super-store-data-analysis

November 21, 2024

1 Super Store Data Analysis

- 1.1 Pandas Workshop
- 1.1.1 Mohd Sameer Hussain mohdsameerhussain28@gmail.com
- 1.1.2 https://github.com/SameerHussain128?tab=repositories

```
[1]: import pandas as pd
    pd.__version__
[2]: '2.2.2'
     store = pd.read_csv('dataset.csv')
     store
[4]:
[4]:
                                       City Country/Region
                   Category
                                                                  Customer Name
     0
            Office Supplies
                                    Houston
                                             United States
                                                                  Darren Powers
     1
            Office Supplies
                                 Naperville
                                             United States
                                                                 Phillina Ober
     2
            Office Supplies
                                 Naperville
                                             United States
                                                                 Phillina Ober
     3
            Office Supplies
                                 Naperville
                                                                 Phillina Ober
                                             United States
     4
            Office Supplies
                               Philadelphia
                                             United States
                                                                    Mick Brown
     10189
            Office Supplies
                              New York City
                                             United States
                                                             Patrick O'Donnell
            Office Supplies
                                  Fairfield
     10190
                                             United States
                                                                     Erica Bern
     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 Date
                                             Order ID Postal Code
     0
            Message Book
                           03-01-2020
                                       US-2020-103800
                                                             77095
     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
     10189
           Wilson Jones
                          30-12-2023
                                      US-2023-143259
                                                             10009
```

```
10190
                 GBC
                      30-12-2023
                                  US-2023-115427
                                                         94533
10191
                      30-12-2023
                                   US-2023-156720
                                                         80538
              Other
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
10192
                                 Gear Head AU3700S Headset
                                                                 East
10193
                              Wilson Jones Impact Binders
                                                                 East
           Segment
                      Ship Date
                                       Ship Mode
                                                         State/Province
0
          Consumer
                     07-01-2020
                                  Standard Class
                                                                   Texas
       Home Office
1
                     08-01-2020
                                  Standard Class
                                                               Illinois
2
       Home Office
                                  Standard Class
                     08-01-2020
                                                               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
                     03-01-2024
10190
         Corporate
                                  Standard Class
                                                             California
10191
          Consumer
                     03-01-2024
                                  Standard Class
                                                               Colorado
10192
          Consumer
                     03-01-2024
                                  Standard Class
                                                               New York
                                 Standard Class Prince Edward Island
10193
          Consumer
                     03-01-2024
      Sub-Category
                                                      Sales
                     Discount
                                Profit
                                         Quantity
0
             Paper
                          0.2
                                                     16.448
                                5.5512
                                                 2
1
           Binders
                          0.8
                               -5.4870
                                                      3.540
2
            Labels
                          0.2
                                4.2717
                                                 3
                                                     11.784
3
           Storage
                          0.2 - 64.7748
                                                 3
                                                    272.736
4
                          0.2
                                4.8840
                                                 3
                                                     19.536
               Art
10189
           Binders
                          0.2 19.7910
                                                 3
                                                     52.776
10190
           Binders
                          0.2
                                6.4750
                                                 2
                                                     20.720
                                                 3
10191
         Fasteners
                          0.2
                               -0.6048
                                                      3.024
10192
            Phones
                          0.0
                                2.7279
                                                 7
                                                     90.930
10193
           Binders
                          0.2 - 0.6048
                                                      3.024
```

[5]: len(store)

```
[5]: 10194
[6]: store.columns
[6]: 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')
[7]: store.shape
[7]: (10194, 19)
[8]: store.isnull().sum()
                       0
[8]: Category
                       0
     City
     Country/Region
                       0
     Customer Name
                       0
     Manufacturer
                       0
     Order Date
                       0
     Order ID
                       0
     Postal Code
                       0
     Product Name
                       0
     Region
                       0
                       0
     Segment
     Ship Date
                       0
     Ship Mode
                       0
     State/Province
                       0
                       0
     Sub-Category
    Discount
                       0
                       0
     Profit
                       0
     Quantity
     Sales
     dtype: int64
```

2 Dataset is clean No missing values

```
[9]: store.dtypes

[9]: Category object
    City object
    Country/Region object
    Customer Name object
    Manufacturer object
```

object Order Date object Order ID Postal Code object Product Name object Region object Segment object Ship Date object object Ship Mode object State/Province object Sub-Category Discount float64 Profit float64 Quantity int64 Sales float64

dtype: object

[10]: store.isna()

[10]:		Category	City	Country/Region	Customer Name	Manufacturer	\
	0	False	•	False	False	False	
	1	False	False	False	False	False	
	2	False		False	False	False	
	3	False	False	False	False	False	
	4	False		False	False	False	
	•••	•••		•••			
	10189	False	False	False	False	False	
	10190	False		False	False	False	
		False		False	False	False	
	10192	False		False	False	False	
	10193	False		False	False	False	
		Order Date	e Order	r ID Postal Coo	le Product Name	Region Seg	ment \
	0	False	e Fa	alse Fals	se False	False F	alse
	1	False	e Fa	alse Fals	se False	False F	alse
	2	False	e Fa	alse Fals	se False	False F	alse
	3	False	e Fa	alse Fals	se False	False F	alse
	4	False	e Fa	alse Fals	se False	False F	alse
		•••	•••	•••	***	•••	
	10189	False	e Fa	alse Fals	se False	False F	alse
	10190	False	e Fa	alse Fals	se False	False F	alse
	10191	False	e Fa	alse Fals	se False	False F	alse
	10192	False	e Fa	alse Fals	se False	False F	alse
	10193	False	e Fa	alse Fals	se False	False F	alse
		Ship Date	Ship N	Mode State/Prov	vince Sub-Categ	orv Discount	Profit \
	0	False	_		_	lse False	
	1	False				lse False	

	2		False		False	F	alse		False		False	False	
	3		False	:	False	F	alse		False		False	False	
	4		False		False	F	alse		False		False	False	
	•••		•••	•••		•••			•••	•••			
	101	L89	False		False	F	alse		False		False	False	
	101	L90	False		False	F	alse		False		False	False	
	101	L91	False		False	F	alse		False		False	False	
	101	192	False		False	F	alse		False		False	False	
	101		False		False		alse		False		False	False	
		Qua	antity	Sales									
	0	•	False	False									
	1			False									
	2			False									
	3			False									
	4		False										
	 101	190	 False	False									
			False	False									
	101												
	101			False									
	101			False									
	101	193	False	False									
	Γ1.	104	10	7									
	LIC)194 ro	WS X 19	COLUM	nsj								
	g+	ore.hea	d()										
•	500	ore mea	u()										
:			Catego	rv	City	Countr	y/Region	Cus	tomer	Name	Manıı	facturer	. \
•	0	Office	Suppli	-	Houston		d States		ren Po			age Book	
	1		Suppli		Naperville				llina		11000	GBC	
	2		Suppli		Naperville		d States		llina			Avery	
	3		Suppli		Naperville Naperville		d States		llina			SAFCO	
					-								
	4	oiiice	Suppli	es Pii	iladelphia	onre	d States		Mick E	PLOMII		Avery	'
		Order 1	Dato	O-	rder ID Pos	r+al Ca	4o \						
	^	03-01-			-103800	770	,						
	0												
	1	04-01-			-112326	605							
		04-01-			-112326	605							
	3	04-01-			-112326	605							
	4	05-01-	2020 U	S-2020	-141817	191	43						
							D 1 :	3.7			~		
	^	М -	- D '	17.	1 -	E 4 /0"	Product		_	gion		O	\
	0	Message	-		ound, Four				Centra		Cons		
	1		GBC St	andard	Plastic B	inding	•					Office	
	2						Aver	y 508	Cent	ral	Home (Office	
	2				0 4 E 0 C E -	~	~ .				TT .	~ ~ ~ ·	

[11]

[11]

4 Avery Hi-Liter EverBold Pen Style Fluorescent ...

SAFCO Boltless Steel Shelving Central Home Office

East

Consumer

```
07-01-2020
                      Standard Class
                                               Texas
                                                            Paper
                                                                         0.2
                                                                               5.5512
                      Standard Class
         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
                                                                         0.2
                                                                               4.8840
                                       Pennsylvania
                                                              Art
         Quantity
                     Sales
      0
                2
                     16.448
                2
      1
                      3.540
      2
                3
                     11.784
      3
                3
                   272.736
                3
                     19.536
      store.tail()
[12]:
[12]:
                     Category
                                        City Country/Region
                                                                   Customer Name
                               New York City United States
             Office Supplies
                                                              Patrick O'Donnell
      10189
             Office Supplies
                                   Fairfield United States
                                                                      Erica Bern
      10190
      10191
             Office Supplies
                                    Loveland United States
                                                                   Jill Matthias
      10192
                  Technology
                               New York City United States Patrick O'Donnell
      10193
             Office Supplies
                               Charlottetown
                                                      Canada
                                                                     Harry Olson
             Manufacturer
                            Order Date
                                               Order ID Postal Code
             Wilson Jones
      10189
                            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
      10192
                     Other
                            30-12-2023
                                        US-2023-143259
                                                               10009
                            30-12-2023
                                        CA-2023-143500
                                                                 COA
      10193
             Wilson Jones
                                      Product Name Region
                                                              Segment
                                                                         Ship Date
             Wilson Jones Legal Size Ring Binders
                                                      East
                                                                        03-01-2024
      10189
                                                             Consumer
                                GBC Binding covers
      10190
                                                      West
                                                            Corporate
                                                                        03-01-2024
      10191
                               Bagged Rubber Bands
                                                              Consumer
                                                      West
                                                                        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
      10189
             Standard Class
                                          New York
                                                         Binders
                                                                        0.2
                                                                             19.7910
      10190
             Standard Class
                                        California
                                                         Binders
                                                                        0.2
                                                                              6.4750
             Standard Class
                                                                        0.2 - 0.6048
      10191
                                          Colorado
                                                       Fasteners
      10192
             Standard Class
                                          New York
                                                          Phones
                                                                        0.0
                                                                              2.7279
      10193
             Standard Class Prince Edward Island
                                                         Binders
                                                                        0.2 - 0.6048
             Quantity
                        Sales
      10189
                     3
                       52.776
```

Ship Mode State/Province Sub-Category

Discount

Profit \

Ship Date

```
10191
                    3
                        3.024
      10192
                    7 90.930
      10193
                        3.024
[13]: store.head(2)
[13]:
                                 City Country/Region
                                                                     Manufacturer \
                Category
                                                      Customer Name
      O Office Supplies
                             Houston United States
                                                      Darren Powers
                                                                      Message Book
      1 Office Supplies
                          Naperville
                                      United States
                                                      Phillina Ober
                                                                               GBC
         Order Date
                            Order ID Postal Code
      0 03-01-2020
                     US-2020-103800
                                           77095
      1 04-01-2020
                     US-2020-112326
                                           60540
                                               Product Name
                                                              Region
                                                                           Segment \
         Message Book, Wirebound, Four 5 1/2" X 4" Form... Central
                                                                        Consumer
      1
                GBC Standard Plastic Binding Systems Combs Central
                                                                       Home Office
                          Ship Mode State/Province Sub-Category
          Ship Date
                                                                   Discount
                                                                             Profit
      0 07-01-2020
                     Standard Class
                                              Texas
                                                            Paper
                                                                        0.2 5.5512
      1 08-01-2020
                     Standard Class
                                           Illinois
                                                         Binders
                                                                        0.8 - 5.4870
         Quantity
                    Sales
      0
                2
                   16.448
      1
                    3.540
[14]: store[:]
[14]:
                    Category
                                        City Country/Region
                                                                  Customer Name
             Office Supplies
                                     Houston United States
                                                                  Darren Powers
      0
             Office Supplies
                                  Naperville
                                              United States
                                                                  Phillina Ober
      1
      2
             Office Supplies
                                  Naperville
                                              United States
                                                                  Phillina Ober
      3
             Office Supplies
                                                                  Phillina Ober
                                  Naperville
                                              United States
      4
             Office Supplies
                                Philadelphia
                                              United States
                                                                     Mick Brown
      10189
             Office Supplies
                              New York City
                                              United States
                                                             Patrick O'Donnell
             Office Supplies
                                   Fairfield
                                              United States
      10190
                                                                     Erica Bern
      10191
             Office Supplies
                                    Loveland United States
                                                                  Jill Matthias
      10192
                  Technology
                              New York City
                                              United States
                                                             Patrick O'Donnell
      10193
             Office Supplies
                              Charlottetown
                                                     Canada
                                                                    Harry Olson
                                              Order ID Postal Code \
             Manufacturer
                           Order Date
                                                              77095
      0
             Message Book
                           03-01-2020
                                        US-2020-103800
      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
```

10190

2 20.720

4	Avery	05-01-2020 US-2020-141817 19143	
	•••		
10189	Wilson Jones		
10190	GBC		
10191	Other		
10192	Other	30-12-2023 US-2023-143259 10009	
10193	Wilson Jones	30-12-2023 CA-2023-143500 COA	
		Duradurat Name - Davidson \	
^	Managana Daala	Product Name Region \	
0	~	, Wirebound, Four 5 1/2" X 4" Form Central	
1	GDC 5	tandard Plastic Binding Systems Combs Central	
2		Avery 508 Central	
	Arramer Hi Ti+	SAFCO Boltless Steel Shelving Central	
4	Avery HI-LIC	er EverBold Pen Style Fluorescent East	
 10100		Wilson Jones Legal Size Ring Binders East	
10189			
10190 10191		GBC Binding covers West	
		Bagged Rubber Bands West Gear Head AU3700S Headset East	
10192			
10193		Wilson Jones Impact Binders East	
	Segment	Ship Date Ship Mode State/Province	\
0	Consumer	07-01-2020 Standard Class Texas	
1	Home Office	08-01-2020 Standard Class 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	
10191	Consumer	03-01-2024 Standard Class Colorado	
10192	Consumer	03-01-2024 Standard Class New York	
10193	Consumer	03-01-2024 Standard Class Prince Edward Island	
	a 1 a .		
		Discount Profit Quantity Sales	
0	Paper	0.2 5.5512 2 16.448	
1	Binders	0.8 -5.4870 2 3.540	
2	Labels	0.2 4.2717 3 11.784	
3	Storage	0.2 -64.7748 3 272.736	
4	Art	0.2 4.8840 3 19.536	
 10189	 Binders		
10109	Binders	0.2 19.7910 3 52.776 0.2 6.4750 2 20.720	
10190	Fasteners	0.2 -0.6048 3 3.024	
10191	Phones	0.0 2.7279 7 90.930	
10192	Binders	0.0 2.7279 7 90.930 0.2 -0.6048 3 3.024	
10193	pringers	0.2 -0.0040 3 3.024	

	Categ	forv	City	Country/Regi	on Cu	stomer Name
10193	Office Suppl		ttetown	Cana		Harry Olson
10192	Technol		rk City			k O'Donnell
10191	Office Suppl	00	oveland			ll Matthias
10190	Office Suppl		irfield			Erica Beri
10189	Office Suppl		rk City			k O'Donnell
4	Office Suppl	ies Phila	delphia	United Stat	es	Mick Brown
3	Office Suppl		erville		es Ph	illina Obe
2	Office Suppl	_	erville		es Ph	illina Obe
1	Office Suppl	-	erville		es Ph	illina Obe
0	Office Suppl	_	Houston		es Da	rren Power:
	Manufacturer	Order Dat	е	Order ID Po	stal Code	\
10193	Wilson Jones	30-12-202	3 CA-2	023-143500	COA	
10192	Other	30-12-202	3 US-20	023-143259	10009	
10191	Other	30-12-202	3 US-20	023-156720	80538	
10190	GBC	30-12-202	3 US-20	023-115427	94533	
10189	Wilson Jones	30-12-202	3 US-20	023-143259	10009	
•••	•••	•••				
4	Avery	05-01-202	0 US-20	020-141817	19143	
3	SAFCO			020-112326	60540	
2	Avery			020-112326	60540	
1	GBC			020-112326	60540	
0	Message Book	03-01-202	0 US-20	020-103800	77095	
				ъ.,		. \
10102		T.:	diaon I			egion \
10193 10192		W		ones Impact E ead AU3700S H		East East
10192				Bagged Rubber		West
10191			1	GBC Binding		West
10130		Wilson Ion	os Ioma	l Size Ring E		East
		WIISON JON	.es Lega.	r bize ming r	THUELD	Last
4	Avery Hi-Tit	er EverBold	Pen St	yle Fluoresce	nt F	ast
3	Avery III EIG			less Steel Sh		ntral
2		DAI	OU DOIG		•	ntral
1	GBC S	Standard Pla	stic Bi	nding Systems	•	
0				5 1/2" X 4" F		
Ü	11022460 2001	i, wii obouiid	, 1041	3 1, 2 11 1 1	01111111	- 41
	Segment	Ship Date	:	Ship Mode	State/	Province
10193	Consumer	03-01-2024		_	rince Edwar	
10192	Consumer	03-01-2024		ard Class		New York
10101	2 2 11 D am 0 1		2 2 2 2		•	

Colorado

10191 Consumer 03-01-2024 Standard Class

10190	Corporate	03-01-2024	Standa	rd Class		California
10189	Consumer	03-01-2024	Standa	rd Class		New York
•••	•••	•••	•••			•
4	Consumer	12-01-2020	Standa	rd Class	Pe	nnsylvania
3	Home Office	08-01-2020	Standa	rd Class		Illinois
2	Home Office	08-01-2020	Standa	rd Class		Illinois
1	Home Office	08-01-2020	Standa	rd Class		Illinois
0	Consumer	07-01-2020	Standa	rd Class		Texas
	Sub-Category	Discount	Profit	Quantity	Sales	
10193	Binders	0.2 -	0.6048	3	3.024	
10192	Phones	0.0	2.7279	7	90.930	
10191	Fasteners	0.2 -	0.6048	3	3.024	
10190	Binders	0.2	6.4750	2	20.720	
10189	Binders	0.2 1	9.7910	3	52.776	
	•••		•••	•••		
4	Art	0.2	4.8840	3	19.536	
3	Storage	0.2 -6	4.7748	3	272.736	
2	Labels	0.2	4.2717	3	11.784	
1	Binders	0.8 -	5.4870	2	3.540	
0	Paper	0.2	5.5512	2	16.448	

```
[16]: store[1:10:3]
```

```
[16]:
                Category
                                  City Country/Region Customer Name Manufacturer
      1 Office Supplies
                            Naperville
                                        United States
                                                       Phillina Ober
                                                                              GBC
      4 Office Supplies
                         Philadelphia
                                        United States
                                                          Mick Brown
                                                                            Avery
      7 Office Supplies
                                Athens
                                        United States
                                                       Jack O'Briant
                                                                            Dixon
                           Order ID Postal Code \
         Order Date
      1 04-01-2020 US-2020-112326
                                          60540
      4 05-01-2020
                     US-2020-141817
                                          19143
      7 06-01-2020
                    US-2020-106054
                                          30605
                                              Product Name
                                                             Region
                                                                         Segment \
      1
                GBC Standard Plastic Binding Systems Combs
                                                            Central
                                                                     Home Office
      4 Avery Hi-Liter EverBold Pen Style Fluorescent ...
                                                             East
                                                                      Consumer
      7 Dixon Prang Watercolor Pencils, 10-Color Set w...
                                                            South
                                                                     Corporate
          Ship Date
                          Ship Mode State/Province Sub-Category
                                                                 Discount
                                                                           Profit
      1 08-01-2020
                                                        Binders
                                                                      0.8 -5.4870
                     Standard Class
                                          Illinois
      4 12-01-2020
                     Standard Class
                                      Pennsylvania
                                                            Art
                                                                      0.2 4.8840
```

Quantity Sales

First Class

7 07-01-2020

Georgia

0.0 5.2398

Art

```
4
                3 19.536
      7
                3 12.780
[17]: store.info()
     <class 'pandas.core.frame.DataFrame'>
     RangeIndex: 10194 entries, 0 to 10193
     Data columns (total 19 columns):
                          Non-Null Count Dtype
          Column
          _____
                          -----
      0
                                          object
          Category
                          10194 non-null
      1
                          10194 non-null
                                          object
          City
      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
          Product Name
      8
                          10194 non-null object
      9
          Region
                          10194 non-null object
          Segment
      10
                          10194 non-null object
          Ship Date
      11
                          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
          Discount
                          10194 non-null
                                          float64
      15
      16 Profit
                          10194 non-null float64
      17
          Quantity
                          10194 non-null
                                          int64
                          10194 non-null
      18
          Sales
                                          float64
     dtypes: float64(3), int64(1), object(15)
     memory usage: 1.5+ MB
[18]: store.columns
[18]: 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')
[19]:
      store['Category']
[19]: 0
               Office Supplies
               Office Supplies
      1
               Office Supplies
      2
      3
               Office Supplies
```

1

2

3.540

```
4
               Office Supplies
      10189
               Office Supplies
      10190
               Office Supplies
      10191
               Office Supplies
      10192
                    Technology
      10193
               Office Supplies
      Name: Category, Length: 10194, dtype: object
[20]: store['City']
[20]: 0
                     Houston
                  Naperville
      1
      2
                  Naperville
      3
                  Naperville
      4
                Philadelphia
      10189
               New York City
                   Fairfield
      10190
      10191
                    Loveland
      10192
               New York City
      10193
               Charlottetown
      Name: City, Length: 10194, dtype: object
[21]: store['City'].value_counts()
[21]: City
      New York City
                       915
      Los Angeles
                       747
      Philadelphia
                       537
      San Francisco
                       510
      Seattle
                        428
      Manhattan
                          1
      Iowa City
                          1
      Cedar Rapids
                          1
      Baytown
                          1
      Danbury
                          1
      Name: count, Length: 542, dtype: int64
[22]: store[['Category','City']]
[22]:
                    Category
                                        City
             Office Supplies
      0
                                     Houston
      1
             Office Supplies
                                  Naperville
      2
             Office Supplies
                                  Naperville
      3
             Office Supplies
                                  Naperville
```

```
4
             Office Supplies
                                Philadelphia
      10189
             Office Supplies
                               New York City
             Office Supplies
      10190
                                   Fairfield
      10191
             Office Supplies
                                    Loveland
      10192
                  Technology
                              New York City
      10193 Office Supplies
                              Charlottetown
      [10194 rows x 2 columns]
      store[['Category','City','Order ID']]
[23]:
                                                    Order ID
                    Category
                                        City
      0
             Office Supplies
                                     Houston
                                              US-2020-103800
      1
             Office Supplies
                                  Naperville
                                              US-2020-112326
                                  Naperville
      2
             Office Supplies
                                              US-2020-112326
             Office Supplies
      3
                                  Naperville
                                              US-2020-112326
      4
             Office Supplies
                                Philadelphia
                                              US-2020-141817
             Office Supplies
                              New York City
      10189
                                              US-2023-143259
             Office Supplies
      10190
                                   Fairfield US-2023-115427
                                    Loveland US-2023-156720
             Office Supplies
      10191
      10192
                  Technology
                              New York City US-2023-143259
             Office Supplies
                              Charlottetown CA-2023-143500
      10193
      [10194 rows x 3 columns]
      store.dtypes
[24]: Category
                         object
                         object
      City
      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

Quantity

float64 int64

```
dtype: object
[25]: len(store.dtypes)
[25]: 19
[26]:
      store.columns
[26]: 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')
[27]: store_categorical=store[['Category', 'City', 'Country/Region', 'Customer Name', __
       'Order Date', 'Order ID', 'Postal Code', 'Product Name', 'Region',
             'Segment', 'Ship Date', 'Ship Mode', 'State/Province', 'Sub-Category']]
[28]: store_categorical
[28]:
                                                                Customer Name
                    Category
                                       City Country/Region
             Office Supplies
                                    Houston United States
                                                                Darren Powers
      0
      1
             Office Supplies
                                 Naperville United States
                                                                Phillina Ober
      2
             Office Supplies
                                 Naperville United States
                                                                Phillina Ober
                                                                Phillina Ober
      3
             Office Supplies
                                 Naperville United States
      4
             Office Supplies
                               Philadelphia United States
                                                                   Mick Brown
      10189
            Office Supplies
                              New York City United States
                                                            Patrick O'Donnell
             Office Supplies
                                  Fairfield United States
      10190
                                                                   Erica Bern
      10191
             Office Supplies
                                   Loveland United States
                                                                 Jill Matthias
      10192
                  Technology
                              New York City
                                             United States
                                                            Patrick O'Donnell
      10193
            Office Supplies
                              Charlottetown
                                                    Canada
                                                                  Harry Olson
             Manufacturer
                           Order Date
                                             Order ID Postal Code \
      0
             Message Book
                           03-01-2020
                                       US-2020-103800
                                                            77095
      1
                      GBC
                           04-01-2020
                                       US-2020-112326
                                                             60540
      2
                           04-01-2020
                                       US-2020-112326
                                                             60540
                    Avery
      3
                    SAFCO
                           04-01-2020
                                       US-2020-112326
                                                             60540
      4
                    Avery
                           05-01-2020
                                       US-2020-141817
                                                             19143
            Wilson Jones
                                                             10009
      10189
                           30-12-2023
                                       US-2023-143259
      10190
                      GBC
                           30-12-2023
                                       US-2023-115427
                                                            94533
      10191
                    Other
                           30-12-2023
                                       US-2023-156720
                                                            80538
      10192
                           30-12-2023
                                                             10009
                    Other
                                       US-2023-143259
      10193 Wilson Jones
                           30-12-2023
                                       CA-2023-143500
                                                              COA
```

Sales

float64

```
Product Name
                                                              Region \
       Message Book, Wirebound, Four 5 1/2" X 4" Form... Central
0
              GBC Standard Plastic Binding Systems Combs
1
                                                             Central
2
                                                 Avery 508
                                                             Central
                            SAFCO Boltless Steel Shelving
3
                                                            Central
4
       Avery Hi-Liter EverBold Pen Style Fluorescent ...
                                                             East
10189
                    Wilson Jones Legal Size Ring Binders
                                                                East
10190
                                       GBC Binding covers
                                                                West
                                       Bagged Rubber Bands
10191
                                                                West
10192
                                Gear Head AU3700S Headset
                                                                East
10193
                              Wilson Jones Impact Binders
                                                                East
                      Ship Date
                                                        State/Province
           Segment
                                       Ship Mode
0
          Consumer
                    07-01-2020
                                 Standard Class
                                                                  Texas
1
       Home Office
                    08-01-2020
                                 Standard Class
                                                               Illinois
2
       Home Office
                                 Standard Class
                    08-01-2020
                                                               Illinois
3
       Home Office
                    08-01-2020
                                 Standard Class
                                                               Illinois
          Consumer
                     12-01-2020
                                 Standard Class
                                                          Pennsylvania
                                 Standard Class
10189
                    03-01-2024
                                                              New York
          Consumer
10190
                    03-01-2024
                                 Standard Class
                                                            California
         Corporate
          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
0
             Paper
           Binders
1
2
            Labels
3
           Storage
4
               Art
10189
           Binders
10190
           Binders
         Fasteners
10191
10192
            Phones
10193
           Binders
[10194 rows x 15 columns]
```

[29]: store_categorical.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10194 entries, 0 to 10193
Data columns (total 15 columns):

```
Column
                           Non-Null Count
                                           Dtype
      #
          _____
                           _____
      0
          Category
                           10194 non-null
                                           object
      1
          City
                           10194 non-null
                                           object
      2
                           10194 non-null
          Country/Region
                                           object
      3
          Customer Name
                           10194 non-null
                                           object
      4
          Manufacturer
                           10194 non-null
                                           object
          Order Date
      5
                           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
          Segment
      10
                           10194 non-null
                                           object
          Ship Date
                           10194 non-null
                                           object
      11
          Ship Mode
      12
                           10194 non-null
                                           object
          State/Province
                           10194 non-null
                                           object
          Sub-Category
                           10194 non-null
                                           object
     dtypes: object(15)
     memory usage: 1.2+ MB
[30]: store_categorical.dtypes #Text data
[30]: Category
                        object
                        object
      City
      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
      dtype: object
[31]: store_numerical=store[['Discount', 'Profit', 'Quantity', 'Sales']] #Numbers
[32]:
      store_numerical
[32]:
             Discount
                        Profit
                                Quantity
                                             Sales
      0
                  0.2
                        5.5512
                                        2
                                            16.448
                                             3.540
      1
                  0.8 - 5.4870
                                        2
                  0.2
      2
                        4.2717
                                        3
                                            11.784
```

```
4
                  0.2 4.8840
                                       3
                                           19.536
                  0.2 19.7910
                                           52.776
      10189
                                       3
      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 x 4 columns]
[33]: store_numerical.dtypes
[33]: Discount
                  float64
     Profit
                  float64
                    int64
      Quantity
      Sales
                  float64
      dtype: object
[34]: store.columns
[34]: 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')
[35]: len(store.columns)
[35]: 19
[36]: store_categorical.columns
[36]: 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')
[37]: len(store_categorical.columns)
[37]: 15
[38]: store_numerical.columns
[38]: Index(['Discount', 'Profit', 'Quantity', 'Sales'], dtype='object')
[39]: len(store_numerical.columns)
```

3 272.736

3

0.2 - 64.7748

```
[40]: print(len(store.columns))
      print(len(store_categorical.columns))
      print(len(store_numerical.columns))
     19
     15
     4
[41]: store.head()
[41]:
                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
                                           77095
      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
                                                                          Segment \
                                               Product Name
                                                              Region
         Message Book, Wirebound, Four 5 1/2" X 4" Form... Central
                                                                       Consumer
      0
                GBC Standard Plastic Binding Systems Combs
      1
                                                             Central
                                                                      Home Office
      2
                                                  Avery 508
                                                             Central
                                                                      Home Office
      3
                             SAFCO Boltless Steel Shelving
                                                             Central
                                                                      Home Office
        Avery Hi-Liter EverBold Pen Style Fluorescent ...
                                                                       Consumer
                                                              East
                          Ship Mode State/Province Sub-Category
          Ship Date
                                                                  Discount
                                                                             Profit \
                     Standard Class
      0 07-01-2020
                                              Texas
                                                           Paper
                                                                       0.2
                                                                             5.5512
      1 08-01-2020
                     Standard Class
                                           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
                                      Pennsylvania
      4 12-01-2020
                     Standard Class
                                                             Art
                                                                       0.2
                                                                             4.8840
         Quantity
                     Sales
      0
                    16.448
                2
                2
                     3.540
      1
      2
                3
                    11.784
      3
                3
                   272.736
                    19.536
```

[39]: 4

```
[42]: store.tail()
[42]:
                     Category
                                        City Country/Region
                                                                   Customer Name
             Office Supplies
                                                               Patrick O'Donnell
                               New York City
                                               United States
      10189
      10190
             Office Supplies
                                   Fairfield
                                               United States
                                                                      Erica Bern
             Office Supplies
      10191
                                    Loveland
                                               United States
                                                                   Jill Matthias
      10192
                   Technology
                               New York City
                                               United States
                                                               Patrick O'Donnell
      10193
             Office Supplies
                               Charlottetown
                                                      Canada
                                                                     Harry Olson
             Manufacturer
                            Order Date
                                               Order ID Postal Code
      10189
             Wilson Jones
                            30-12-2023
                                        US-2023-143259
                                                               10009
      10190
                       GBC
                            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
                                                               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
                                    State/Province Sub-Category
                   Ship Mode
                                                                   Discount
                                                                              Profit
      10189
             Standard Class
                                           New York
                                                         Binders
                                                                        0.2 19.7910
      10190
             Standard Class
                                        California
                                                                        0.2
                                                         Binders
                                                                               6.4750
             Standard Class
      10191
                                           Colorado
                                                       Fasteners
                                                                        0.2
                                                                             -0.6048
      10192
             Standard Class
                                           New York
                                                          Phones
                                                                        0.0
                                                                               2.7279
             Standard Class Prince Edward Island
      10193
                                                         Binders
                                                                        0.2 - 0.6048
                         Sales
             Quantity
      10189
                     3 52.776
      10190
                     2
                        20.720
                     3
                         3.024
      10191
                     7
      10192
                        90.930
      10193
                         3.024
[43]: store[:]
[43]:
                                        City Country/Region
                                                                   Customer Name
                     Category
      0
             Office Supplies
                                     Houston United States
                                                                   Darren Powers
      1
             Office Supplies
                                               United States
                                                                   Phillina Ober
                                  Naperville
      2
             Office Supplies
                                  Naperville
                                               United States
                                                                   Phillina Ober
      3
             Office Supplies
                                  Naperville
                                                                   Phillina Ober
                                               United States
      4
             Office Supplies
                                Philadelphia
                                               United States
                                                                      Mick Brown
```

United States

New York City

10189

Office Supplies

Patrick O'Donnell

```
10190
       Office Supplies
                             Fairfield United States
                                                                Erica Bern
                              Loveland United States
10191
       Office Supplies
                                                             Jill Matthias
10192
            Technology
                         New York City
                                        United States
                                                        Patrick O'Donnell
10193
       Office Supplies
                         Charlottetown
                                                Canada
                                                               Harry Olson
       Manufacturer
                      Order Date
                                         Order ID Postal Code
0
                      03-01-2020
                                  US-2020-103800
       Message Book
                                                         77095
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
              Averv
                      05-01-2020
                                  US-2020-141817
                                                         19143
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
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
              GBC Standard Plastic Binding Systems Combs
1
                                                             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
                                        GBC Binding covers
10190
                                                                West
                                       Bagged Rubber Bands
10191
                                                                West
10192
                                Gear Head AU3700S Headset
                                                                East
10193
                              Wilson Jones Impact Binders
                                                                East
           Segment
                      Ship Date
                                       Ship Mode
                                                         State/Province
0
          Consumer
                     07-01-2020
                                                                  Texas
                                 Standard Class
1
       Home Office
                     08-01-2020
                                 Standard Class
                                                               Illinois
2
       Home Office
                     08-01-2020
                                 Standard Class
                                                               Illinois
3
       Home Office
                     08-01-2020
                                 Standard Class
                                                               Illinois
4
                     12-01-2020
                                 Standard Class
          Consumer
                                                           Pennsylvania
10189
          Consumer
                     03-01-2024
                                 Standard Class
                                                               New York
                                 Standard Class
                                                             California
10190
         Corporate
                     03-01-2024
                                 Standard Class
10191
          Consumer
                     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
                                Profit
                                         Quantity
                                                     Sales
                    Discount
0
                                5.5512
                                                2
                                                     16.448
             Paper
                          0.2
1
                                                2
           Binders
                          0.8
                               -5.4870
                                                     3.540
```

2	Labels		0.2	4.2717		3	11.784
3	Storage		0.2	-64.7748		3	272.736
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
10191	Fasteners		0.2	-0.6048		3	3.024
10192	Phones		0.0	2.7279		7	90.930
10193	Binders		0.2	-0.6048		3	3.024

[44]: store[::-1]

[44]:		Catego	ory	City	Country/Region	Customer Name	\
	10193	Office Suppli	ies Charlott	etown	Canada	Harry Olson	
	10192	Technolo	ogy New York	City	United States	Patrick O'Donnell	
	10191	Office Suppli	ies Lov	eland	United States	Jill Matthias	
	10190	Office Suppli	ies Fair	field	United States	Erica Bern	
	10189	Office Suppli		City	United States	Patrick O'Donnell	
	•••	•••	•••		•••	•••	
	4	Office Suppli	ies Philade	lphia	United States	Mick Brown	
	3	Office Suppli		ville		Phillina Ober	
	2	Office Suppli	_	ville		Phillina Ober	
	1	Office Suppli	_	ville		Phillina Ober	
	0	Office Suppli	_	uston		Darren Powers	
		Manufacturer	Order Date		Order ID Posta	al Code \	
	10193	Wilson Jones	30-12-2023	CA-20	23-143500	COA	
	10192	Other	30-12-2023	US-20	23-143259	10009	
	10191	Other	30-12-2023	US-20	23-156720	80538	
	10190	GBC	30-12-2023	US-20	23-115427	94533	
	10189	Wilson Jones	30-12-2023	US-20	23-143259	10009	
	•••	•••	•••				
	4	Avery	05-01-2020	US-20	20-141817	19143	
	3	SAFCO	04-01-2020	US-20	20-112326	60540	
	2	Avery	04-01-2020	US-20	20-112326	60540	
	1	GBC	04-01-2020	US-20	20-112326	60540	
	0	Message Book			20-103800	77095	
		_					
					Product N	Name Region \	
	10193		Wil	son Jo	nes Impact Bind	lers East	
	10192		C	ear He	ad AU3700S Head	lset East	
	10191			Е	Bagged Rubber Ba	ands West	
	10190				GBC Binding cov		
	10189		Wilson Jones		Size Ring Bind		
	•••			0.	J	•••	

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

[45]: store[0:100:10]

[45]:		Category	City	Country/Region	Customer Name	١
	0	Office Supplies	Houston	United States	Darren Powers	
	10	Office Supplies	Henderson	United States	Maria Etezadi	
	20	Furniture	Dover	United States	Seth Vernon	
	30	Office Supplies	San Francisco	United States	Brian Dahlen	
	40	Office Supplies	Scottsdale	United States	Toby Swindell	
	50	Office Supplies	Jonesboro	United States	Hunter Lopez	
	60	Furniture	Miami	United States	Tom Boeckenhauer	
	70	Furniture	Alexandria	United States	Shirley Daniels	
	80	Office Supplies	New York City	United States	Jasper Cacioppo	
	90	Office Supplies	Escondido	United States	Mark Haberlin	

```
Order ID Postal Code \
    Manufacturer
                  Order Date
0
    Message Book
                  03-01-2020
                               US-2020-103800
                                                     77095
      Southworth
10
                  06-01-2020
                               US-2020-167199
                                                     42420
20
                  11-01-2020
                               US-2020-130092
                                                     19901
             DAX
30
         Tennsco
                  13-01-2020
                               US-2020-157147
                                                     94109
40
             GBC
                  19-01-2020
                               US-2020-146591
                                                     85254
                  20-01-2020
50
          Newell
                               US-2020-147627
                                                     72401
                  21-01-2020
                                                     33180
60
           Ultra
                               US-2020-110422
70
              GE
                  26-01-2020
                               US-2020-155502
                                                     22304
80
                  28-01-2020
      Pressboard
                               US-2020-100328
                                                     10024
90
             GBC
                  04-02-2020
                               US-2020-104808
                                                     92025
                                           Product Name
                                                           Region
                                                                       Segment \
    Message Book, Wirebound, Four 5 1/2" X 4" Form... Central
0
                                                                    Consumer
10
      Southworth 25% Cotton Granite Paper & Envelopes
                                                            South
                                                                   Home Office
20
      DAX Value U-Channel Document Frames, Easel Back
                                                             East
                                                                      Consumer
30
                Tennsco 6- and 18-Compartment Lockers
                                                             West
                                                                      Consumer
40
    GBC Standard Recycled Report Covers, Clear Pla...
                                                           West
                                                                    Consumer
50
                                             Newell 327
                                                            South
                                                                      Consumer
60
                                Ultra Door Pull Handle
                                                            South
                                                                      Consumer
70
    GE General Use Halogen Bulbs, 100 Watts, 1 Bul...
                                                         South Home Office
80
    Pressboard Covers with Storage Hooks, 9 1/2" x...
                                                           East
                                                                    Consumer
    GBC Standard Recycled Report Covers, Clear Pla...
90
                                                           West
                                                                   Corporate
     Ship Date
                      Ship Mode State/Province Sub-Category
                                                              Discount
    07-01-2020
                                                       Paper
0
                Standard Class
                                          Texas
                                                                    0.2
10
    10-01-2020
                Standard Class
                                                       Paper
                                                                    0.0
                                       Kentucky
20
    14-01-2020
                    First Class
                                       Delaware
                                                 Furnishings
                                                                    0.0
    18-01-2020
                Standard Class
                                                                    0.0
30
                                    California
                                                     Storage
40
    20-01-2020
                    First Class
                                                     Binders
                                                                    0.7
                                       Arizona
                Standard Class
                                                                    0.0
50
    26-01-2020
                                       Arkansas
                                                         Art
    23-01-2020
                  Second Class
                                                                    0.2
60
                                       Florida
                                                 Furnishings
70
    31-01-2020
                Standard Class
                                       Virginia
                                                 Furnishings
                                                                    0.0
80
    03-02-2020
                Standard Class
                                       New York
                                                     Binders
                                                                    0.2
90
    08-02-2020
                  Second Class
                                    California
                                                     Binders
                                                                    0.2
              Quantity
                            Sales
      Profit
                      2
0
      5.5512
                           16.448
10
      3.0084
                      1
                            6.540
20
      3.0814
                      2
                            9.940
30
    238.6530
                      5
                         1325.850
40
    -23.7160
                     10
                           32.340
50
      1.7901
                      3
                            6.630
      4.1028
                      3
                           25.248
60
70
     30.7818
                      3
                           62.820
                            3.928
      1.3257
80
                      1
```

[46]: store[10:11] [46]: Category City Country/Region Customer Name Manufacturer \ 10 Office Supplies Henderson United States Maria Etezadi Order Date Order ID Postal Code \ 10 06-01-2020 US-2020-167199 42420 Product Name Region Segment \ 10 Southworth 25% Cotton Granite Paper & Envelopes South Home Office Ship Mode State/Province Sub-Category Discount Profit \ Ship Date 10 10-01-2020 Standard Class Kentucky Paper 0.0 3.0084 Quantity Sales 10 1 6.54 [47]: store_categorical[:4] [47]: Manufacturer \ Category City Country/Region Customer Name 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 Order Date Order ID Postal Code 0 03-01-2020 US-2020-103800 77095 1 04-01-2020 US-2020-112326 60540 2 04-01-2020 US-2020-112326 60540 3 04-01-2020 US-2020-112326 60540 Product Name Segment \ Region 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 3 Home Office SAFCO Boltless Steel Shelving Central Ship Date Ship Mode State/Province Sub-Category 0 07-01-2020 Standard Class Texas Paper 1 08-01-2020 Standard Class Illinois Binders 2 08-01-2020 Standard Class Illinois Labels 3 08-01-2020 Standard Class Illinois Storage [48]: store_numerical[::-1]

90

6.0368

2

17.248

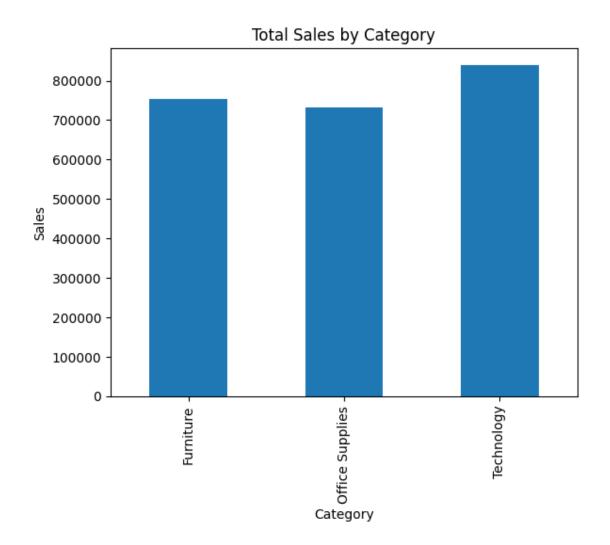
```
[48]:
           Discount Profit Quantity
                                       Sales
     10193
               0.2 -0.6048
                                       3.024
     10192
               0.0 2.7279
                                  7
                                      90.930
     10191
               0.2 -0.6048
                                  3
                                      3.024
     10190
               0.2
                     6.4750
                                  2 20.720
     10189
               0.2 19.7910
                                  3 52.776
                                  3 19.536
               0.2
                    4.8840
     3
               0.2 -64.7748
                                  3 272.736
     2
               0.2 4.2717
                                  3 11.784
               0.8 -5.4870
                                  2 3.540
     1
     0
               0.2 5.5512
                                  2 16.448
```

3 VISUALIZATION

```
[49]: # 1. Bar chart of sales by category
import matplotlib.pyplot as plt

# Group sales by category and sum
category_sales = store.groupby('Category')['Sales'].sum()

# Create a bar chart
category_sales.plot(kind='bar')
plt.title('Total Sales by Category')
plt.xlabel('Category')
plt.ylabel('Sales')
plt.show()
```



```
[50]: # 2. Histogram of profit

# Create a histogram of profit

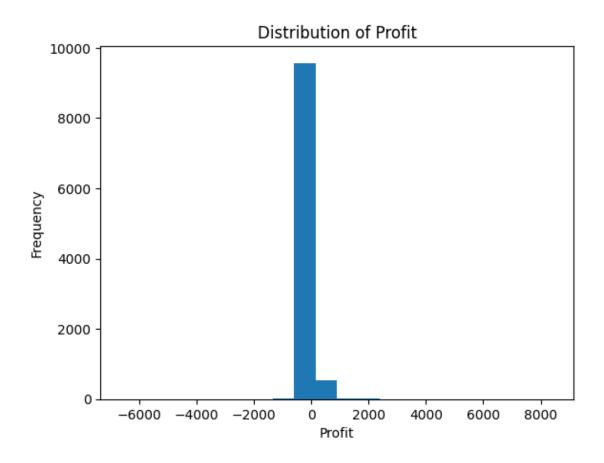
plt.hist(store['Profit'], bins=20)

plt.title('Distribution of Profit')

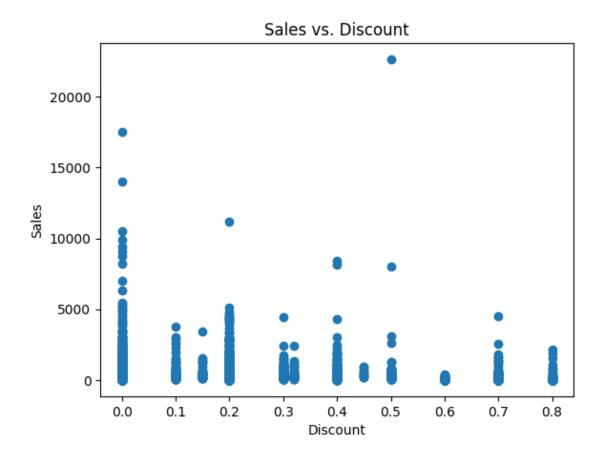
plt.xlabel('Profit')

plt.ylabel('Frequency')

plt.show()
```



```
[51]: # 3. Scatter plot of sales vs discount
    # Create a scatter plot
    plt.scatter(store['Discount'], store['Sales'])
    plt.title('Sales vs. Discount')
    plt.xlabel('Discount')
    plt.ylabel('Sales')
    plt.show()
```



```
[52]: # 4. Pie chart of sales by region
# Group sales by region and sum
region_sales = store.groupby('Region')['Sales'].sum()

# Create a pie chart
region_sales.plot(kind='pie', autopct='%1.1f%%')
plt.title('Sales Distribution by Region')
plt.ylabel('') # Remove the default ylabel
plt.show()
```

Sales Distribution by Region

