

```
In [1]: import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings("ignore")
```

```
In [2]: df=pd.read_csv("ECOMM DATA.csv")
df
```

Out[2]:

	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	...
0	32298	CA-2012-124891	31-07-2012	31-07-2012	Same Day	RH-19495	Rick Hansen	Consumer	New York City	New York	...
1	26341	IN-2013-77878	05-02-2013	07-02-2013	Second Class	JR-16210	Justin Ritter	Corporate	Wollongong	New South Wales	...
2	25330	IN-2013-71249	17-10-2013	18-10-2013	First Class	CR-12730	Craig Reiter	Consumer	Brisbane	Queensland	...
3	13524	ES-2013-1579342	28-01-2013	30-01-2013	First Class	KM-16375	Katherine Murray	Home Office	Berlin	Berlin	...
4	47221	SG-2013-4320	05-11-2013	06-11-2013	Same Day	RH-9495	Rick Hansen	Consumer	Dakar	Dakar	...
...
51285	29002	IN-2014-62366	19-06-2014	19-06-2014	Same Day	KE-16420	Katrina Edelman	Corporate	Kure	Hiroshima	...
51286	35398	US-2014-102288	20-06-2014	24-06-2014	Standard Class	ZC-21910	Zuschuss Carroll	Consumer	Houston	Texas	...
51287	40470	US-2013-155768	02-12-2013	02-12-2013	Same Day	LB-16795	Laurel Beltran	Home Office	Oxnard	California	...
51288	9596	MX-2012-140767	18-02-2012	22-02-2012	Standard Class	RB-19795	Ross Baird	Home Office	Valinhos	São Paulo	...
51289	6147	MX-2012-134460	22-05-2012	26-05-2012	Second Class	MC-18100	Mick Crebagga	Consumer	Tipitapa	Managua	...

51290 rows × 24 columns

```
In [3]: df1=pd.read_csv("ECOMM DATA (Return).csv")
df1
```

Out[3]:

	Returned	Order ID	Market	Unnamed: 3	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9
0	Yes	MX-2013-168137	LATAM	NaN						
1	Yes	US-2011-165316	LATAM	NaN						
2	Yes	ES-2013-1525878	EU	NaN						
3	Yes	CA-2013-118311	United States	NaN						
4	Yes	ES-2011-1276768	EU	NaN						
...
1168	Yes	ES-2013-2639112	EU	NaN						
1169	Yes	CA-2014-134194	United States	NaN						
1170	Yes	ES-2012-3246286	EU	NaN						
1171	Yes	ES-2012-4379168	EU	NaN						
1172	Yes	CA-2014-168193	United States	NaN						

1173 rows × 256 columns

```
In [4]: df2=pd.read_csv("ECOMM DATA (People).csv")
df2
```

Out[4]:

	Person	Region	Unnamed: 2	Unnamed: 3	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8
0	Anna Andreadi	Central	NaN						
1	Chuck Magee	South	NaN						
2	Kelly Williams	East	NaN						
3	Matt Collister	West	NaN						
4	Deborah Brumfield	Africa	NaN						
5	Larry Hughes	EMEA	NaN						
6	Nicole Hansen	Canada	NaN						
7	Giulietta Dorch	Caribbean	NaN						
8	Nora Preis	Central Asia	NaN						
9	Jack Lebron	North	NaN						
10	Shirley Daniels	North Asia	NaN						
11	Anthony Jacobs	Oceania	NaN						
12	Alejandro Ballentine	Southeast Asia	NaN						

13 rows × 256 columns

In [5]: df2.iloc[:, 0:2]

Out[5]:

	Person	Region
0	Anna Andreadi	Central
1	Chuck Magee	South
2	Kelly Williams	East
3	Matt Collister	West
4	Deborah Brumfield	Africa
5	Larry Hughes	EMEA
6	Nicole Hansen	Canada
7	Giulietta Dorch	Caribbean
8	Nora Preis	Central Asia
9	Jack Lebron	North
10	Shirley Daniels	North Asia
11	Anthony Jacobs	Oceania
12	Alejandro Ballentine	Southeast Asia

In [6]: `df1.iloc[:, 0:3]`

Out[6]:

	Returned	Order ID	Market
0	Yes	MX-2013-168137	LATAM
1	Yes	US-2011-165316	LATAM
2	Yes	ES-2013-1525878	EU
3	Yes	CA-2013-118311	United States
4	Yes	ES-2011-1276768	EU
...
1168	Yes	ES-2013-2639112	EU
1169	Yes	CA-2014-134194	United States
1170	Yes	ES-2012-3246286	EU
1171	Yes	ES-2012-4379168	EU
1172	Yes	CA-2014-168193	United States

1173 rows × 3 columns

In [7]: `df.isnull().sum()`

Out[7]:

Row ID	0
Order ID	0
Order Date	0
Ship Date	0
Ship Mode	0
Customer ID	0
Customer Name	0
Segment	0
City	0
State	0
Country	0
Postal Code	41296
Market	0
Region	0
Product ID	0
Category	0
Sub-Category	0
Product Name	0
Sales	0
Quantity	0
Discount	0
Profit	0
Shipping Cost	0
Order Priority	0

dtype: int64

In [8]: `df.drop("Postal Code", axis=1, inplace=True)`

In [9]: `df`

Out[9]:	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	...
0	32298	CA-2012-124891	31-07-2012	31-07-2012	Same Day	RH-19495	Rick Hansen	Consumer	New York City	New York	...
1	26341	IN-2013-77878	05-02-2013	07-02-2013	Second Class	JR-16210	Justin Ritter	Corporate	Wollongong	New South Wales	...
2	25330	IN-2013-71249	17-10-2013	18-10-2013	First Class	CR-12730	Craig Reiter	Consumer	Brisbane	Queensland	...
3	13524	ES-2013-1579342	28-01-2013	30-01-2013	First Class	KM-16375	Katherine Murray	Home Office	Berlin	Berlin	...
4	47221	SG-2013-4320	05-11-2013	06-11-2013	Same Day	RH-9495	Rick Hansen	Consumer	Dakar	Dakar	...
...
51285	29002	IN-2014-62366	19-06-2014	19-06-2014	Same Day	KE-16420	Katrina Edelman	Corporate	Kure	Hiroshima	...
51286	35398	US-2014-102288	20-06-2014	24-06-2014	Standard Class	ZC-21910	Zuschuss Carroll	Consumer	Houston	Texas	...
51287	40470	US-2013-155768	02-12-2013	02-12-2013	Same Day	LB-16795	Laurel Beltran	Home Office	Oxnard	California	...
51288	9596	MX-2012-140767	18-02-2012	22-02-2012	Standard Class	RB-19795	Ross Baird	Home Office	Valinhos	São Paulo	...
51289	6147	MX-2012-134460	22-05-2012	26-05-2012	Second Class	MC-18100	Mick Crebagga	Consumer	Tipitapa	Managua	...

51290 rows × 23 columns

In [10]: df.isnull().sum()

```
Out[10]: Row ID      0  
          Order ID     0  
          Order Date   0  
          Ship Date    0  
          Ship Mode    0  
          Customer ID  0  
          Customer Name 0  
          Segment       0  
          City          0  
          State         0  
          Country        0  
          Market         0  
          Region         0  
          Product ID    0  
          Category       0  
          Sub-Category   0  
          Product Name   0  
          Sales          0  
          Quantity       0  
          Discount       0  
          Profit          0  
          Shipping Cost  0  
          Order Priority 0  
          dtype: int64
```

```
In [11]: df["Product Name"].value_counts()
```

```
Out[11]: Product Name  
          Staples                227  
          Cardinal Index Tab, Clear 92  
          Eldon File Cart, Single Width 90  
          Rogers File Cart, Single Width 84  
          Ibico Index Tab, Clear     83  
          ...  
          Penpower WorldCard Pro Card Scanner 1  
          Cisco SPA525G2 5-Line IP Phone      1  
          Park Ridge Embossed Executive Business Envelopes 1  
          Rediform S.O.S. Phone Message Books 1  
          Eureka Disposable Bags for Sanitaire Vibra Groomer I Upright Vac 1  
          Name: count, Length: 3788, dtype: int64
```

```
In [12]: df["Product Name"].max()
```

```
Out[12]: 'netTALK DUO VoIP Telephone Service'
```

```
In [13]: df["Product Name"].min()
```

```
Out[13]: '"While you Were Out" Message Book, One Form per Page'
```

```
In [14]: df["Sales"].max()
```

```
Out[14]: 22638.48
```

```
In [15]: df["Sales"].sum()
```

```
Out[15]: 12642501.909880001
```

```
In [16]: df
```

Out[16]:	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	...
0	32298	CA-2012-124891	31-07-2012	31-07-2012	Same Day	RH-19495	Rick Hansen	Consumer	New York City	New York	...
1	26341	IN-2013-77878	05-02-2013	07-02-2013	Second Class	JR-16210	Justin Ritter	Corporate	Wollongong	New South Wales	...
2	25330	IN-2013-71249	17-10-2013	18-10-2013	First Class	CR-12730	Craig Reiter	Consumer	Brisbane	Queensland	...
3	13524	ES-2013-1579342	28-01-2013	30-01-2013	First Class	KM-16375	Katherine Murray	Home Office	Berlin	Berlin	...
4	47221	SG-2013-4320	05-11-2013	06-11-2013	Same Day	RH-9495	Rick Hansen	Consumer	Dakar	Dakar	...
...
51285	29002	IN-2014-62366	19-06-2014	19-06-2014	Same Day	KE-16420	Katrina Edelman	Corporate	Kure	Hiroshima	...
51286	35398	US-2014-102288	20-06-2014	24-06-2014	Standard Class	ZC-21910	Zuschuss Carroll	Consumer	Houston	Texas	...
51287	40470	US-2013-155768	02-12-2013	02-12-2013	Same Day	LB-16795	Laurel Beltran	Home Office	Oxnard	California	...
51288	9596	MX-2012-140767	18-02-2012	22-02-2012	Standard Class	RB-19795	Ross Baird	Home Office	Valinhos	São Paulo	...
51289	6147	MX-2012-134460	22-05-2012	26-05-2012	Second Class	MC-18100	Mick Crebagga	Consumer	Tipitapa	Managua	...

51290 rows × 23 columns

In [17]: df.describe()

Out[17]:

	Row ID	Sales	Quantity	Discount	Profit	Shipping Cost
count	51290.00000	51290.000000	51290.000000	51290.000000	51290.000000	51290.000000
mean	25645.50000	246.490581	3.476545	0.142908	28.610982	26.375915
std	14806.29199	487.565361	2.278766	0.212280	174.340972	57.296804
min	1.00000	0.444000	1.000000	0.000000	-6599.978000	0.000000
25%	12823.25000	30.758625	2.000000	0.000000	0.000000	2.610000
50%	25645.50000	85.053000	3.000000	0.000000	9.240000	7.790000
75%	38467.75000	251.053200	5.000000	0.200000	36.810000	24.450000
max	51290.00000	22638.480000	14.000000	0.850000	8399.976000	933.570000

In [19]:

```
df["Order Date"] = pd.to_datetime(df["Order Date"])
df
```

Out[19]:	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	...
0	32298	CA-2012-124891	2012-07-31	31-07-2012	Same Day	RH-19495	Rick Hansen	Consumer	New York City	New York	...
1	26341	IN-2013-77878	2013-02-05	07-02-2013	Second Class	JR-16210	Justin Ritter	Corporate	Wollongong	New South Wales	...
2	25330	IN-2013-71249	2013-10-17	18-10-2013	First Class	CR-12730	Craig Reiter	Consumer	Brisbane	Queensland	...
3	13524	ES-2013-1579342	2013-01-28	30-01-2013	First Class	KM-16375	Katherine Murray	Home Office	Berlin	Berlin	...
4	47221	SG-2013-4320	2013-11-05	06-11-2013	Same Day	RH-9495	Rick Hansen	Consumer	Dakar	Dakar	...
...
51285	29002	IN-2014-62366	2014-06-19	19-06-2014	Same Day	KE-16420	Katrina Edelman	Corporate	Kure	Hiroshima	...
51286	35398	US-2014-102288	2014-06-20	24-06-2014	Standard Class	ZC-21910	Zuschuss Carroll	Consumer	Houston	Texas	...
51287	40470	US-2013-155768	2013-12-02	02-12-2013	Same Day	LB-16795	Laurel Beltran	Home Office	Oxnard	California	...
51288	9596	MX-2012-140767	2012-02-18	22-02-2012	Standard Class	RB-19795	Ross Baird	Home Office	Valinhos	São Paulo	...
51289	6147	MX-2012-134460	2012-05-22	26-05-2012	Second Class	MC-18100	Mick Crebaga	Consumer	Tipitapa	Managua	...

51290 rows × 23 columns

In [20]: `df["Ship Date"] = pd.to_datetime(df["Ship Date"])`

Out[20]:	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	...
0	32298	CA-2012-124891	2012-07-31	2012-07-31	Same Day	RH-19495	Rick Hansen	Consumer	New York City	New York	...
1	26341	IN-2013-77878	2013-02-05	2013-02-07	Second Class	JR-16210	Justin Ritter	Corporate	Wollongong	New South Wales	...
2	25330	IN-2013-71249	2013-10-17	2013-10-18	First Class	CR-12730	Craig Reiter	Consumer	Brisbane	Queensland	...
3	13524	ES-2013-1579342	2013-01-28	2013-01-30	First Class	KM-16375	Katherine Murray	Home Office	Berlin	Berlin	...
4	47221	SG-2013-4320	2013-11-05	2013-11-06	Same Day	RH-9495	Rick Hansen	Consumer	Dakar	Dakar	...
...
51285	29002	IN-2014-62366	2014-06-19	2014-06-19	Same Day	KE-16420	Katrina Edelman	Corporate	Kure	Hiroshima	...
51286	35398	US-2014-102288	2014-06-20	2014-06-24	Standard Class	ZC-21910	Zuschuss Carroll	Consumer	Houston	Texas	...
51287	40470	US-2013-155768	2013-12-02	2013-12-02	Same Day	LB-16795	Laurel Beltran	Home Office	Oxnard	California	...
51288	9596	MX-2012-140767	2012-02-18	2012-02-22	Standard Class	RB-19795	Ross Baird	Home Office	Valinhos	São Paulo	...
51289	6147	MX-2012-134460	2012-05-22	2012-05-26	Second Class	MC-18100	Mick Crebagga	Consumer	Tipitapa	Managua	...

51290 rows × 23 columns

In [21]: `df["year"] = df["Order Date"].dt.year
df`

Out[21]:	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	...
0	32298	CA-2012-124891	2012-07-31	2012-07-31	Same Day	RH-19495	Rick Hansen	Consumer	New York City	New York	...
1	26341	IN-2013-77878	2013-02-05	2013-02-07	Second Class	JR-16210	Justin Ritter	Corporate	Wollongong	New South Wales	...
2	25330	IN-2013-71249	2013-10-17	2013-10-18	First Class	CR-12730	Craig Reiter	Consumer	Brisbane	Queensland	...
3	13524	ES-2013-1579342	2013-01-28	2013-01-30	First Class	KM-16375	Katherine Murray	Home Office	Berlin	Berlin	...
4	47221	SG-2013-4320	2013-11-05	2013-11-06	Same Day	RH-9495	Rick Hansen	Consumer	Dakar	Dakar	...
...
51285	29002	IN-2014-62366	2014-06-19	2014-06-19	Same Day	KE-16420	Katrina Edelman	Corporate	Kure	Hiroshima	...
51286	35398	US-2014-102288	2014-06-20	2014-06-24	Standard Class	ZC-21910	Zuschuss Carroll	Consumer	Houston	Texas	...
51287	40470	US-2013-155768	2013-12-02	2013-12-02	Same Day	LB-16795	Laurel Beltran	Home Office	Oxnard	California	...
51288	9596	MX-2012-140767	2012-02-18	2012-02-22	Standard Class	RB-19795	Ross Baird	Home Office	Valinhos	São Paulo	...
51289	6147	MX-2012-134460	2012-05-22	2012-05-26	Second Class	MC-18100	Mick Crebagga	Consumer	Tipitapa	Managua	...

51290 rows × 24 columns

In [22]: df1=df[['year', "Sales"]]
df1

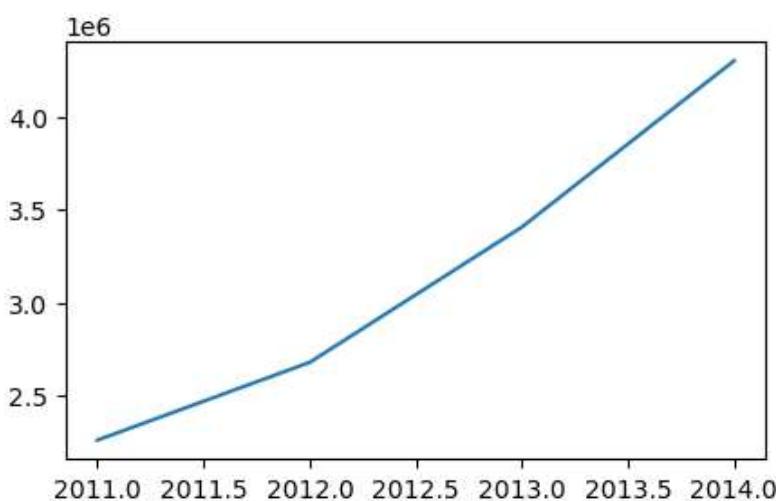
```
Out[22]:   year      Sales
          0  2012  2309.650
          1  2013  3709.395
          2  2013  5175.171
          3  2013  2892.510
          4  2013  2832.960
          ...
          ...
          ...
          51285 2014    65.100
          51286 2014     0.444
          51287 2013   22.920
          51288 2012   13.440
          51289 2012   61.380
```

51290 rows × 2 columns

```
In [24]: year_sales=df1.groupby("year").sum()
year_sales
```

```
Out[24]:      Sales
      year
      2011  2.259451e+06
      2012  2.677439e+06
      2013  3.405746e+06
      2014  4.299866e+06
```

```
In [25]: # Line plot
plt.figure(figsize=(5,3))
plt.plot(year_sales.index,year_sales["Sales"])
plt.show()
```



```
In [26]: df2=df[["year","Profit"]]
df2
```

Out[26]:

	year	Profit
0	2012	762.1845
1	2013	-288.7650
2	2013	919.9710
3	2013	-96.5400
4	2013	311.5200
...
51285	2014	4.5000
51286	2014	-1.1100
51287	2013	11.2308
51288	2012	2.4000
51289	2012	1.8000

51290 rows × 2 columns

In [27]:

```
year_profit=df2.groupby("year").sum()
year_profit
```

Out[27]:

	Profit
year	
2011	248940.81154
2012	307415.27910
2013	406935.23018
2014	504165.97046

In [28]:

```
from matplotlib import style
style.use("ggplot")
plt.figure(figsize=(5,4))
plt.plot(year_sales.index,year_sales["Sales"],color="g",linewidth=2,marker="*",ms=10,mec="k",mfc="w",label="sales",linestyle=":")
plt.plot(year_profit.index,year_profit["Profit"],color="m",linewidth=2,marker="o",ms=10,mec="k",mfc="y",label="profit",linestyle="--")
plt.xlabel("years",fontsize=16, color="r")
plt.ylabel("sales & profit",fontsize=16, color="r")
plt.title("year wise sales & profit",fontsize=20, color="b")
plt.legend(loc=2)
plt.grid(color="k")
plt.show()
```



```
In [29]: # subplot
plt.subplot(2,1,1)

plt.plot(year_sales.index,year_sales["Sales"],color="g",linewidth=2,marker="o",ms=10,mec="k",mew=2,
         label="sales",linestyle=":")
plt.ylabel("sales",fontsize=16, color="r")

plt.title("year wise sales & profit",fontsize=20, color="b")
plt.grid(color="k")
plt.legend(loc=2)

plt.subplot(2,1,2)
plt.plot(year_profit.index,year_profit["Profit"],color="b",linewidth=2,marker="^",ms=10,mec="k",
         label="profit",linestyle=":")
plt.legend(loc=2)

plt.ylabel("profit",fontsize=16, color="r")

plt.xlabel("years",fontsize=16, color="r")
plt.grid(color="k")
plt.show()
```



```
In [30]: df2=df[["Sales","Profit"]]
df2
```

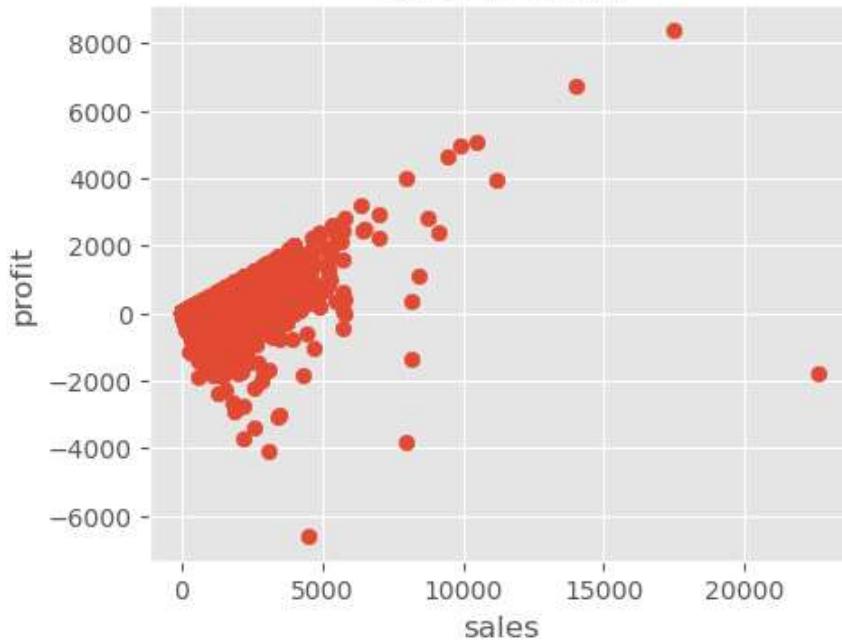
```
Out[30]:
```

	Sales	Profit
0	2309.650	762.1845
1	3709.395	-288.7650
2	5175.171	919.9710
3	2892.510	-96.5400
4	2832.960	311.5200
...
51285	65.100	4.5000
51286	0.444	-1.1100
51287	22.920	11.2308
51288	13.440	2.4000
51289	61.380	1.8000

51290 rows × 2 columns

```
In [31]: plt.figure(figsize=(5,4))
plt.scatter(df2.Sales,df2.Profit)
plt.xlabel("sales")
plt.ylabel("profit")
plt.title("sales vs profit")
plt.show()
```

sales vs profit



In [32]: df

Out[32]:	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	...
0	32298	CA-2012-124891	2012-07-31	2012-07-31	Same Day	RH-19495	Rick Hansen	Consumer	New York City	New York	...
1	26341	IN-2013-77878	2013-02-05	2013-02-07	Second Class	JR-16210	Justin Ritter	Corporate	Wollongong	New South Wales	...
2	25330	IN-2013-71249	2013-10-17	2013-10-18	First Class	CR-12730	Craig Reiter	Consumer	Brisbane	Queensland	...
3	13524	ES-2013-1579342	2013-01-28	2013-01-30	First Class	KM-16375	Katherine Murray	Home Office	Berlin	Berlin	...
4	47221	SG-2013-4320	2013-11-05	2013-11-06	Same Day	RH-9495	Rick Hansen	Consumer	Dakar	Dakar	...
...
51285	29002	IN-2014-62366	2014-06-19	2014-06-19	Same Day	KE-16420	Katrina Edelman	Corporate	Kure	Hiroshima	...
51286	35398	US-2014-102288	2014-06-20	2014-06-24	Standard Class	ZC-21910	Zuschuss Carroll	Consumer	Houston	Texas	...
51287	40470	US-2013-155768	2013-12-02	2013-12-02	Same Day	LB-16795	Laurel Beltran	Home Office	Oxnard	California	...
51288	9596	MX-2012-140767	2012-02-18	2012-02-22	Standard Class	RB-19795	Ross Baird	Home Office	Valinhos	São Paulo	...
51289	6147	MX-2012-134460	2012-05-22	2012-05-26	Second Class	MC-18100	Mick Crebagga	Consumer	Tipitapa	Managua	...

51290 rows × 24 columns

In [33]: `df["cp"] = df["Sales"] - df["Profit"]`
`df`

Out[33]:	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	...
0	32298	CA-2012-124891	2012-07-31	2012-07-31	Same Day	RH-19495	Rick Hansen	Consumer	New York City	New York	...
1	26341	IN-2013-77878	2013-02-05	2013-02-07	Second Class	JR-16210	Justin Ritter	Corporate	Wollongong	New South Wales	...
2	25330	IN-2013-71249	2013-10-17	2013-10-18	First Class	CR-12730	Craig Reiter	Consumer	Brisbane	Queensland	...
3	13524	ES-2013-1579342	2013-01-28	2013-01-30	First Class	KM-16375	Katherine Murray	Home Office	Berlin	Berlin	...
4	47221	SG-2013-4320	2013-11-05	2013-11-06	Same Day	RH-9495	Rick Hansen	Consumer	Dakar	Dakar	...
...
51285	29002	IN-2014-62366	2014-06-19	2014-06-19	Same Day	KE-16420	Katrina Edelman	Corporate	Kure	Hiroshima	...
51286	35398	US-2014-102288	2014-06-20	2014-06-24	Standard Class	ZC-21910	Zuschuss Carroll	Consumer	Houston	Texas	...
51287	40470	US-2013-155768	2013-12-02	2013-12-02	Same Day	LB-16795	Laurel Beltran	Home Office	Oxnard	California	...
51288	9596	MX-2012-140767	2012-02-18	2012-02-22	Standard Class	RB-19795	Ross Baird	Home Office	Valinhos	São Paulo	...
51289	6147	MX-2012-134460	2012-05-22	2012-05-26	Second Class	MC-18100	Mick Crebagga	Consumer	Tipitapa	Managua	...

51290 rows × 25 columns

In [34]: df1=df[['Sales", "Profit", "cp"]]
df1

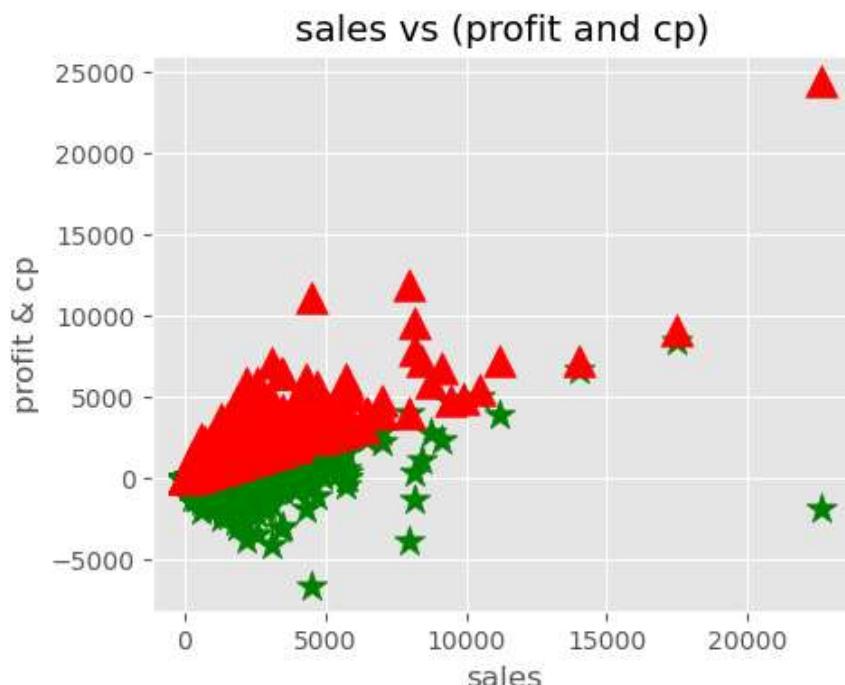
Out[34]:

	Sales	Profit	cp
0	2309.650	762.1845	1547.4655
1	3709.395	-288.7650	3998.1600
2	5175.171	919.9710	4255.2000
3	2892.510	-96.5400	2989.0500
4	2832.960	311.5200	2521.4400
...
51285	65.100	4.5000	60.6000
51286	0.444	-1.1100	1.5540
51287	22.920	11.2308	11.6892
51288	13.440	2.4000	11.0400
51289	61.380	1.8000	59.5800

51290 rows × 3 columns

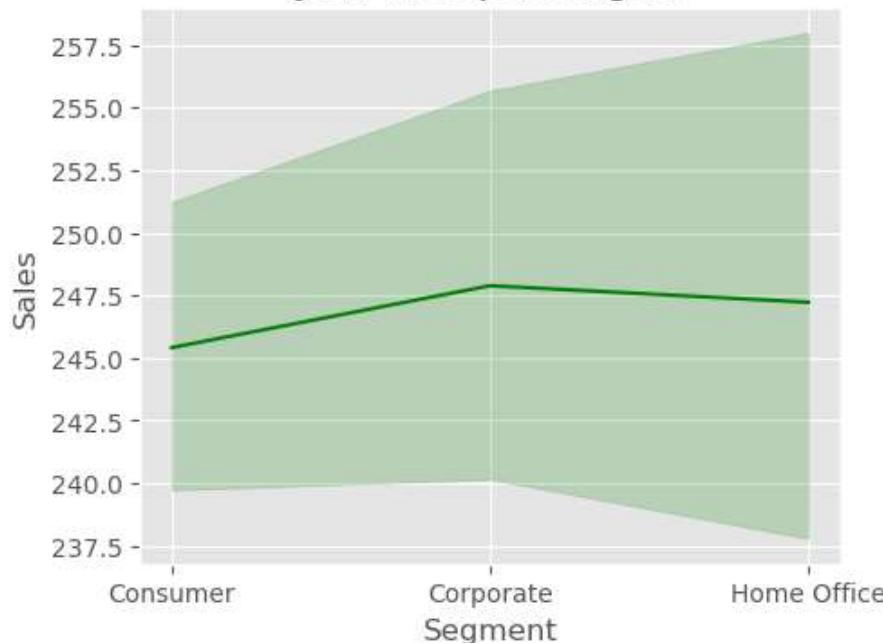
In [35]:

```
plt.figure(figsize=(5,4))
plt.scatter(df1.Sales,df1.Profit,marker="*",color="g",s=160)
plt.scatter(df1.Sales,df1.cp,marker="^",color="r",s=160)
plt.xlabel("sales")
plt.ylabel("profit & cp")
plt.title("sales vs (profit and cp)")
plt.show()
```

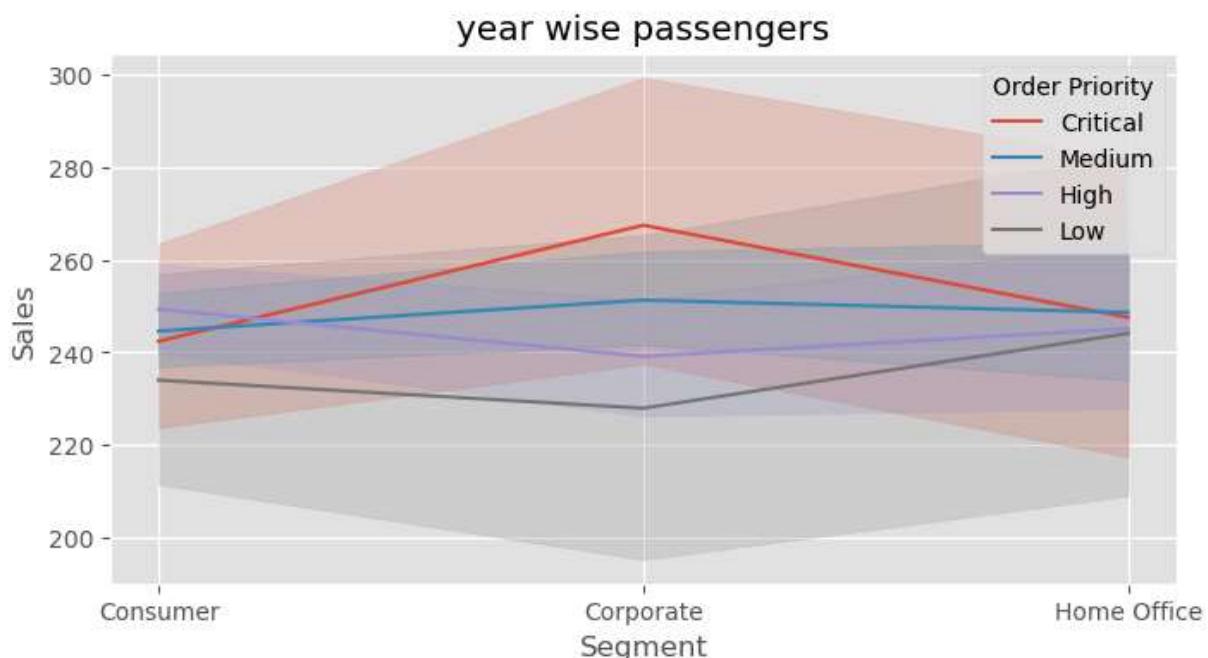


In [42]:

```
# Line plot
plt.figure(figsize=(5,4))
sns.lineplot(x="Segment",y="Sales",data=df,color="g")
plt.title("year wise passengers")
plt.show()
```

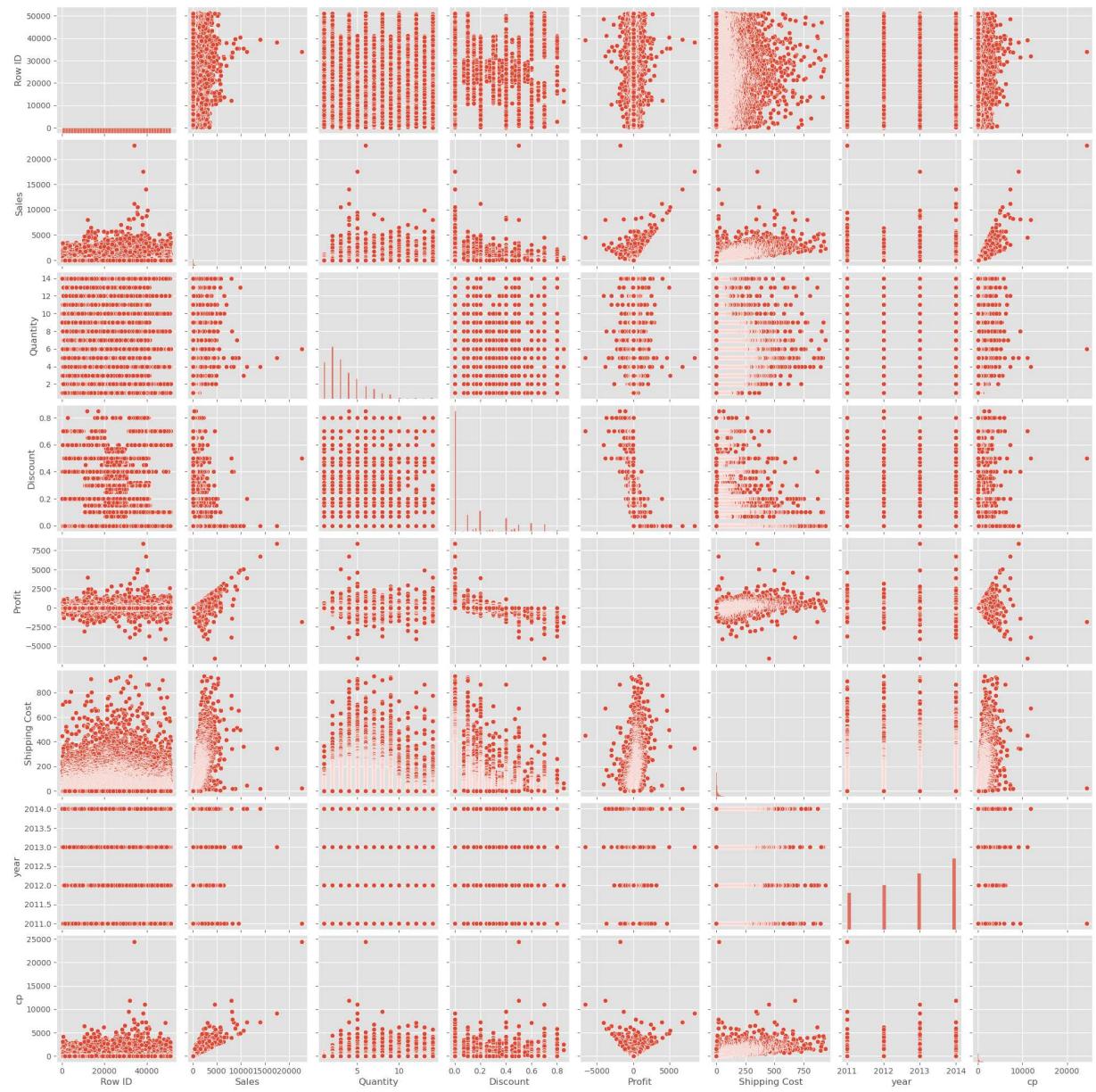
year wise passengers

```
In [44]: # Line plot
plt.figure(figsize=(8,4))
sns.lineplot(x="Segment",y="Sales",data=df,color="g",hue="Order Priority")
plt.title("year wise passengers")
plt.show()
```



```
In [46]: sns.pairplot(df)
plt.show()
```

Sales Data Analysis



In [47]: df

Out[47]:	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	...
0	32298	CA-2012-124891	2012-07-31	2012-07-31	Same Day	RH-19495	Rick Hansen	Consumer	New York City	New York	...
1	26341	IN-2013-77878	2013-02-05	2013-02-07	Second Class	JR-16210	Justin Ritter	Corporate	Wollongong	New South Wales	...
2	25330	IN-2013-71249	2013-10-17	2013-10-18	First Class	CR-12730	Craig Reiter	Consumer	Brisbane	Queensland	...
3	13524	ES-2013-1579342	2013-01-28	2013-01-30	First Class	KM-16375	Katherine Murray	Home Office	Berlin	Berlin	...
4	47221	SG-2013-4320	2013-11-05	2013-11-06	Same Day	RH-9495	Rick Hansen	Consumer	Dakar	Dakar	...
...
51285	29002	IN-2014-62366	2014-06-19	2014-06-19	Same Day	KE-16420	Katrina Edelman	Corporate	Kure	Hiroshima	...
51286	35398	US-2014-102288	2014-06-20	2014-06-24	Standard Class	ZC-21910	Zuschuss Carroll	Consumer	Houston	Texas	...
51287	40470	US-2013-155768	2013-12-02	2013-12-02	Same Day	LB-16795	Laurel Beltran	Home Office	Oxnard	California	...
51288	9596	MX-2012-140767	2012-02-18	2012-02-22	Standard Class	RB-19795	Ross Baird	Home Office	Valinhos	São Paulo	...
51289	6147	MX-2012-134460	2012-05-22	2012-05-26	Second Class	MC-18100	Mick Crebagga	Consumer	Tipitapa	Managua	...

51290 rows × 25 columns

◀ ▶

In []: