

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
```

```
df=pd.read_csv('business.retailsales.csv')
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1775 entries, 0 to 1774
Data columns (total 6 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Product Type          1767 non-null   object
1   Net Quantity          1775 non-null   int64
2   Gross Sales           1775 non-null   float64
3   Discounts              1775 non-null   float64
4   Returns                1775 non-null   float64
5   Total Net Sales        1775 non-null   float64
dtypes: float64(4), int64(1), object(1)
memory usage: 83.3+ KB
```

```
df.describe()
```

	Net Quantity	Gross Sales	Discounts	Returns	Total Net
Sales count	1775.000000	1775.000000	1775.000000	1775.000000	
mean	3.712676	199.671746	-6.317623	-5.385437	
std	6.243078	464.880638	20.903517	46.654269	
min	-1.000000	0.000000	-594.000000	-1609.000000	-
25%	1.000000	48.000000	-6.000000	0.000000	
50%	2.000000	100.000000	0.000000	0.000000	
75%	4.000000	185.500000	0.000000	0.000000	
max	96.000000	14935.000000	0.000000	0.000000	

```
df.dropna()
```

	Product Type	Net Quantity	Gross Sales	Discounts	Returns
0	Art & Sculpture	34	14935.0	-594.00	-1609.00
1	Basket	13	3744.0	-316.80	0.00

2	Basket	12	3825.0	-201.60	-288.00
3	Basket	17	3035.0	-63.25	0.00
4	Art & Sculpture	47	2696.8	-44.16	0.00
...
1770	Kitchen	0	28.0	-2.81	-25.19
1771	Jewelry	0	28.0	0.00	-28.00
1772	Basket	0	116.0	-23.20	-92.80
1773	Kitchen	0	16.5	0.00	-16.50
1774	Kitchen	-1	0.0	0.00	-106.25

	Total Net Sales
0	12732.00
1	3427.20
2	3335.40
3	2971.75
4	2652.64
...	...
1770	0.00
1771	0.00
1772	0.00
1773	0.00
1774	-106.25

[1767 rows x 6 columns]

```
sales = df.groupby('Product Type').sum().sort_values('Gross Sales').reset_index()
sales
```

	Product Type	Net Quantity	Gross Sales	Discounts	Returns	\
0	Gift Baskets	1	19.50	0.00	0.00	
1	Easter	1	38.00	-3.80	0.00	
2	Textiles	43	1889.00	-112.90	-97.00	
3	Furniture	27	2034.00	-169.04	0.00	
4	One-of-a-Kind	12	2180.00	-71.99	0.00	
5	Fair Trade Gifts	110	2258.00	-53.33	0.00	
6	Skin Care	101	2609.50	-37.70	0.00	
7	Music	98	2643.50	-82.19	-142.41	
8	Recycled Art	99	3792.80	-88.64	0.00	

9	Kids	140	3838.00	-116.66	0.00
10	Accessories	84	3892.40	-107.02	0.00
11	Soapstone	199	4795.50	-96.91	-69.50
12	Christmas	575	15476.00	-345.19	-670.00
13	Kitchen	809	16096.00	-431.11	-328.07
14	Home Decor	404	27114.55	-991.21	-423.35
15	Jewelry	991	31048.00	-965.85	-509.20
16	Art & Sculpture	1427	90316.60	-2955.82	-2879.93
17	Basket	1461	143815.50	-4584.42	-4439.69

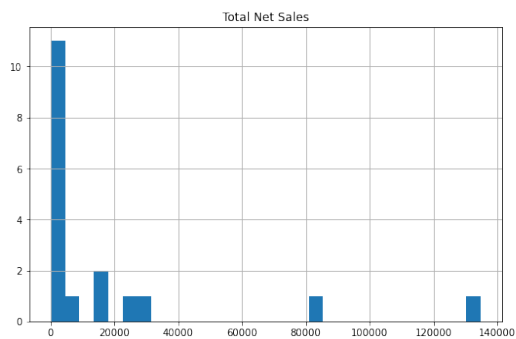
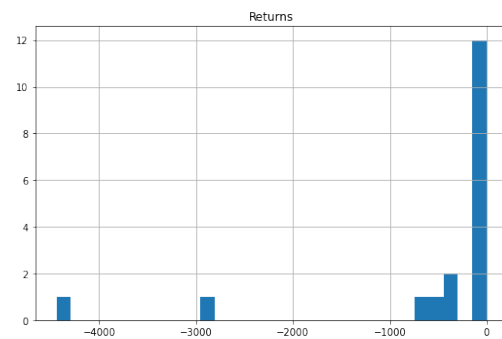
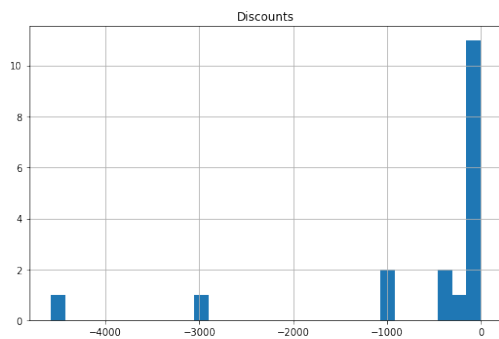
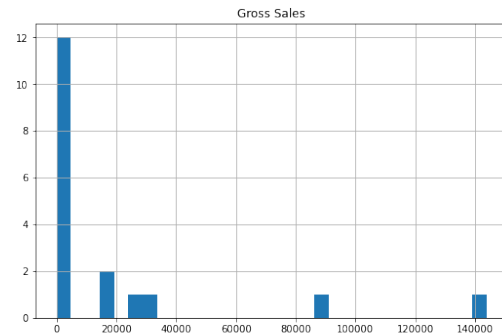
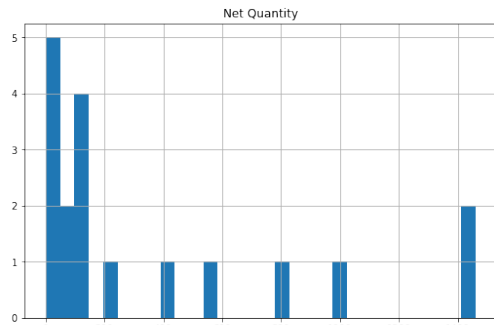
	Total Net Sales
0	19.50
1	34.20
2	1679.10
3	1864.96
4	2108.01
5	2204.67
6	2571.80
7	2418.90
8	3704.16
9	3721.34
10	3785.38
11	4629.09
12	14460.81
13	15336.82
14	25699.99
15	29572.95
16	84480.85
17	134791.39

```

sales.hist(bins=30, figsize=(20,20))

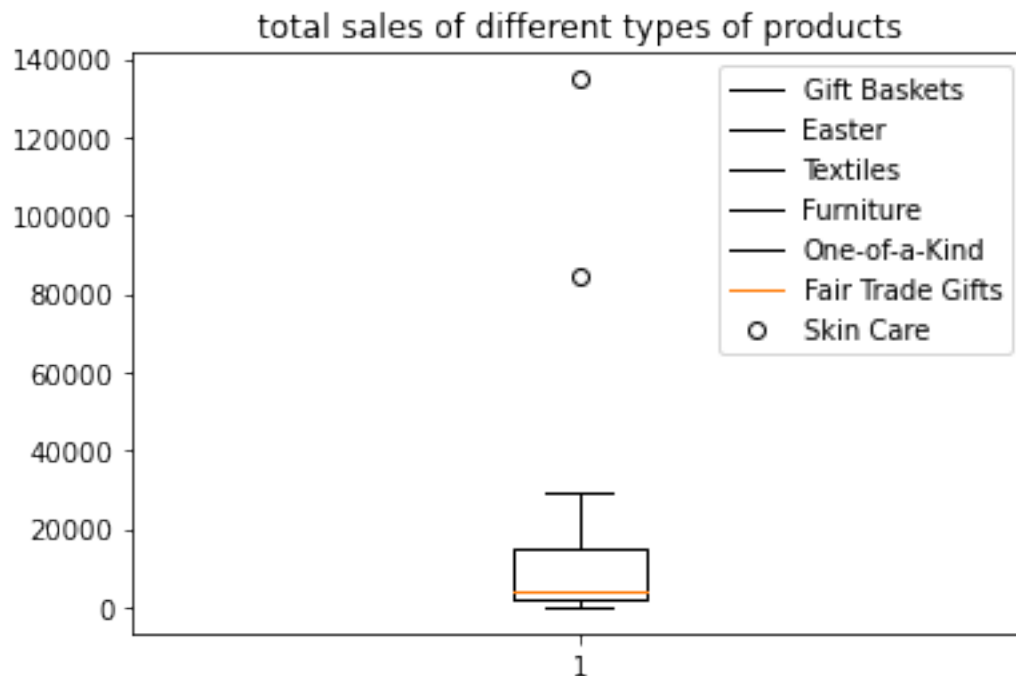
array([[<AxesSubplot:title={'center':'Net Quantity'}>,
        <AxesSubplot:title={'center':'Gross Sales'}>],
       [<AxesSubplot:title={'center':'Discounts'}>,
        <AxesSubplot:title={'center':'Returns'}>],
       [<AxesSubplot:title={'center':'Total Net Sales'}>,
        <AxesSubplot:>]],
       dtype=object)

```



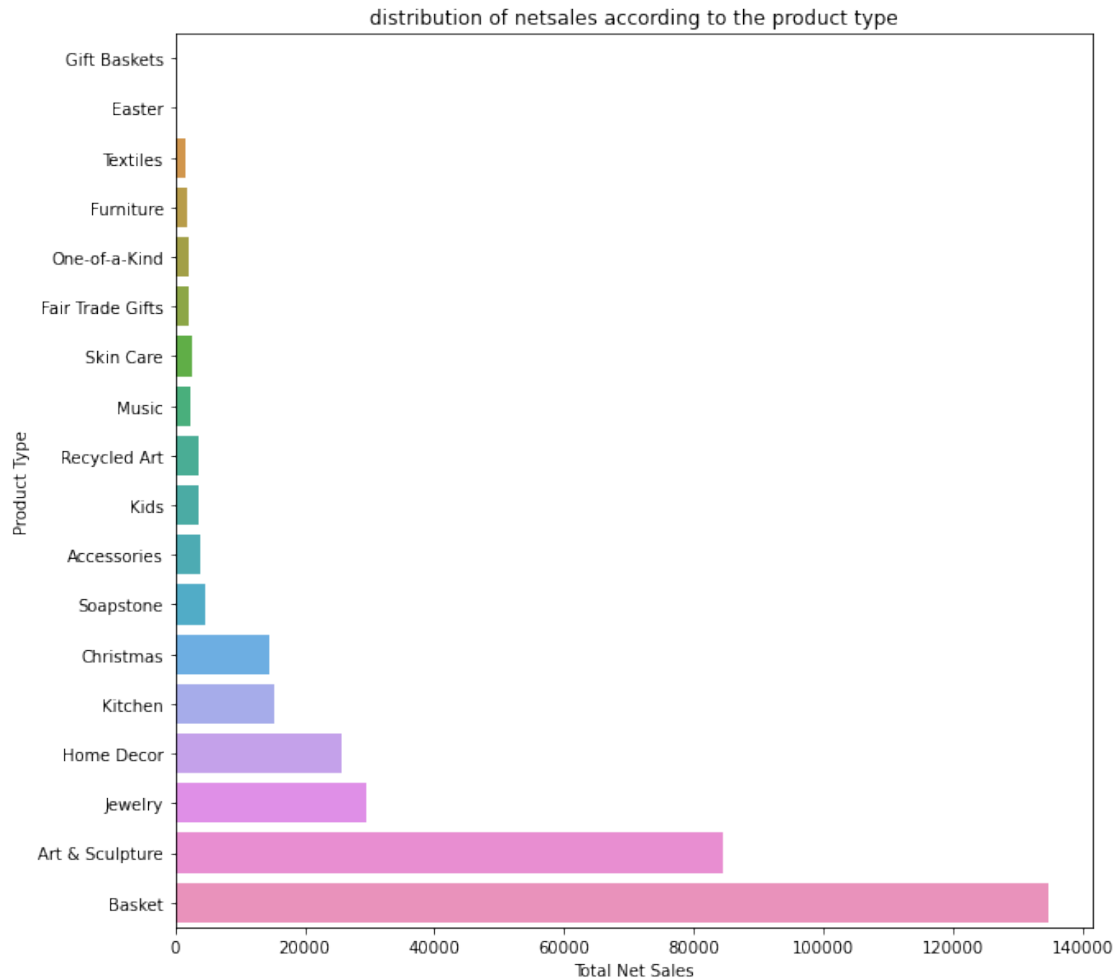
```
y=sales['Total Net Sales']
x=sales['Product Type']
plt.boxplot(y)
plt.legend(x)
plt.title("total sales of different types of products")
```

Text(0.5, 1.0, 'total sales of different types of products')



```
x=sales['Total Net Sales']
y=sales['Product Type']
sns.barplot(x=x,y=y)
plt.rcParams["figure.figsize"]=(10,10)
plt.title("distribution of netsales according to the product type")

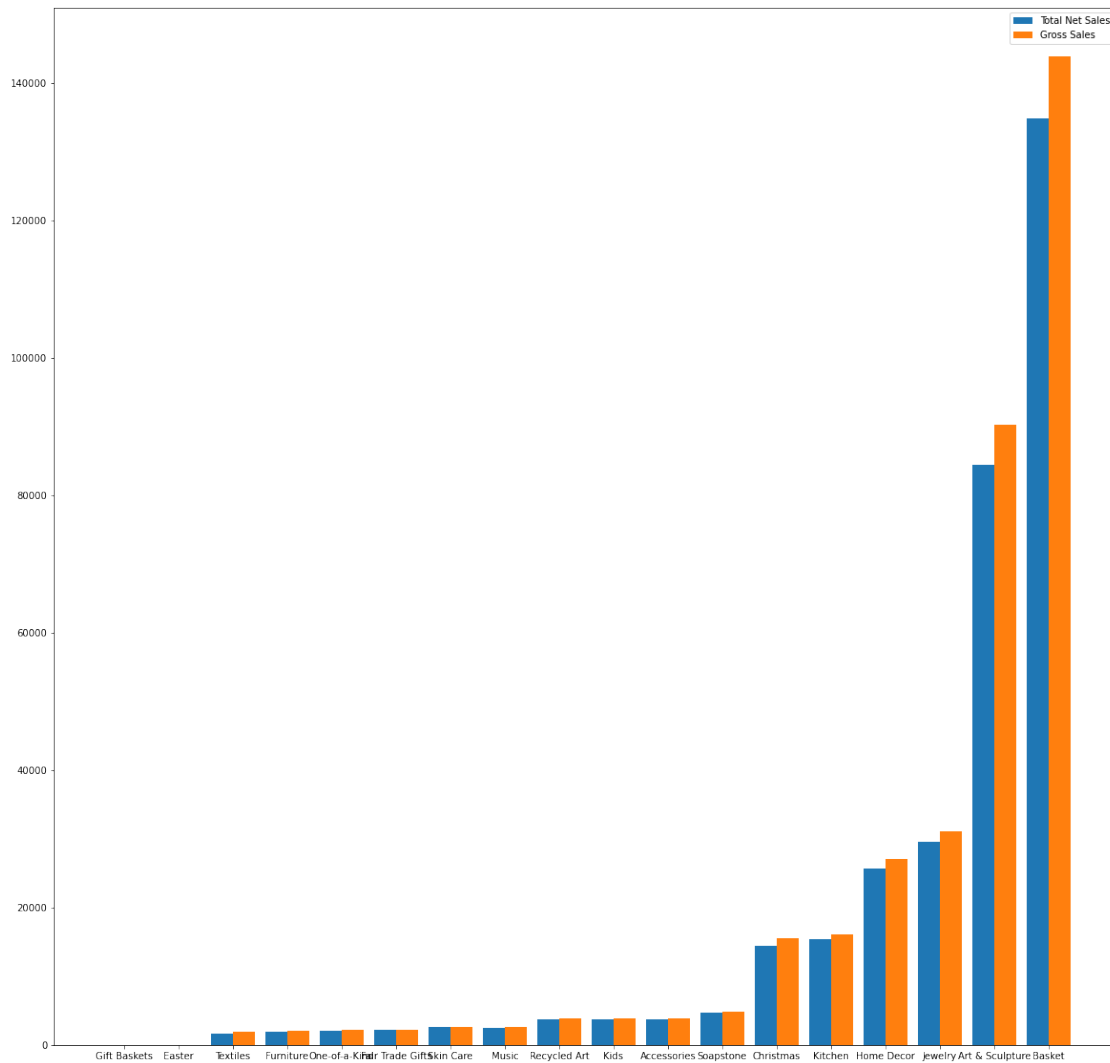
Text(0.5, 1.0, 'distribution of netsales according to the product
type')
```



```

products=sales['Product Type']
x_axis = np.arange(len(products))
x=sales['Total Net Sales']
y=sales['Gross Sales']
plt.bar(x_axis -0.2,x, width=0.4, label = 'Total Net Sales')
plt.bar(x_axis +0.2,y , width=0.4, label = 'Gross Sales')
plt.xticks(x_axis, products)
plt.legend()
plt.show()
plt.rcParams["figure.figsize"]=(20,20)

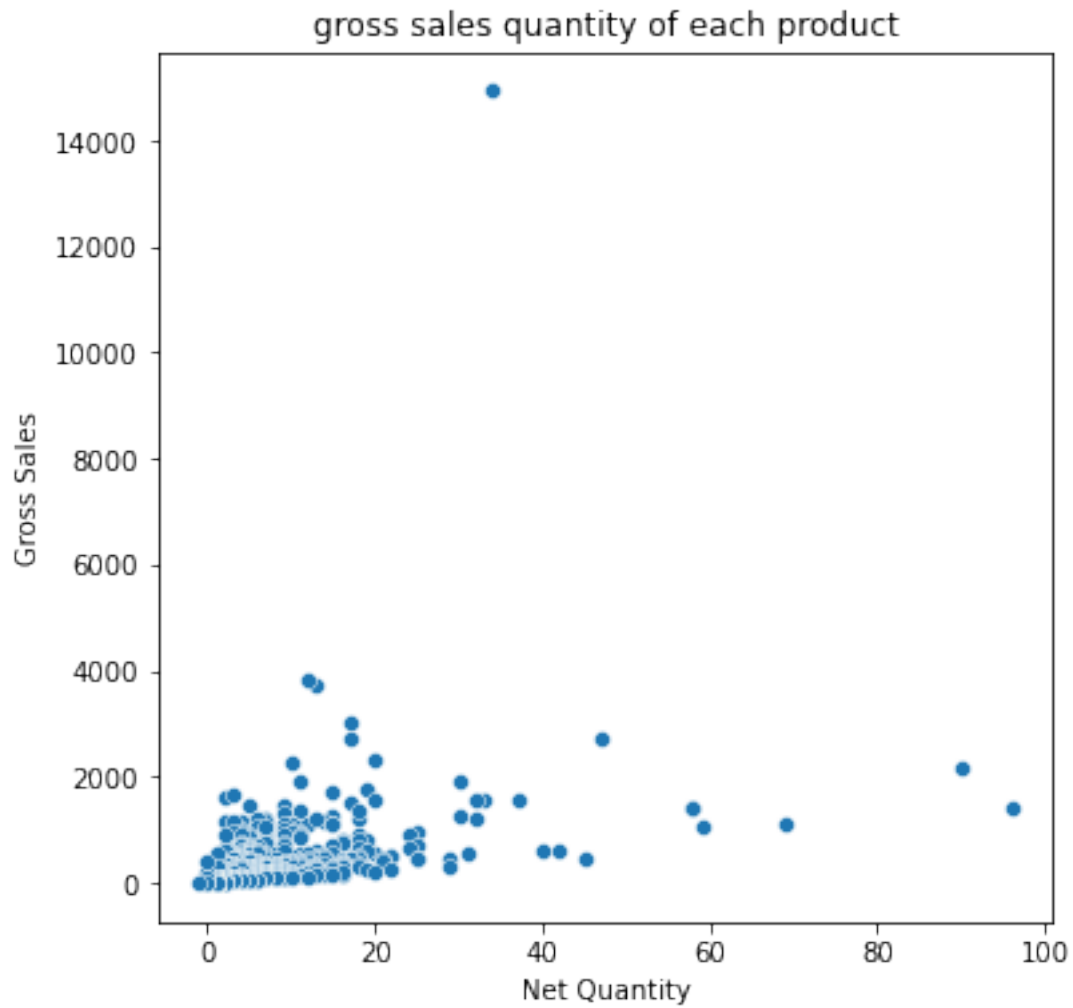
```



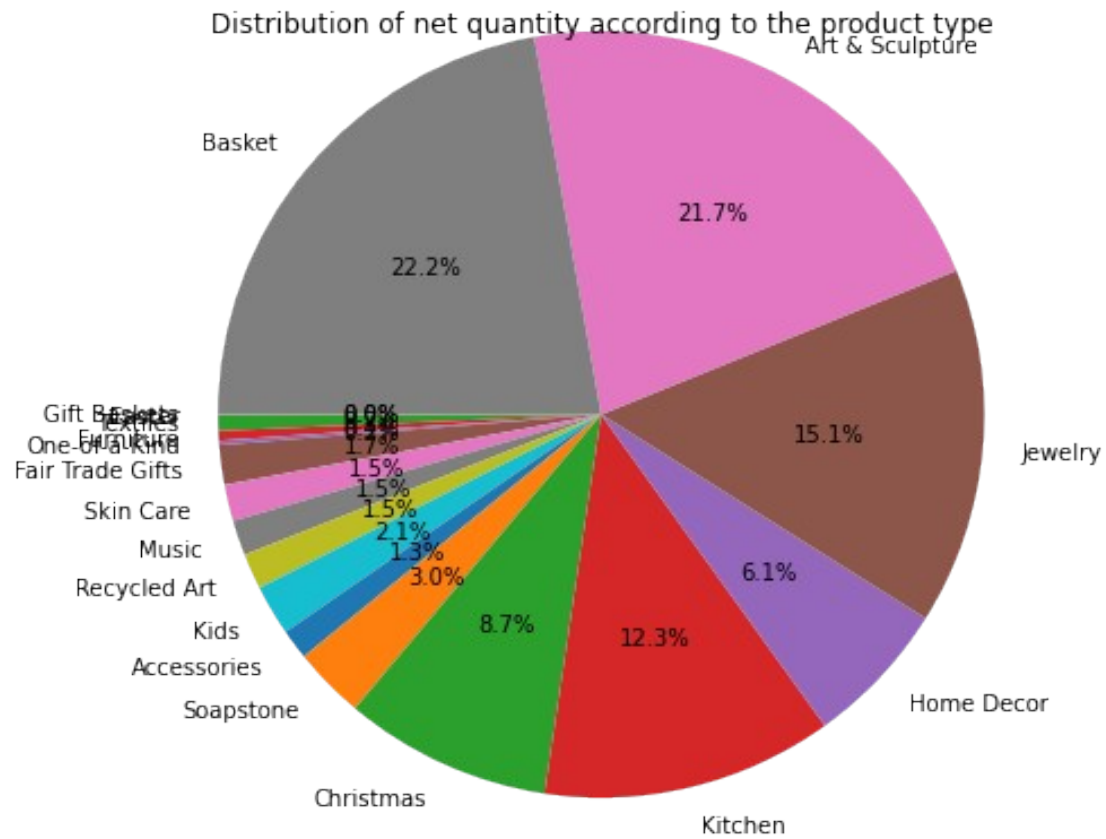
```

x=df['Net Quantity']
y=df['Gross Sales']
sns.scatterplot(x=x,y=y)
plt.title('gross sales quantity of each product')
plt.rcParams["figure.figsize"]=(6,6)

```



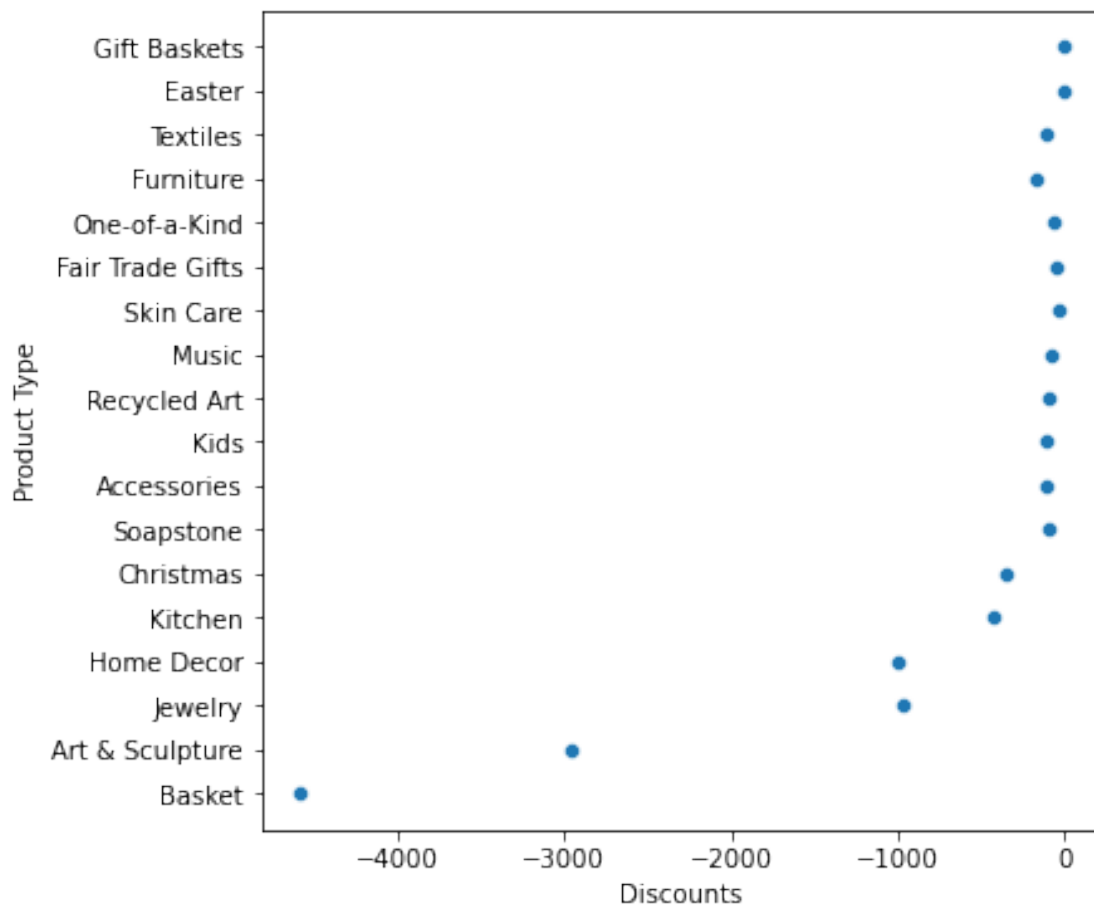
```
y=sales['Product Type']
x=sales['Net Quantity']
plt.pie(x,labels = y,startangle=180,autopct='%.1f%%',radius=1.3)
plt.title("Distribution of net quantity according to the product
type")
plt.show()
```

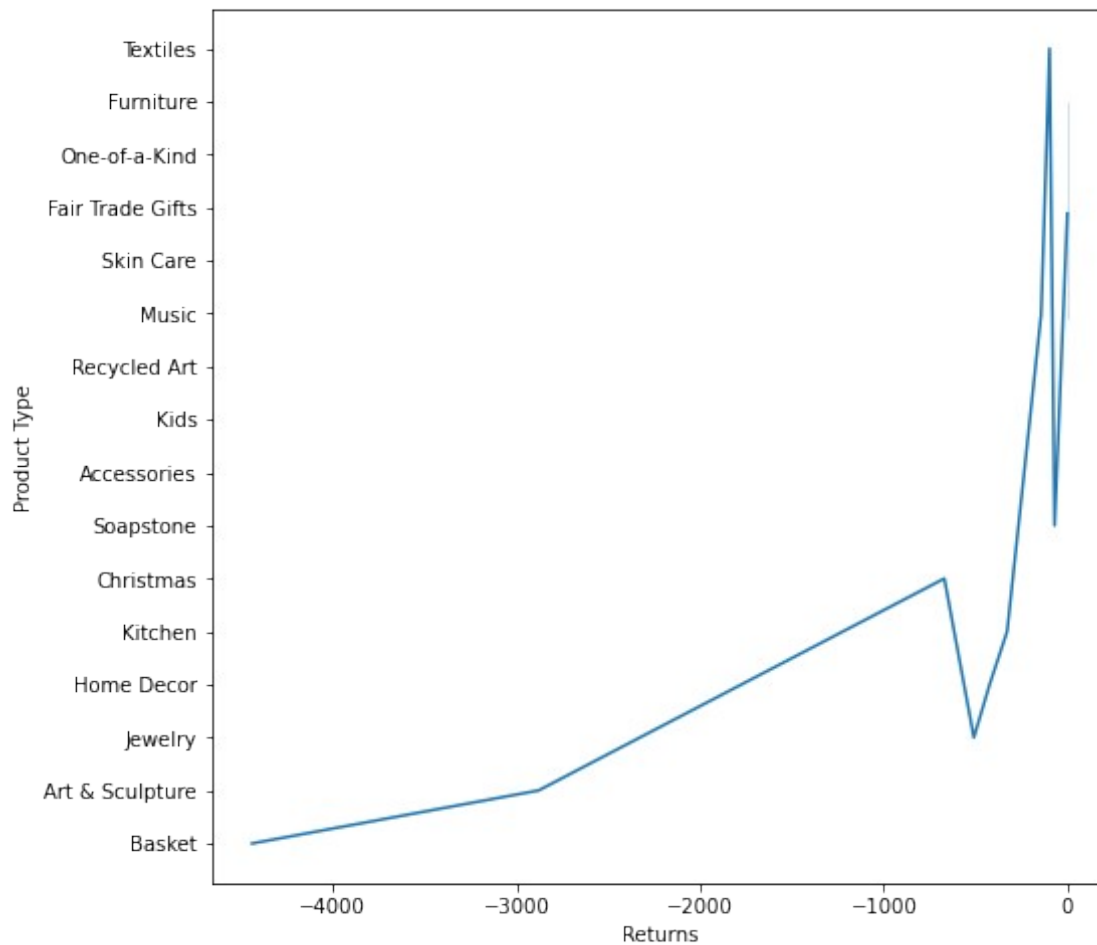
```

y=sales['Product Type']
x=sales['Discounts']
sns.scatterplot(x=x,y=y)
plt.rcParams['figure.figsize'] = [5,5]

```



```
y=sales['Product Type']
x=sales['Returns']
sns.lineplot(x=x,y=y)
plt.rcParams['figure.figsize'] = [10,7]
```



```
table1=pd.pivot_table(df,values=['Net Quantity','Total Net Sales'],index=['Product Type'],aggfunc=np.sum)
table1
```

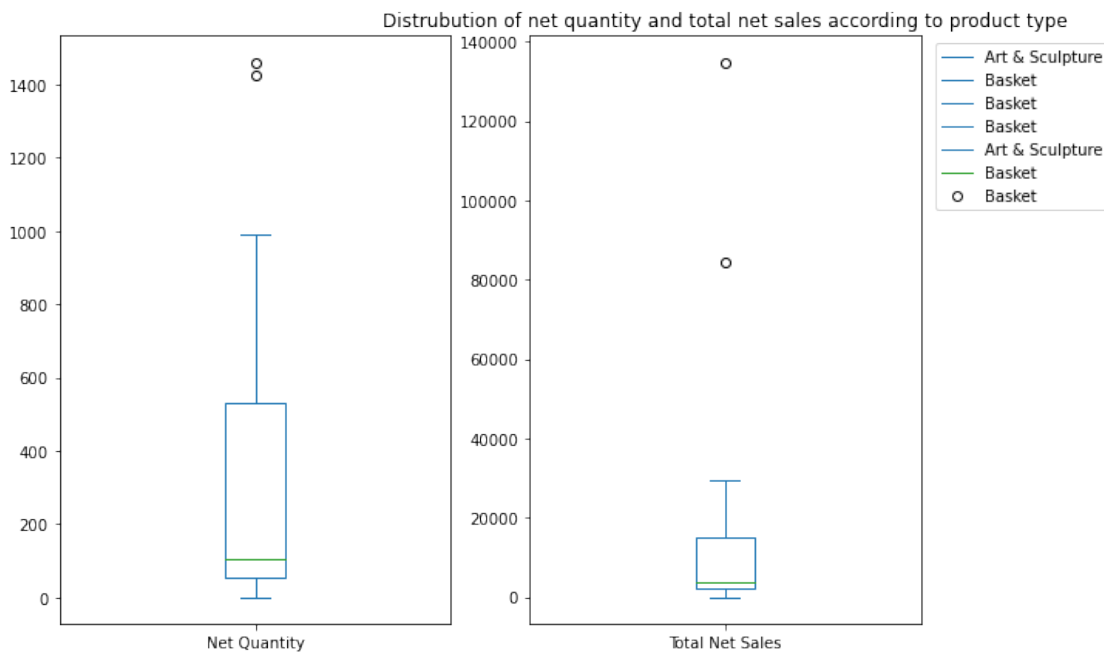
Product Type	Net Quantity	Total Net Sales
Accessories	84	3785.38
Art & Sculpture	1427	84480.85
Basket	1461	134791.39
Christmas	575	14460.81
Easter	1	34.20
Fair Trade Gifts	110	2204.67
Furniture	27	1864.96
Gift Baskets	1	19.50
Home Decor	404	25699.99
Jewelry	991	29572.95
Kids	140	3721.34
Kitchen	809	15336.82
Music	98	2418.90
One-of-a-Kind	12	2108.01
Recycled Art	99	3704.16

Skin Care	101	2571.80
Soapstone	199	4629.09
Textiles	43	1679.10

```
table1.head()
```

	Net Quantity	Total Net Sales
Product Type		
Accessories	84	3785.38
Art & Sculpture	1427	84480.85
Basket	1461	134791.39
Christmas	575	14460.81
Easter	1	34.20

```
x=df['Product Type']
table1.plot(kind="box",subplots=True)
plt.legend(x,bbox_to_anchor=(1.5, 1))
plt.title('Distribution of net quantity and total net sales according
to product type',loc='center')
plt.rcParams['figure.figsize'] = [12,12]
```



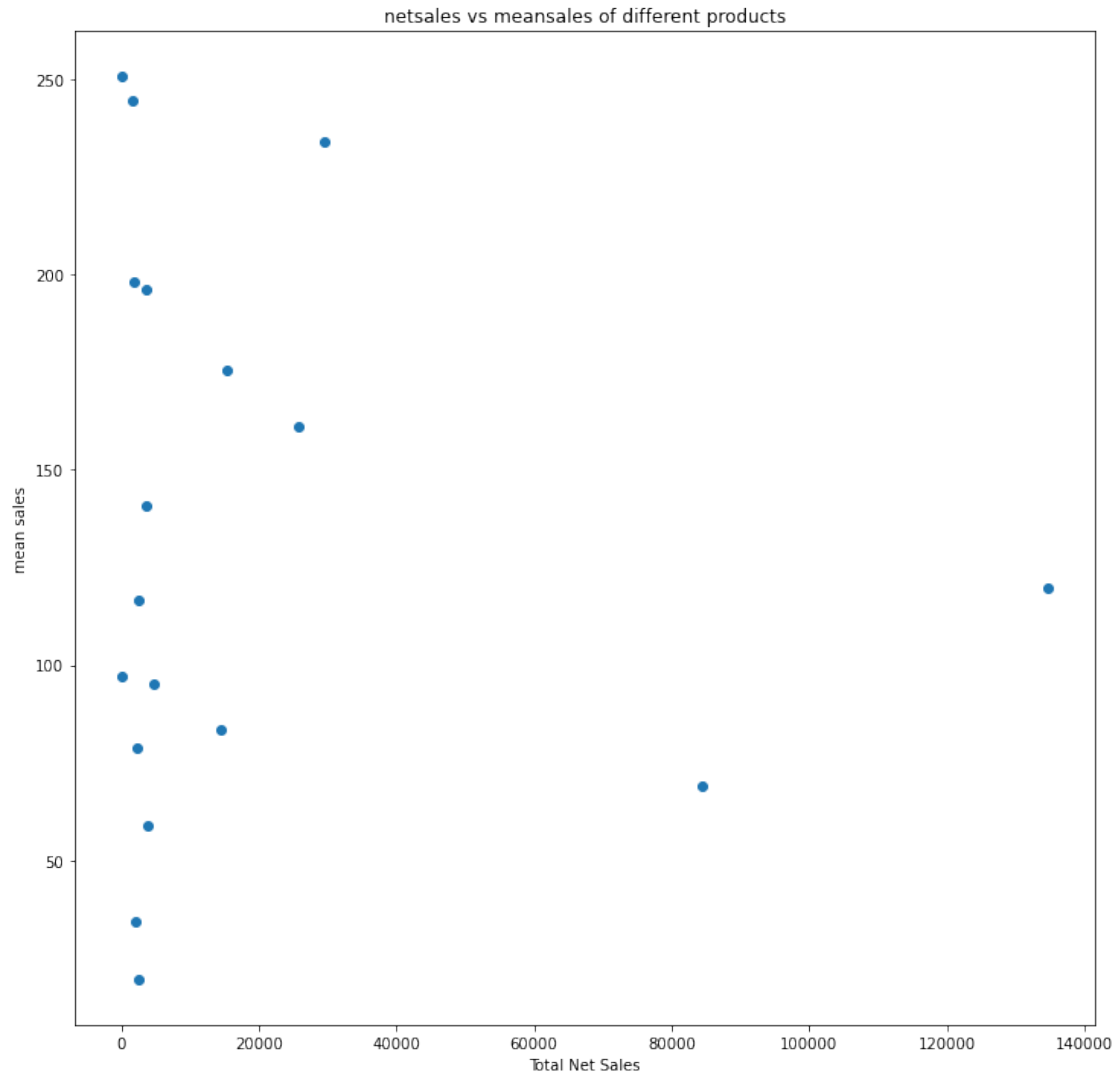
```
z=df.groupby('Product Type')['Total Net Sales'].mean()
z
```

Product Type	
Accessories	97.061026
Art & Sculpture	250.685015
Basket	244.630472
Christmas	198.093288
Easter	34.200000
Fair Trade Gifts	78.738214

Furniture	116.560000
Gift Baskets	19.500000
Home Decor	196.183130
Jewelry	140.823571
Kids	59.068889
Kitchen	95.259752
Music	83.410345
One-of-a-Kind	175.667500
Recycled Art	161.050435
Skin Care	233.800000
Soapstone	69.090896
Textiles	119.935714

Name: Total Net Sales, dtype: float64

```
z=df.groupby('Product Type')['Total Net Sales'].mean()
z
y=sales['Total Net Sales']
len(y)
k=sales['Product Type']
plt.scatter(x=y,y=z)
plt.rcParams['figure.figsize'] = [8,8]
plt.xlabel('Total Net Sales')
plt.ylabel('mean sales')
plt.title('netsales vs meansales of different products')
Text(0.5, 1.0, 'netsales vs meansales of different products')
```



```
table2=pd.pivot_table(df,values=['Gross Sales','Total Net Sales'],index=['Product Type'],aggfunc=np.sum)
table2
```

Product Type	Gross Sales	Total Net Sales
Accessories	3892.40	3785.38
Art & Sculpture	90316.60	84480.85
Basket	143815.50	134791.39
Christmas	15476.00	14460.81
Easter	38.00	34.20
Fair Trade Gifts	2258.00	2204.67
Furniture	2034.00	1864.96
Gift Baskets	19.50	19.50
Home Decor	27114.55	25699.99
Jewelry	31048.00	29572.95
Kids	3838.00	3721.34
Kitchen	16096.00	15336.82

Music	2643.50	2418.90
One-of-a-Kind	2180.00	2108.01
Recycled Art	3792.80	3704.16
Skin Care	2609.50	2571.80
Soapstone	4795.50	4629.09
Textiles	1889.00	1679.10

```
table3=pd.pivot_table(df,values=['Discounts'],index=['Product
Type'],aggfunc=np.sum)
table3
```

Product Type	Discounts
Accessories	-107.02
Art & Sculpture	-2955.82
Basket	-4584.42
Christmas	-345.19
Easter	-3.80
Fair Trade Gifts	-53.33
Furniture	-169.04
Gift Baskets	0.00
Home Decor	-991.21
Jewelry	-965.85
Kids	-116.66
Kitchen	-431.11
Music	-82.19
One-of-a-Kind	-71.99
Recycled Art	-88.64
Skin Care	-37.70
Soapstone	-96.91
Textiles	-112.90