                                              Texas
                                                           Paper
                                                                        0.2
                                                                              5.5512
```

US-2020-167199

42420

5

Global

06-01-2020

```
Standard Class
      2 08-01-2020
                                           Illinois
                                                          Labels
                                                                        0.2
                                                                              4.2717
      3 08-01-2020
                     Standard Class
                                           Illinois
                                                         Storage
                                                                        0.2 - 64.7748
         12-01-2020
                     Standard Class
                                       Pennsylvania
                                                             Art
                                                                        0.2
                                                                              4.8840
         Quantity
                     Sales
      0
                2
                    16.448
                2
      1
                     3.540
      2
                3
                    11.784
      3
                3
                   272.736
      4
                3
                    19.536
[16]: store.head(3)
[16]:
                Category
                                City Country/Region
                                                      Customer Name Manufacturer \
      O Office Supplies
                             Houston United States
                                                      Darren Powers
                                                                      Message Book
      1 Office Supplies
                                      United States
                                                      Phillina Ober
                                                                               GBC
                          Naperville
      2 Office Supplies
                          Naperville
                                      United States
                                                      Phillina Ober
                                                                             Avery
         Order Date
                           Order ID Postal Code
      0 03-01-2020
                     US-2020-103800
                                           77015
      1 04-01-2020
                     US-2020-112326
                                           60540
      2 04-01-2020
                     US-2020-112326
                                           60540
                                               Product Name
                                                                           Segment \
                                                               Region
         Message Book, Wirebound, Four 5 1/2" X 4" Form... Central
                                                                        Consumer
                GBC Standard Plastic Binding Systems Combs
                                                                       Home Office
      1
                                                             Central
      2
                                                  Avery 508 Central
                                                                       Home Office
                          Ship Mode State/Province Sub-Category
                                                                   Discount
          Ship Date
                                                                            Profit
         07-01-2020
                     Standard Class
                                              Texas
                                                           Paper
                                                                        0.2 5.5512
      1 08-01-2020
                     Standard Class
                                                                        0.8 -5.4870
                                           Illinois
                                                         Binders
      2 08-01-2020
                                                                        0.2 4.2717
                     Standard Class
                                           Illinois
                                                          Labels
                    Sales
         Quantity
      0
                2
                   16.448
                2
                    3.540
      1
      2
                3
                   11.784
[17]:
     store.tail()
[17]:
                    Category
                                        City Country/Region
                                                                  Customer Name
             Office Supplies
                              New York City
                                             United States
                                                            Patrick O'Donnell
      10189
      10190
             Office Supplies
                                  Fairfield United States
                                                                     Erica Bern
      10191
             Office Supplies
                                    Loveland United States
                                                                  Jill Matthias
      10192
                  Technology
                              New York City
                                              United States
                                                             Patrick O'Donnell
             Office Supplies
      10193
                              Charlottetown
                                                     Canada
                                                                    Harry Olson
```

Illinois

Binders

0.8 - 5.4870

1 08-01-2020

Standard Class

```
10189
             Wilson Jones
                            30-12-2023
                                        US-2023-143259
                                                              10009
      10190
                      GBC
                            30-12-2023
                                        US-2023-115427
                                                              94533
      10191
                    Other
                           30-12-2023
                                        US-2023-156720
                                                              80538
                    Other
      10192
                           30-12-2023
                                        US-2023-143259
                                                              10009
      10193 Wilson Jones
                           30-12-2023
                                        CA-2023-143500
                                                                COA
                                      Product Name Region
                                                              Segment
                                                                        Ship Date \
             Wilson Jones Legal Size Ring Binders
                                                     East
                                                             Consumer
                                                                       03-01-2024
      10189
      10190
                                GBC Binding covers
                                                     West
                                                            Corporate
                                                                       03-01-2024
      10191
                               Bagged Rubber Bands
                                                     West
                                                             Consumer
                                                                       03-01-2024
      10192
                        Gear Head AU3700S Headset
                                                     East
                                                             Consumer
                                                                       03-01-2024
      10193
                      Wilson Jones Impact Binders
                                                     East
                                                             Consumer
                                                                       03-01-2024
                  Ship Mode
                                    State/Province Sub-Category
                                                                 Discount
                                                                             Profit
             Standard Class
                                          New York
                                                        Binders
                                                                       0.2
      10189
                                                                            19.7910
             Standard Class
                                        California
                                                                       0.2
      10190
                                                         Binders
                                                                             6.4750
      10191
             Standard Class
                                          Colorado
                                                      Fasteners
                                                                       0.2 -0.6048
                                                          Phones
      10192
             Standard Class
                                          New York
                                                                       0.0
                                                                             2.7279
             Standard Class Prince Edward Island
      10193
                                                        Binders
                                                                       0.2 -0.6048
             Quantity
                        Sales
      10189
                    3
                       52.776
                       20.720
      10190
      10191
                    3
                        3.024
                       90.930
      10192
                    7
      10193
                    3
                        3.024
[18]:
      store.isna()
[18]:
             Category
                        City
                               Country/Region Customer Name
                                                               Manufacturer \
                False False
                                        False
                                                        False
                                                                      False
      0
      1
                                        False
                                                                      False
                False False
                                                        False
      2
                False False
                                        False
                                                        False
                                                                      False
      3
                False False
                                        False
                                                        False
                                                                      False
      4
                False False
                                                        False
                                        False
                                                                      False
                False False
      10189
                                        False
                                                        False
                                                                      False
                False False
                                        False
                                                        False
                                                                      False
      10190
      10191
                False False
                                        False
                                                        False
                                                                      False
      10192
                False False
                                        False
                                                        False
                                                                      False
      10193
                False False
                                        False
                                                        False
                                                                      False
             Order Date
                         Order ID Postal Code Product Name
                                                                Region Segment \
      0
                  False
                                                                 False
                                                                          False
                             False
                                          False
                                                        False
      1
                  False
                             False
                                          False
                                                        False
                                                                 False
                                                                          False
```

Order ID Postal Code

Manufacturer

Order Date

2	False	e False	False	False	False Fa	alse	
3	False	e False	False	False	False Fa	alse	
4	False	e False	False	False	False Fa	alse	
•••	•••	•••			•		
10189	False	e False	False	False	False Fa	alse	
10190	False	e False	False	False	False Fa	alse	
10191	False	e False	False	False	False Fa	alse	
10192	False	e False	False	False	False Fa	alse	
10193	False	e False	False	False	False Fa	alse	
	Ship Date	Ship Mode	State/Province	Sub-Category	Discount	Profit	\
0	False	False	False	False	e False	False	
1	False	False	False	False	e False	False	
2	False	False	False	False	e False	False	
3	False	False	False	False	e False	False	
4	False	False	False	False	e False	False	
•••	***	•••	•••		•••		
10189	False	False	False	False	e False	False	
10190	False	False	False	False	e False	False	
10191	False	False	False	False	e False	False	
10192	False	False	False	False		False	
10193	False	False	False	False	e False	False	
	${\tt Quantity}$						
0		False					
1		False					
2		False					
3		False					
4	False	False					
10189	False						
10190		False					
10191		False					
10192		False					
10193	False	False					

# 1 Introudce statistical concept in pandas ==>

[10194 rows x 19 columns]

#### [26]: store.describe() [26]: Discount ${\tt Profit}$ Quantity Sales count 10194.000000 10194.000000 10194.000000 10194.000000 0.155385 28.673417 3.791838 228.225854 mean 0.206249 232.465115 2.228317 619.906839 std

```
min
                 0.000000
                           -6599.978000
                                              1.000000
                                                            0.444000
      25%
                 0.000000
                               1.760800
                                              2.000000
                                                           17.220000
      50%
                 0.200000
                               8.690000
                                              3.000000
                                                           53.910000
      75%
                 0.200000
                               29.297925
                                              5.000000
                                                          209.500000
      max
                 0.800000
                            8399.976000
                                             14.000000
                                                        22638.480000
[31]: store.head()
[31]:
                                  City Country/Region
                                                        Customer Name
                                                                        Manufacturer
                Category
         Office Supplies
                               Houston United States
                                                        Darren Powers
                                                                        Message Book
         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
         Order Date
                           Order ID Postal Code
      0 03-01-2020 US-2020-103800
                                           77015
      1 04-01-2020
                     US-2020-112326
                                           60540
      2 04-01-2020
                     US-2020-112326
                                           60540
      3 04-01-2020
                     US-2020-112326
                                           60540
      4 05-01-2020 US-2020-141817
                                           19143
                                               Product Name
                                                              Region
                                                                           Segment \
         Message Book, Wirebound, Four 5 1/2" X 4" Form... Central
                                                                        Consumer
                GBC Standard Plastic Binding Systems Combs
      1
                                                             Central
                                                                      Home Office
      2
                                                  Avery 508
                                                             Central
                                                                      Home Office
                                                                      Home Office
      3
                             SAFCO Boltless Steel Shelving
                                                             Central
         Avery Hi-Liter EverBold Pen Style Fluorescent ...
                                                              East
                                                                        Consumer
                          Ship Mode State/Province Sub-Category
          Ship Date
                                                                  Discount
                                                                              Profit
         07-01-2020
                     Standard Class
                                                           Paper
                                                                        0.2
                                              Texas
                                                                              5.5512
                     Standard Class
      1 08-01-2020
                                           Illinois
                                                         Binders
                                                                        0.8
                                                                            -5.4870
      2 08-01-2020
                     Standard Class
                                           Illinois
                                                          Labels
                                                                        0.2
                                                                              4.2717
      3 08-01-2020
                     Standard Class
                                           Illinois
                                                         Storage
                                                                        0.2 - 64.7748
      4 12-01-2020
                     Standard Class
                                       Pennsylvania
                                                                        0.2
                                                                              4.8840
                                                             Art
         Quantity
                     Sales
      0
                2
                    16.448
                2
      1
                     3.540
      2
                3
                    11.784
      3
                3
                   272.736
                3
                    19.536
[32]:
      store.columns
```

[32]: Index(['Category', 'City', 'Country/Region', 'Customer Name', 'Manufacturer', 'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',

```
dtype='object')
      store['Category']
[34]: 0
               Office Supplies
               Office Supplies
      1
      2
               Office Supplies
      3
               Office Supplies
               Office Supplies
      4
      10189
               Office Supplies
               Office Supplies
      10190
      10191
               Office Supplies
      10192
                    Technology
      10193
               Office Supplies
      Name: Category, Length: 10194, dtype: object
[35]: store[['Category','City']]
[35]:
                    Category
                                        City
      0
             Office Supplies
                                     Houston
      1
             Office Supplies
                                  Naperville
      2
             Office Supplies
                                  Naperville
      3
             Office Supplies
                                  Naperville
      4
             Office Supplies
                                Philadelphia
             Office Supplies
      10189
                               New York City
             Office Supplies
      10190
                                   Fairfield
      10191
             Office Supplies
                                    Loveland
      10192
                  Technology
                               New York City
             Office Supplies
      10193
                               Charlottetown
      [10194 rows x 2 columns]
[36]: store[['Category','City','Customer Name']]
[36]:
                    Category
                                                   Customer Name
                                        City
             Office Supplies
      0
                                     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
             Office Supplies
                               New York City Patrick O'Donnell
      10189
             Office Supplies
      10190
                                   Fairfield
                                                      Erica Bern
```

'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category',

'Discount', 'Profit', 'Quantity', 'Sales'],

```
10191Office SuppliesLovelandJill Matthias10192TechnologyNew York CityPatrick O'Donnell10193Office SuppliesCharlottetownHarry Olson
```

[10194 rows x 3 columns]

```
[37]: store.columns #store_cat=store
```

## [39]: 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
dtyp	es: float64(3),	int64(1), object	(15)

```
[40]: store_cat.info()
    <class 'pandas.core.frame.DataFrame'>
    RangeIndex: 10194 entries, 0 to 10193
    Data columns (total 15 columns):
        Column
                    Non-Null Count Dtype
       -----
                     _____
     0
                     10194 non-null object
        Category
     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
                     10194 non-null object
     5
        Order Date
     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
[41]: store.columns
[41]: 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')
[]: store_num = store[['Discount', 'Profit', 'Quantity', _
     []: print(len(store.columns))
    print(len(store_num.
     print(len(store cat.columns))
    19
    4
```

memory usage: 1.5+ MB

15

```
[44]:
                  Discount
                                   Profit
                                                Quantity
                                                                  Sales
                            10194.000000
                                           10194.000000
                                                          10194.000000
             10194.000000
      count
      mean
                  0.155385
                                28.673417
                                                3.791838
                                                            228.225854
      std
                  0.206249
                               232.465115
                                                2.228317
                                                            619.906839
                            -6599.978000
      min
                  0.000000
                                                1.000000
                                                              0.444000
      25%
                  0.00000
                                 1.760800
                                                2.000000
                                                             17.220000
      50%
                  0.200000
                                                3.000000
                                                             53.910000
                                 8.690000
      75%
                  0.200000
                                29.297925
                                                5.000000
                                                            209.500000
      max
                  0.800000
                             8399.976000
                                               14.000000
                                                          22638.480000
[45]: rev=store[::-1]
      rev
[45]:
                                         City Country/Region
                                                                    Customer Name
                     Category
             Office Supplies
                                Charlottetown
                                                       Canada
                                                                      Harry Olson
      10193
                   Technology
                                               United States
                                                               Patrick O'Donnell
      10192
                               New York City
             Office Supplies
                                     Loveland
                                               United States
                                                                    Jill Matthias
      10191
      10190
             Office Supplies
                                    Fairfield
                                               United States
                                                                       Erica Bern
                                                               Patrick O'Donnell
      10189
             Office Supplies
                               New York City
                                               United States
      ...
      4
             Office Supplies
                                 Philadelphia
                                               United States
                                                                       Mick Brown
      3
             Office Supplies
                                   Naperville
                                               United States
                                                                    Phillina Ober
      2
             Office Supplies
                                                                    Phillina Ober
                                   Naperville
                                               United States
             Office Supplies
      1
                                   Naperville
                                               United States
                                                                    Phillina Ober
             Office Supplies
                                                                    Darren Powers
                                      Houston
                                               United States
             Manufacturer
                            Order Date
                                                Order ID Postal Code
      10193
             Wilson Jones
                            30-12-2023
                                         CA-2023-143500
                                                                  COA
      10192
                     Other
                            30-12-2023
                                         US-2023-143259
                                                                10009
                     Other
                            30-12-2023
                                         US-2023-156720
      10191
                                                               80538
      10190
                       GBC
                            30-12-2023
                                         US-2023-115427
                                                               94533
      10189
                            30-12-2023
                                         US-2023-143259
                                                                10009
             Wilson Jones
      4
                     Avery
                            05-01-2020
                                         US-2020-141817
                                                               19143
      3
                     SAFCO
                            04-01-2020
                                         US-2020-112326
                                                                60540
      2
                     Avery
                            04-01-2020
                                         US-2020-112326
                                                                60540
      1
                       GBC
                            04-01-2020
                                         US-2020-112326
                                                               60540
             Message Book
      0
                            03-01-2020
                                         US-2020-103800
                                                               77015
                                                     Product Name
                                                                     Region
                                     Wilson Jones Impact Binders
      10193
                                                                       East
      10192
                                       Gear Head AU3700S Headset
                                                                       East
                                             Bagged Rubber Bands
      10191
                                                                       West
                                              GBC Binding covers
      10190
                                                                       West
      10189
                           Wilson Jones Legal Size Ring Binders
                                                                       East
```

[44]: store\_num.describe()

```
SAFCO Boltless Steel Shelving
      3
                                                                   Central
      2
                                                        Avery 508
                                                                    Central
                     GBC Standard Plastic Binding Systems Combs
      1
                                                                   Central
             Message Book, Wirebound, Four 5 1/2" X 4" Form...
      0
                  Segment
                            Ship Date
                                             Ship Mode
                                                               State/Province
                 Consumer
                           03-01-2024
                                        Standard Class
                                                         Prince Edward Island
      10193
      10192
                 Consumer
                           03-01-2024
                                        Standard Class
                                                                      New York
      10191
                 Consumer
                           03-01-2024
                                        Standard Class
                                                                      Colorado
      10190
                                        Standard Class
                                                                    California
               Corporate
                           03-01-2024
      10189
                 Consumer
                           03-01-2024
                                        Standard Class
                                                                      New York
      4
                           12-01-2020
                                                                 Pennsylvania
                 Consumer
                                        Standard Class
      3
             Home Office
                           08-01-2020
                                        Standard Class
                                                                      Illinois
      2
             Home Office
                                        Standard Class
                                                                      Illinois
                           08-01-2020
             Home Office
                                        Standard Class
      1
                           08-01-2020
                                                                      Illinois
      0
                 Consumer
                           07-01-2020
                                        Standard Class
                                                                         Texas
            Sub-Category
                           Discount
                                       Profit
                                               Quantity
                                                            Sales
      10193
                  Binders
                                0.2
                                                            3.024
                                      -0.6048
                                                       3
      10192
                                0.0
                                                       7
                                                           90.930
                   Phones
                                       2.7279
                                                       3
      10191
               Fasteners
                                0.2
                                      -0.6048
                                                            3.024
                                0.2
                                                       2
                                                           20.720
      10190
                  Binders
                                       6.4750
      10189
                  Binders
                                 0.2
                                      19.7910
                                                           52.776
      4
                                0.2
                                       4.8840
                                                       3
                                                           19.536
                      Art
      3
                  Storage
                                0.2 - 64.7748
                                                       3
                                                          272.736
      2
                   Labels
                                0.2
                                                       3
                                                           11.784
                                       4.2717
                                                       2
      1
                  Binders
                                0.8
                                      -5.4870
                                                            3.540
                    Paper
                                       5.5512
                                                           16.448
                                0.2
      [10194 rows x 19 columns]
[46]:
      rev.head()
[46]:
                                         City Country/Region
                                                                    Customer Name \
                     Category
             Office Supplies
                                                       Canada
                                                                      Harry Olson
      10193
                               Charlottetown
                   Technology
                                                               Patrick O'Donnell
      10192
                               New York City
                                               United States
      10191
             Office Supplies
                                     Loveland
                                               United States
                                                                    Jill Matthias
      10190
             Office Supplies
                                    Fairfield
                                               United States
                                                                       Erica Bern
      10189
             Office Supplies
                               New York City
                                               United States
                                                              Patrick O'Donnell
             Manufacturer
                                               Order ID Postal Code
                            Order Date
             Wilson Jones
                                         CA-2023-143500
      10193
                            30-12-2023
                                                                  COA
                            30-12-2023
      10192
                     Other
                                         US-2023-143259
                                                               10009
```

Avery Hi-Liter EverBold Pen Style Fluorescent ...

East

4

```
GBC
      10190
                            30-12-2023
                                        US-2023-115427
                                                              94533
      10189
             Wilson Jones
                            30-12-2023
                                        US-2023-143259
                                                              10009
                                                                        Ship Date
                                      Product Name Region
                                                              Segment
      10193
                      Wilson Jones Impact Binders
                                                      East
                                                             Consumer
                                                                       03-01-2024
      10192
                         Gear Head AU3700S Headset
                                                      East
                                                             Consumer
                                                                       03-01-2024
      10191
                               Bagged Rubber Bands
                                                      West
                                                             Consumer
                                                                       03-01-2024
      10190
                                GBC Binding covers
                                                      West
                                                            Corporate
                                                                       03-01-2024
      10189
             Wilson Jones Legal Size Ring Binders
                                                             Consumer
                                                      East
                                                                       03-01-2024
                  Ship Mode
                                    State/Province Sub-Category
                                                                  Discount
                                                                              Profit
      10193
             Standard Class
                              Prince Edward Island
                                                         Binders
                                                                       0.2
                                                                             -0.6048
                                                          Phones
      10192
             Standard Class
                                          New York
                                                                       0.0
                                                                              2.7279
      10191
             Standard Class
                                          Colorado
                                                                       0.2
                                                       Fasteners
                                                                             -0.6048
      10190
             Standard Class
                                        California
                                                         Binders
                                                                       0.2
                                                                              6.4750
             Standard Class
                                                                       0.2 19.7910
      10189
                                          New York
                                                         Binders
             Quantity
                        Sales
      10193
                    3
                         3.024
      10192
                    7
                       90.930
                    3
                        3.024
      10191
                       20.720
      10190
                    2
      10189
                       52.776
[47]: rev.tail()
[47]:
                                   City Country/Region
                Category
                                                         Customer Name
                                                                        Manufacturer \
      4 Office Supplies
                          Philadelphia
                                         United States
                                                            Mick Brown
                                                                                Avery
      3 Office Supplies
                             Naperville
                                         United States
                                                         Phillina Ober
                                                                                SAFCO
      2 Office Supplies
                             Naperville
                                         United States
                                                         Phillina Ober
                                                                                Avery
      1 Office Supplies
                             Naperville
                                                                                  GBC
                                         United States
                                                         Phillina Ober
      O Office Supplies
                                Houston
                                         United States
                                                         Darren Powers
                                                                        Message Book
                            Order ID Postal Code
         Order Date
      4 05-01-2020
                     US-2020-141817
                                            19143
      3 04-01-2020
                     US-2020-112326
                                           60540
      2 04-01-2020
                     US-2020-112326
                                           60540
         04-01-2020
                     US-2020-112326
                                           60540
         03-01-2020
                     US-2020-103800
                                           77015
                                               Product Name
                                                               Region
                                                                            Segment \
      4
         Avery Hi-Liter EverBold Pen Style Fluorescent ...
                                                               East
                                                                         Consumer
      3
                              SAFCO Boltless Steel Shelving
                                                                       Home Office
                                                              Central
      2
                                                  Avery 508
                                                              Central
                                                                       Home Office
      1
                GBC Standard Plastic Binding Systems Combs
                                                              Central
                                                                       Home Office
         Message Book, Wirebound, Four 5 1/2" X 4" Form... Central
                                                                         Consumer
```

10191

Other

30-12-2023

US-2023-156720

80538

```
Ship Mode State/Province Sub-Category Discount
   Ship Date
                                                                      Profit \
4 12-01-2020
               Standard Class
                                Pennsylvania
                                                                0.2
                                                                       4.8840
                                                      Art
3 08-01-2020
               Standard Class
                                                  Storage
                                                                0.2 -64.7748
                                    Illinois
                                                                0.2
2 08-01-2020
              Standard Class
                                    Illinois
                                                   Labels
                                                                       4.2717
                                                  Binders
1 08-01-2020
              Standard Class
                                    Illinois
                                                                0.8 -5.4870
0 07-01-2020
              Standard Class
                                       Texas
                                                    Paper
                                                                0.2
                                                                      5.5512
  Quantity
               Sales
4
          3
              19.536
             272.736
3
          3
2
          3
             11.784
          2
              3.540
1
0
          2
              16.448
```