

```
import numpy as np
import pandas as pd
```

```
df=pd.read_csv('/content/
Global_Superstore2.csv',encoding='unicode_escape',engine='python')
print(df)
```

	Row ID	Order ID	Order Date	Ship Date	Ship Mode
\					
0	32298	CA-2012-124891	31-07-2012	31-07-2012	Same Day
1	26341	IN-2013-77878	05-02-2013	07-02-2013	Second Class
2	25330	IN-2013-71249	17-10-2013	18-10-2013	First Class
3	13524	ES-2013-1579342	28-01-2013	30-01-2013	First Class
4	47221	SG-2013-4320	05-11-2013	06-11-2013	Same Day
...
51285	29002	IN-2014-62366	19-06-2014	19-06-2014	Same Day
51286	35398	US-2014-102288	20-06-2014	24-06-2014	Standard Class
51287	40470	US-2013-155768	02-12-2013	02-12-2013	Same Day
51288	9596	MX-2012-140767	18-02-2012	22-02-2012	Standard Class
51289	6147	MX-2012-134460	22-05-2012	26-05-2012	Second Class

	Customer ID	Customer Name	Segment	City	\
0	RH-19495	Rick Hansen	Consumer	New York City	
1	JR-16210	Justin Ritter	Corporate	Wollongong	
2	CR-12730	Craig Reiter	Consumer	Brisbane	
3	KM-16375	Katherine Murray	Home Office	Berlin	
4	RH-9495	Rick Hansen	Consumer	Dakar	
...
51285	KE-16420	Katrina Edelman	Corporate	Kure	
51286	ZC-21910	Zuschuss Carroll	Consumer	Houston	
51287	LB-16795	Laurel Beltran	Home Office	Oxnard	
51288	RB-19795	Ross Baird	Home Office	Valinhos	
51289	MC-18100	Mick Crebagga	Consumer	Tipitapa	

	State	...	Product ID	Category Sub-
Category \				
0	New York	...	TEC-AC-10003033	Technology
Accessories				

1	New South Wales	...	FUR-CH-10003950	Furniture
Chairs				
2	Queensland	...	TEC-PH-10004664	Technology
Phones				
3	Berlin	...	TEC-PH-10004583	Technology
Phones				
4	Dakar	...	TEC-SHA-10000501	Technology
Copiers				
...
...				
51285	Hiroshima	...	OFF-FA-10000746	Office Supplies
Fasteners				
51286	Texas	...	OFF-AP-10002906	Office Supplies
Appliances				
51287	California	...	OFF-EN-10001219	Office Supplies
Envelopes				
51288	São Paulo	...	OFF-BI-10000806	Office Supplies
Binders				
51289	Managua	...	OFF-PA-10004155	Office Supplies
Paper				

Quantity \	Product Name	Sales
0	Plantronics CS510 - Over-the-Head monaural Wir...	2309.650
7		
1	Novimex Executive Leather Armchair, Black	3709.395
9		
2	Nokia Smart Phone, with Caller ID	5175.171
9		
3	Motorola Smart Phone, Cordless	2892.510
5		
4	Sharp Wireless Fax, High-Speed	2832.960
8		
...
...		
51285	Advantus Thumb Tacks, 12 Pack	65.100
5		
51286	Hoover Replacement Belt for Commercial Guardsm...	0.444
1		
51287	#10- 4 1/8" x 9 1/2" Security-Tint Envelopes	22.920
3		
51288	Acco Index Tab, Economy	13.440
2		
51289	Eaton Computer Printout Paper, 8.5 x 11	61.380
3		

	Discount	Profit	Shipping Cost	Order Priority
0	0.0	762.1845	933.57	Critical
1	0.1	-288.7650	923.63	Critical
2	0.1	919.9710	915.49	Medium

3	0.1	-96.5400	910.16	Medium
4	0.0	311.5200	903.04	Critical
...
51285	0.0	4.5000	0.01	Medium
51286	0.8	-1.1100	0.01	Medium
51287	0.0	11.2308	0.01	High
51288	0.0	2.4000	0.00	Medium
51289	0.0	1.8000	0.00	High

[51290 rows x 24 columns]

df

	Row ID	Order ID	Order Date	Ship Date	Ship Mode
\					
0	32298	CA-2012-124891	31-07-2012	31-07-2012	Same Day
1	26341	IN-2013-77878	05-02-2013	07-02-2013	Second Class
2	25330	IN-2013-71249	17-10-2013	18-10-2013	First Class
3	13524	ES-2013-1579342	28-01-2013	30-01-2013	First Class
4	47221	SG-2013-4320	05-11-2013	06-11-2013	Same Day
...
51285	29002	IN-2014-62366	19-06-2014	19-06-2014	Same Day
51286	35398	US-2014-102288	20-06-2014	24-06-2014	Standard Class
51287	40470	US-2013-155768	02-12-2013	02-12-2013	Same Day
51288	9596	MX-2012-140767	18-02-2012	22-02-2012	Standard Class
51289	6147	MX-2012-134460	22-05-2012	26-05-2012	Second Class

	Customer ID	Customer Name	Segment	City	\
0	RH-19495	Rick Hansen	Consumer	New York City	
1	JR-16210	Justin Ritter	Corporate	Wollongong	
2	CR-12730	Craig Reiter	Consumer	Brisbane	
3	KM-16375	Katherine Murray	Home Office	Berlin	
4	RH-9495	Rick Hansen	Consumer	Dakar	
...
51285	KE-16420	Katrina Edelman	Corporate	Kure	
51286	ZC-21910	Zuschuss Carroll	Consumer	Houston	
51287	LB-16795	Laurel Beltran	Home Office	Oxnard	
51288	RB-19795	Ross Baird	Home Office	Valinhos	

51289	MC-18100	Mick Crebagga	Consumer	Tipitapa
Category \	State	...	Product ID	Category Sub-
0 Accessories	New York	...	TEC-AC-10003033	Technology
1 Chairs	New South Wales	...	FUR-CH-10003950	Furniture
2 Phones	Queensland	...	TEC-PH-10004664	Technology
3 Phones	Berlin	...	TEC-PH-10004583	Technology
4 Copiers	Dakar	...	TEC-SHA-10000501	Technology
...
51285 Fasteners	Hiroshima	...	OFF-FA-10000746	Office Supplies
51286 Appliances	Texas	...	OFF-AP-10002906	Office Supplies
51287 Envelopes	California	...	OFF-EN-10001219	Office Supplies
51288 Binders	São Paulo	...	OFF-BI-10000806	Office Supplies
51289 Paper	Managua	...	OFF-PA-10004155	Office Supplies

Quantity \	Product Name	Sales
0	Plantronics CS510 - Over-the-Head monaural Wir...	2309.650
7		
1	Novimex Executive Leather Armchair, Black	3709.395
9		
2	Nokia Smart Phone, with Caller ID	5175.171
9		
3	Motorola Smart Phone, Cordless	2892.510
5		
4	Sharp Wireless Fax, High-Speed	2832.960
8		
...
...		
51285	Advantus Thumb Tacks, 12 Pack	65.100
5		
51286	Hoover Replacement Belt for Commercial Guardsm...	0.444
1		
51287	#10- 4 1/8" x 9 1/2" Security-Tint Envelopes	22.920
3		
51288	Acco Index Tab, Economy	13.440
2		
51289	Eaton Computer Printout Paper, 8.5 x 11	61.380

3

	Discount	Profit	Shipping Cost	Order	Priority
0	0.0	762.1845	933.57		Critical
1	0.1	-288.7650	923.63		Critical
2	0.1	919.9710	915.49		Medium
3	0.1	-96.5400	910.16		Medium
4	0.0	311.5200	903.04		Critical
...
51285	0.0	4.5000	0.01		Medium
51286	0.8	-1.1100	0.01		Medium
51287	0.0	11.2308	0.01		High
51288	0.0	2.4000	0.00		Medium
51289	0.0	1.8000	0.00		High

[51290 rows x 24 columns]

df.dtypes

Row ID	int64
Order ID	object
Order Date	object
Ship Date	object
Ship Mode	object
Customer ID	object
Customer Name	object
Segment	object
City	object
State	object
Country	object
Postal Code	float64
Market	object
Region	object
Product ID	object
Category	object
Sub-Category	object
Product Name	object
Sales	float64
Quantity	int64
Discount	float64
Profit	float64
Shipping Cost	float64
Order Priority	object
dtype:	object

df.head(3)

Row ID	Order ID	Order Date	Ship Date	Ship Mode	
Customer ID \					
0 32298	CA-2012-124891	31-07-2012	31-07-2012	Same Day	RH-
19495					

1	26341	IN-2013-77878	05-02-2013	07-02-2013	Second Class	JR-16210
2	25330	IN-2013-71249	17-10-2013	18-10-2013	First Class	CR-12730

	Customer Name	Segment	City	State	...	\
0	Rick Hansen	Consumer	New York City	New York	...	
1	Justin Ritter	Corporate	Wollongong	New South Wales	...	
2	Craig Reiter	Consumer	Brisbane	Queensland	...	

	Product ID	Category	Sub-Category	\
0	TEC-AC-10003033	Technology	Accessories	
1	FUR-CH-10003950	Furniture	Chairs	
2	TEC-PH-10004664	Technology	Phones	

	Product Name	Sales
Quantity \		
0	Plantronics CS510 - Over-the-Head monaural Wir...	2309.650
7		
1	Novimex Executive Leather Armchair, Black	3709.395
9		
2	Nokia Smart Phone, with Caller ID	5175.171
9		

	Discount	Profit	Shipping Cost	Order Priority
0	0.0	762.1845	933.57	Critical
1	0.1	-288.7650	923.63	Critical
2	0.1	919.9710	915.49	Medium

[3 rows x 24 columns]

df.tail(3)

	Row ID	Order ID	Order Date	Ship Date	Ship Mode
\					
51287	40470	US-2013-155768	02-12-2013	02-12-2013	Same Day
51288	9596	MX-2012-140767	18-02-2012	22-02-2012	Standard Class
51289	6147	MX-2012-134460	22-05-2012	26-05-2012	Second Class

	Customer ID	Customer Name	Segment	City
State ... \				
51287	LB-16795	Laurel Beltran	Home Office	Oxnard
California ...				
51288	RB-19795	Ross Baird	Home Office	Valinhos
Paulo ...				
51289	MC-18100	Mick Crebagga	Consumer	Tipitapa
Managua ...				

	Product ID	Category	Sub-Category	\
51287	OFF-EN-10001219	Office Supplies	Envelopes	
51288	OFF-BI-10000806	Office Supplies	Binders	
51289	OFF-PA-10004155	Office Supplies	Paper	

Discount \	Product Name	Sales	Quantity
51287 0.0	#10- 4 1/8" x 9 1/2" Security-Tint Envelopes	22.92	3
51288 0.0	Acco Index Tab, Economy	13.44	2
51289 0.0	Eaton Computer Printout Paper, 8.5 x 11	61.38	3

	Profit	Shipping	Cost	Order	Priority
51287	11.2308		0.01		High
51288	2.4000		0.00		Medium
51289	1.8000		0.00		High

```
[3 rows x 24 columns]
```

```
df.duplicated()
```

```
0      False
1      False
2      False
3      False
4      False
```

```

51285    False
51286    False
51287    False
51288    False
51289    False

```

Length: 51290, dtype: bool

```
df.isnull()
```

	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID
0	False	False	False	False	False	False
1	False	False	False	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
4	False	False	False	False	False	False

...
51285	False	False	False	False	False	False
51286	False	False	False	False	False	False
51287	False	False	False	False	False	False
51288	False	False	False	False	False	False
51289	False	False	False	False	False	False

	Customer Name	Segment	City	State	...	Product ID	Category
\ 0	False	False	False	False	...	False	False
1	False	False	False	False	...	False	False
2	False	False	False	False	...	False	False
3	False	False	False	False	...	False	False
4	False	False	False	False	...	False	False
...
51285	False	False	False	False	...	False	False
51286	False	False	False	False	...	False	False
51287	False	False	False	False	...	False	False
51288	False	False	False	False	...	False	False
51289	False	False	False	False	...	False	False

Profit	Sub-Category	Product Name	Sales	Quantity	Discount	
	\					
	0	False	False	False	False	False
	1	False	False	False	False	False
	2	False	False	False	False	False
	3	False	False	False	False	False

4	False	False	False	False	False	False
...
51285	False	False	False	False	False	False
51286	False	False	False	False	False	False
51287	False	False	False	False	False	False
51288	False	False	False	False	False	False
51289	False	False	False	False	False	False

	Shipping Cost	Order Priority
0	False	False
1	False	False
2	False	False
3	False	False
4	False	False
...
51285	False	False
51286	False	False
51287	False	False
51288	False	False
51289	False	False

[51290 rows x 24 columns]

```
import matplotlib
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 51290 entries, 0 to 51289
```

```
Data columns (total 24 columns):
```

#	Column	Non-Null Count	Dtype
0	Row ID	51290 non-null	int64
1	Order ID	51290 non-null	object
2	Order Date	51290 non-null	object
3	Ship Date	51290 non-null	object
4	Ship Mode	51290 non-null	object
5	Customer ID	51290 non-null	object
6	Customer Name	51290 non-null	object
7	Segment	51290 non-null	object
8	City	51290 non-null	object
9	State	51290 non-null	object

```

10 Country          51290 non-null object
11 Postal Code      9994 non-null float64
12 Market           51290 non-null object
13 Region           51290 non-null object
14 Product ID       51290 non-null object
15 Category         51290 non-null object
16 Sub-Category     51290 non-null object
17 Product Name     51290 non-null object
18 Sales            51290 non-null float64
19 Quantity         51290 non-null int64
20 Discount         51290 non-null float64
21 Profit           51290 non-null float64
22 Shipping Cost    51290 non-null float64
23 Order Priority    51290 non-null object
dtypes: float64(5), int64(2), object(17)
memory usage: 9.4+ MB

```

```
print(df['Postal Code'])
```

```

0      10024.0
1         NaN
2         NaN
3         NaN
4         NaN

```

```

...
51285      NaN
51286    77095.0
51287    93030.0
51288      NaN
51289      NaN

```

```
Name: Postal Code, Length: 51290, dtype: float64
```

```

a=df.dropna(axis = 1)
a.head()

```

Row ID	Order ID	Order Date	Ship Date	Ship Mode
Customer ID \				
0 32298 RH-19495	CA-2012-124891	31-07-2012	31-07-2012	Same Day
1 26341 JR-16210	IN-2013-77878	05-02-2013	07-02-2013	Second Class
2 25330 CR-12730	IN-2013-71249	17-10-2013	18-10-2013	First Class
3 13524 KM-16375	ES-2013-1579342	28-01-2013	30-01-2013	First Class
4 47221 RH-9495	SG-2013-4320	05-11-2013	06-11-2013	Same Day

\	Customer Name	Segment	City	State	...
0	Rick Hansen	Consumer	New York City	New York	...

1	Justin Ritter	Corporate	Wollongong	New South Wales	...
2	Craig Reiter	Consumer	Brisbane	Queensland	...
3	Katherine Murray	Home Office	Berlin	Berlin	...
4	Rick Hansen	Consumer	Dakar	Dakar	...

	Product ID	Category	Sub-Category	\
0	TEC-AC-10003033	Technology	Accessories	
1	FUR-CH-10003950	Furniture	Chairs	
2	TEC-PH-10004664	Technology	Phones	
3	TEC-PH-10004583	Technology	Phones	
4	TEC-SHA-10000501	Technology	Copiers	

	Product Name	Sales
Quantity \		
0	Plantronics CS510 - Over-the-Head monaural Wir...	2309.650
7		
1	Novimex Executive Leather Armchair, Black	3709.395
9		
2	Nokia Smart Phone, with Caller ID	5175.171
9		
3	Motorola Smart Phone, Cordless	2892.510
5		
4	Sharp Wireless Fax, High-Speed	2832.960
8		

	Discount	Profit	Shipping Cost	Order Priority
0	0.0	762.1845	933.57	Critical
1	0.1	-288.7650	923.63	Critical
2	0.1	919.9710	915.49	Medium
3	0.1	-96.5400	910.16	Medium
4	0.0	311.5200	903.04	Critical

[5 rows x 23 columns]

a.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 51290 entries, 0 to 51289
Data columns (total 23 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Row ID                51290 non-null  int64
1   Order ID              51290 non-null  object
2   Order Date            51290 non-null  object
3   Ship Date             51290 non-null  object
```

```

4   Ship Mode      51290 non-null object
5   Customer ID    51290 non-null object
6   Customer Name  51290 non-null object
7   Segment        51290 non-null object
8   City           51290 non-null object
9   State          51290 non-null object
10  Country        51290 non-null object
11  Market         51290 non-null object
12  Region         51290 non-null object
13  Product ID     51290 non-null object
14  Category       51290 non-null object
15  Sub-Category   51290 non-null object
16  Product Name   51290 non-null object
17  Sales          51290 non-null float64
18  Quantity       51290 non-null int64
19  Discount       51290 non-null float64
20  Profit         51290 non-null float64
21  Shipping Cost  51290 non-null float64
22  Order Priority  51290 non-null object
dtypes: float64(4), int64(2), object(17)
memory usage: 9.0+ MB

```

```
print(df['Order ID'])
```

```

0      CA-2012-124891
1      IN-2013-77878
2      IN-2013-71249
3      ES-2013-1579342
4      SG-2013-4320
...
51285   IN-2014-62366
51286   US-2014-102288
51287   US-2013-155768
51288   MX-2012-140767
51289   MX-2012-134460
Name: Order ID, Length: 51290, dtype: object

```

```
print(df[['Ship Mode', 'Order Priority']])
```

```

      Ship Mode Order Priority
0      Same Day      Critical
1    Second Class      Critical
2    First Class      Medium
3    First Class      Medium
4      Same Day      Critical
...
51285   Same Day      Medium
51286 Standard Class      Medium
51287   Same Day      High
51288 Standard Class      Medium
51289   Second Class      High

```

[51290 rows x 2 columns]

```
print(df[['Ship Mode', 'Product Name', 'Order Priority']])
```

Name \	Ship Mode	Product
0	Same Day	Plantronics CS510 - Over-the-Head monaural
Wir...		
1	Second Class	Novimex Executive Leather Armchair,
Black		
2	First Class	Nokia Smart Phone, with Caller
ID		
3	First Class	Motorola Smart Phone,
Cordless		
4	Same Day	Sharp Wireless Fax, High-
Speed		
...
..		
51285	Same Day	Advantus Thumb Tacks, 12
Pack		
51286	Standard Class	Hoover Replacement Belt for Commercial
Guardsm...		
51287	Same Day	#10- 4 1/8" x 9 1/2" Security-Tint
Envelopes		
51288	Standard Class	Acco Index Tab,
Economy		
51289	Second Class	Eaton Computer Printout Paper, 8.5 x
11		

	Order Priority
0	Critical
1	Critical
2	Medium
3	Medium
4	Critical
...	...
51285	Medium
51286	Medium
51287	High
51288	Medium
51289	High

[51290 rows x 3 columns]

```
print(df[['Order Date', 'Ship Date', 'Ship Mode']])
```

	Order Date	Ship Date	Ship Mode
0	31-07-2012	31-07-2012	Same Day
1	05-02-2013	07-02-2013	Second Class
2	17-10-2013	18-10-2013	First Class

3	28-01-2013	30-01-2013	First Class
4	05-11-2013	06-11-2013	Same Day
...
51285	19-06-2014	19-06-2014	Same Day
51286	20-06-2014	24-06-2014	Standard Class
51287	02-12-2013	02-12-2013	Same Day
51288	18-02-2012	22-02-2012	Standard Class
51289	22-05-2012	26-05-2012	Second Class

[51290 rows x 3 columns]

```
print(df[['Product Name', 'Sales', 'Quantity', 'Profit']])
```

	Product Name	Sales
Quantity \		
0	Plantronics CS510 - Over-the-Head monaural Wir...	2309.650
7		
1	Novimex Executive Leather Armchair, Black	3709.395
9		
2	Nokia Smart Phone, with Caller ID	5175.171
9		
3	Motorola Smart Phone, Cordless	2892.510
5		
4	Sharp Wireless Fax, High-Speed	2832.960
8		
...
...		
51285	Advantus Thumb Tacks, 12 Pack	65.100
5		
51286	Hoover Replacement Belt for Commercial Guardsm...	0.444
1		
51287	#10- 4 1/8" x 9 1/2" Security-Tint Envelopes	22.920
3		
51288	Acco Index Tab, Economy	13.440
2		
51289	Eaton Computer Printout Paper, 8.5 x 11	61.380
3		
	Profit	
0	762.1845	
1	-288.7650	
2	919.9710	
3	-96.5400	
4	311.5200	
...	...	
51285	4.5000	
51286	-1.1100	
51287	11.2308	
51288	2.4000	
51289	1.8000	

```
[51290 rows x 4 columns]
```

```
print(df[['Product Name', 'Sales', 'Profit']])
```

	Product Name	Sales	Profit
0	Plantronics CS510 - Over-the-Head monaural Wir...	2309.650	762.1845
1	Novimex Executive Leather Armchair, Black	3709.395	-288.7650
2	Nokia Smart Phone, with Caller ID	5175.171	919.9710
3	Motorola Smart Phone, Cordless	2892.510	-96.5400
4	Sharp Wireless Fax, High-Speed	2832.960	311.5200
...
51285	Advantus Thumb Tacks, 12 Pack	65.100	4.5000
51286	Hoover Replacement Belt for Commercial Guardsm...	0.444	-1.1100
51287	#10- 4 1/8" x 9 1/2" Security-Tint Envelopes	22.920	11.2308
51288	Acco Index Tab, Economy	13.440	2.4000
51289	Eaton Computer Printout Paper, 8.5 x 11	61.380	1.8000

```
[51290 rows x 3 columns]
```

```
print(df[['Country', 'Sales', 'Profit']])
```

	Country	Sales	Profit
0	United States	2309.650	762.1845
1	Australia	3709.395	-288.7650
2	Australia	5175.171	919.9710
3	Germany	2892.510	-96.5400
4	Senegal	2832.960	311.5200
...
51285	Japan	65.100	4.5000
51286	United States	0.444	-1.1100
51287	United States	22.920	11.2308
51288	Brazil	13.440	2.4000
51289	Nicaragua	61.380	1.8000

```
[51290 rows x 3 columns]
```

```
print(df[['Category', 'Discount']])
```

	Category	Discount
0	Technology	0.0
1	Furniture	0.1
2	Technology	0.1
3	Technology	0.1
4	Technology	0.0
...
51285	Office Supplies	0.0
51286	Office Supplies	0.8
51287	Office Supplies	0.0
51288	Office Supplies	0.0
51289	Office Supplies	0.0

[51290 rows x 2 columns]

```
print(df[df['Category']=='Furniture'])
```

	Row ID	Order ID	Order Date	Ship Date	Ship Mode
\					
1	26341	IN-2013-77878	05-02-2013	07-02-2013	Second Class
6	30570	IN-2011-81826	07-11-2011	09-11-2011	First Class
7	31192	IN-2012-86369	14-04-2012	18-04-2012	Standard Class
9	40936	CA-2012-116638	28-01-2012	31-01-2012	Second Class
11	28879	ID-2012-28402	19-04-2012	22-04-2012	First Class
...
51232	3569	US-2013-148705	16-12-2013	18-12-2013	Second Class
51239	34383	CA-2014-118773	10-02-2014	15-02-2014	Standard Class
51242	36409	CA-2012-153073	13-11-2012	13-11-2012	Same Day
51246	44081	TU-2013-7830	14-06-2013	18-06-2013	Standard Class
51254	31608	CA-2013-142902	13-09-2013	15-09-2013	Second Class

	Customer ID	Customer Name	Segment	City
State \				
1	JR-16210	Justin Ritter	Corporate	Wollongong
Wales				
6	TS-21340	Toby Swindell	Consumer	Porirua
Wellington				
7	MB-18085	Mick Brown	Consumer	Hamilton

Waikato					
9	JH-15985	Joseph Holt	Consumer	Concord	North
Carolina					
11	AJ-10780	Anthony Jacobs	Corporate	Kabul	
Kabul					
...	
...					
51232	GK-14620	Grace Kelly	Corporate	La Vega	La
Vega					
51239	TP-21415	Tom Prescott	Consumer	Houston	
Texas					
51242	HA-14905	Helen Abelman	Consumer	Chicago	
Illinois					
51246	LW-6990	Lindsay Williams	Corporate	Kayseri	
Kayseri					
51254	BP-11185	Ben Peterman	Corporate	Arvada	
Colorado					

	...	Product ID	Category	Sub-Category	\
1	...	FUR-CH-10003950	Furniture	Chairs	
6	...	FUR-CH-10004050	Furniture	Chairs	
7	...	FUR-TA-10002958	Furniture	Tables	
9	...	FUR-TA-10000198	Furniture	Tables	
11	...	FUR-TA-10001889	Furniture	Tables	
...	
51232	...	FUR-FU-10002698	Furniture	Furnishings	
51239	...	FUR-FU-10000550	Furniture	Furnishings	
51242	...	FUR-FU-10001025	Furniture	Furnishings	
51246	...	FUR-ADV-10002329	Furniture	Furnishings	
51254	...	FUR-FU-10001756	Furniture	Furnishings	

		Product Name	Sales
Quantity \			
1	Novimex Executive Leather Armchair, Black	3709.395	
9			
6	Novimex Executive Leather Armchair, Adjustable	1822.080	
4			
7	Chromcraft Conference Table, Fully Assembled	5244.840	
6			
9	Chromcraft Bull-Nose Wood Oval Conference Tabl...	4297.644	
13			
11	Bevis Conference Table, Fully Assembled	4626.150	
5			
...	
...			
51232	Deflect-0 Clock, Ergonomic	48.930	
3			
51239	Stacking Trays by OIC	3.984	
2			
51242	Eldon Image Series Desk Accessories, Clear	17.496	

```

9
51246                      Advantus Light Bulb, Ergonomic          7.476
1
51254 Eldon Expressions Desk Accessory, Wood Photo F...      15.232
1

```

	Discount	Profit	Shipping Cost	Order Priority
1	0.1	-288.7650	923.63	Critical
6	0.0	564.8400	894.77	Critical
7	0.0	996.4800	878.38	High
9	0.4	-1862.3124	865.74	Critical
11	0.0	647.5500	835.57	High
...
51232	0.5	-2.9700	0.05	Medium
51239	0.6	-2.6892	0.05	Medium
51242	0.6	-7.4358	0.05	Medium
51246	0.6	-9.5340	0.05	Medium
51254	0.2	1.7136	0.04	High

[9876 rows x 24 columns]

```
print(df[df['Order Date']=='31-07-2012'])
```

\	Row ID	Order ID	Order Date	Ship Date	Ship Mode
0	32298	CA-2012-124891	31-07-2012	31-07-2012	Same Day
331	32299	CA-2012-124891	31-07-2012	31-07-2012	Same Day
10167	7617	US-2012-149328	31-07-2012	04-08-2012	Standard Class
11701	4557	MX-2012-159380	31-07-2012	02-08-2012	First Class
11890	166	MX-2012-107146	31-07-2012	05-08-2012	Second Class
13280	7618	US-2012-149328	31-07-2012	04-08-2012	Standard Class
14304	10160	US-2012-105753	31-07-2012	05-08-2012	Second Class
16191	167	MX-2012-107146	31-07-2012	05-08-2012	Second Class
16198	29905	ID-2012-79999	31-07-2012	05-08-2012	Standard Class
17640	35877	CA-2012-164427	31-07-2012	06-08-2012	Standard Class
17840	51263	AG-2012-1940	31-07-2012	04-08-2012	Standard Class
19436	7366	US-2012-145849	31-07-2012	07-08-2012	Standard Class

20743	8809	US-2012-121097	31-07-2012	05-08-2012	Standard Class
21147	10161	US-2012-105753	31-07-2012	05-08-2012	Second Class
24212	7615	US-2012-149328	31-07-2012	04-08-2012	Standard Class
25749	29904	ID-2012-79999	31-07-2012	05-08-2012	Standard Class
26575	32300	CA-2012-124891	31-07-2012	31-07-2012	Same Day
30611	24008	IN-2012-70619	31-07-2012	04-08-2012	Second Class
33286	40028	CA-2012-111017	31-07-2012	06-08-2012	Standard Class
34114	8807	US-2012-121097	31-07-2012	05-08-2012	Standard Class
37154	7365	US-2012-145849	31-07-2012	07-08-2012	Standard Class
38452	51262	AG-2012-1940	31-07-2012	04-08-2012	Standard Class
42214	24009	IN-2012-70619	31-07-2012	04-08-2012	Second Class
43475	41779	MO-2012-7280	31-07-2012	05-08-2012	Second Class
45153	8808	US-2012-121097	31-07-2012	05-08-2012	Standard Class
45285	7616	US-2012-149328	31-07-2012	04-08-2012	Standard Class
46089	8810	US-2012-121097	31-07-2012	05-08-2012	Standard Class
46389	8806	US-2012-121097	31-07-2012	05-08-2012	Standard Class
49058	7364	US-2012-145849	31-07-2012	07-08-2012	Standard Class

	Customer ID	Customer Name	Segment	City \
0	RH-19495	Rick Hansen	Consumer	New York City
331	RH-19495	Rick Hansen	Consumer	New York City
10167	TS-21610	Troy Staebel	Consumer	Santo Domingo
11701	MS-17830	Melanie Seite	Consumer	Escuintla
11890	SL-20155	Sara Luxemburg	Home Office	Ilopango
13280	TS-21610	Troy Staebel	Consumer	Santo Domingo
14304	SL-20155	Sara Luxemburg	Home Office	Brumado
16191	SL-20155	Sara Luxemburg	Home Office	Ilopango
16198	GH-14425	Gary Hwang	Consumer	Manila
17640	AR-10405	Allen Rosenblatt	Corporate	Hattiesburg
17840	BP-1230	Benjamin Patterson	Consumer	Laghouat
19436	EB-13705	Ed Braxton	Corporate	Buenos Aires

20743	HD-14785	Harold Dahlen	Home Office	Lima
21147	SL-20155	Sara Luxemburg	Home Office	Brumado
24212	TS-21610	Troy Staebel	Consumer	Santo Domingo
25749	GH-14425	Gary Hwang	Consumer	Manila
26575	RH-19495	Rick Hansen	Consumer	New York City
30611	GM-14455	Gary Mitchum	Home Office	Hyderabad
33286	SC-20695	Steve Chapman	Corporate	Saint Louis
34114	HD-14785	Harold Dahlen	Home Office	Lima
37154	EB-13705	Ed Braxton	Corporate	Buenos Aires
38452	BP-1230	Benjamin Patterson	Consumer	Laghouat
42214	GM-14455	Gary Mitchum	Home Office	Hyderabad
43475	PN-8775	Parhena Norris	Home Office	Casablanca
45153	HD-14785	Harold Dahlen	Home Office	Lima
45285	TS-21610	Troy Staebel	Consumer	Santo Domingo
46089	HD-14785	Harold Dahlen	Home Office	Lima
46389	HD-14785	Harold Dahlen	Home Office	Lima
49058	EB-13705	Ed Braxton	Corporate	Buenos Aires

Category \	State	...	Product ID	Category Sub-
0	New York	...	TEC-AC-10003033	Technology
Accessories				
331	New York	...	FUR-TA-10002903	Furniture
Tables				
10167	Santo Domingo	...	OFF-AP-10004425	Office Supplies
Appliances				
11701	Escuintla	...	FUR-CH-10003107	Furniture
Chairs				
11890	San Salvador	...	FUR-CH-10003853	Furniture
Chairs				
13280	Santo Domingo	...	TEC-PH-10004328	Technology
Phones				
14304	Bahia	...	FUR-CH-10004151	Furniture
Chairs				
16191	San Salvador	...	FUR-CH-10003623	Furniture
Chairs				
16198	National Capital	...	OFF-BI-10004986	Office Supplies
Binders				
17640	Mississippi	...	TEC-AC-10004469	Technology
Accessories				
17840	Laghouat	...	OFF-CAR-10001577	Office Supplies
Binders				
19436	Buenos Aires	...	FUR-CH-10002291	Furniture
Chairs				
20743	Lima (city)	...	OFF-PA-10000562	Office Supplies
Paper				
21147	Bahia	...	FUR-CH-10001347	Furniture
Chairs				
24212	Santo Domingo	...	OFF-BI-10002841	Office Supplies
Binders				

25749	National Capital	...	OFF-ST-10004228	Office Supplies
Storage				
26575	New York	...	OFF-PA-10004621	Office Supplies
Paper				
30611	Telangana	...	FUR-CH-10004572	Furniture
Chairs				
33286	Missouri	...	OFF-SU-10002573	Office Supplies
Supplies				
34114	Lima (city)	...	TEC-AC-10002629	Technology
Accessories				
37154	Buenos Aires	...	OFF-FA-10000522	Office Supplies
Fasteners				
38452	Laghouat	...	TEC-STA-10003925	Technology
Machines				
42214	Telangana	...	OFF-AR-10000242	Office Supplies
Art				
43475	Grand Casablanca	...	OFF-AVE-10000909	Office Supplies
Labels				
45153	Lima (city)	...	OFF-AR-10000821	Office Supplies
Art				
45285	Santo Domingo	...	OFF-FA-10001476	Office Supplies
Fasteners				
46089	Lima (city)	...	OFF-ST-10002546	Office Supplies
Storage				
46389	Lima (city)	...	OFF-ST-10000557	Office Supplies
Storage				
49058	Buenos Aires	...	OFF-SU-10001851	Office Supplies
Supplies				

Quantity \	Product Name	Sales
0	Plantronics CS510 - Over-the-Head monaural Wir...	2309.6500
7		
331	Bevis Round Bullnose 29" High Table Top	1090.7820
7		
10167	Hoover Coffee Grinder, White	214.1760
6		
11701	Office Star Swivel Stool, Red	115.9600
1		
11890	Harbour Creations Steel Folding Chair, Red	326.3000
5		
13280	Motorola Audio Dock, Cordless	469.9200
5		
14304	Harbour Creations Steel Folding Chair, Red	130.5200
5		
16191	SAFCO Steel Folding Chair, Black	277.6000
5		
16198	Ibico 3-Hole Punch, Recycled	235.6965
9		
17640	Microsoft Sculpt Comfort Mouse	239.7000

6			
17840	Cardinal Binding Machine, Economy	199.0800	
4			
19436	Hon Bag Chairs, Set of Two	132.7200	
7			
20743	Enermax Cards & Envelopes, 8.5 x 11	184.4640	
9			
21147	SAFCO Steel Folding Chair, Black	111.0400	
5			
24212	Cardinal Binding Machine, Durable	82.0320	
3			
25749	Tenex Shelving, Industrial	60.6540	
2			
26575	Xerox 212	19.4400	
3			
30611	Harbour Creations Bag Chairs, Red	61.6200	
1			
33286	Acme 10" Easy Grip Assistive Scissors	52.5900	
3			
34114	Belkin Memory Card, Programmable	93.0480	
2			
37154	Stockwell Thumb Tacks, 12 Pack	35.1960	
7			
38452	StarTech Card Printer, Durable	163.6200	
1			
42214	BIC Pencil Sharpener, Blue	29.0100	
1			
43475	Avery Shipping Labels, Laser Printer Compatible	11.8500	
1			
45153	Sanford Highlighters, Fluorescent	13.6080	
2			
45285	OIC Paper Clips, Assorted Sizes	6.3360	
1			
46089	Eldon Folders, Single Width	13.6080	
2			
46389	Tenex Box, Blue	13.2000	
2			
49058	Stiletto Ruler, Steel	22.8480	
4			

	Discount	Profit	Shipping Cost	Order Priority
0	0.00	762.1845	933.57	Critical
331	0.40	-290.8752	348.18	Critical
10167	0.20	61.5360	32.21	High
11701	0.00	32.4600	27.40	High
11890	0.00	16.3000	26.91	Medium
13280	0.20	-17.6800	23.31	High
14304	0.60	-179.4800	21.08	High
16191	0.00	2.7000	17.62	Medium
16198	0.15	88.5465	17.61	Medium

17640	0.00	105.4680	15.46	Medium
17840	0.00	55.6800	15.17	Medium
19436	0.40	-24.3600	13.12	Low
20743	0.40	-6.1560	11.72	Medium
21147	0.60	-163.8600	11.29	High
24212	0.20	-7.1880	8.74	High
25749	0.45	-15.4860	7.73	Medium
26575	0.00	9.3312	7.24	Critical
30611	0.00	23.4000	5.23	Medium
33286	0.00	15.7770	4.17	Low
34114	0.40	9.2880	3.87	Medium
37154	0.40	-12.4040	2.97	Low
38452	0.00	40.8900	2.62	Medium
42214	0.00	5.2200	1.77	Medium
43475	0.00	1.6500	1.52	High
45153	0.40	1.5680	1.21	Medium
45285	0.20	1.8160	1.19	High
46089	0.40	-4.9920	1.05	Medium
46389	0.40	-6.4000	0.99	Medium
49058	0.40	-3.4720	0.55	Low

[29 rows x 24 columns]

df.describe()

	Row ID	Postal Code	Sales	Quantity
Discount \				
count	51290.000000	9994.000000	51290.000000	51290.000000
mean	25645.50000	55190.379428	246.490581	3.476545
std	14806.29199	32063.693350	487.565361	2.278766
min	1.00000	1040.000000	0.444000	1.000000
25%	12823.25000	23223.000000	30.758625	2.000000
50%	25645.50000	56430.500000	85.053000	3.000000
75%	38467.75000	90008.000000	251.053200	5.000000
max	51290.00000	99301.000000	22638.480000	14.000000

	Profit	Shipping Cost
count	51290.000000	51290.000000
mean	28.610982	26.375915
std	174.340972	57.296804
min	-6599.978000	0.000000
25%	0.000000	2.610000

50%	9.240000	7.790000
75%	36.810000	24.450000
max	8399.976000	933.570000