

Plotting Bar Chart, Pie Chart & Pareto Diagram

```
In [1]:  ▶ import matplotlib.pyplot as plt
```

```
In [2]:  ▶ import seaborn as sns
```

German Car Shop Dataset

```
In [3]:  ▶ data={'category':['Audi','BMW','Mercedes'],'frequency':[124,98,113]}
```

```
In [4]:  ▶ data
```

```
Out[4]: {'category': ['Audi', 'BMW', 'Mercedes'], 'frequency': [124, 98, 113]}
```

```
In [5]:  ▶ import pandas as pd
```

```
In [6]:  ▶ df=pd.DataFrame(data)
```

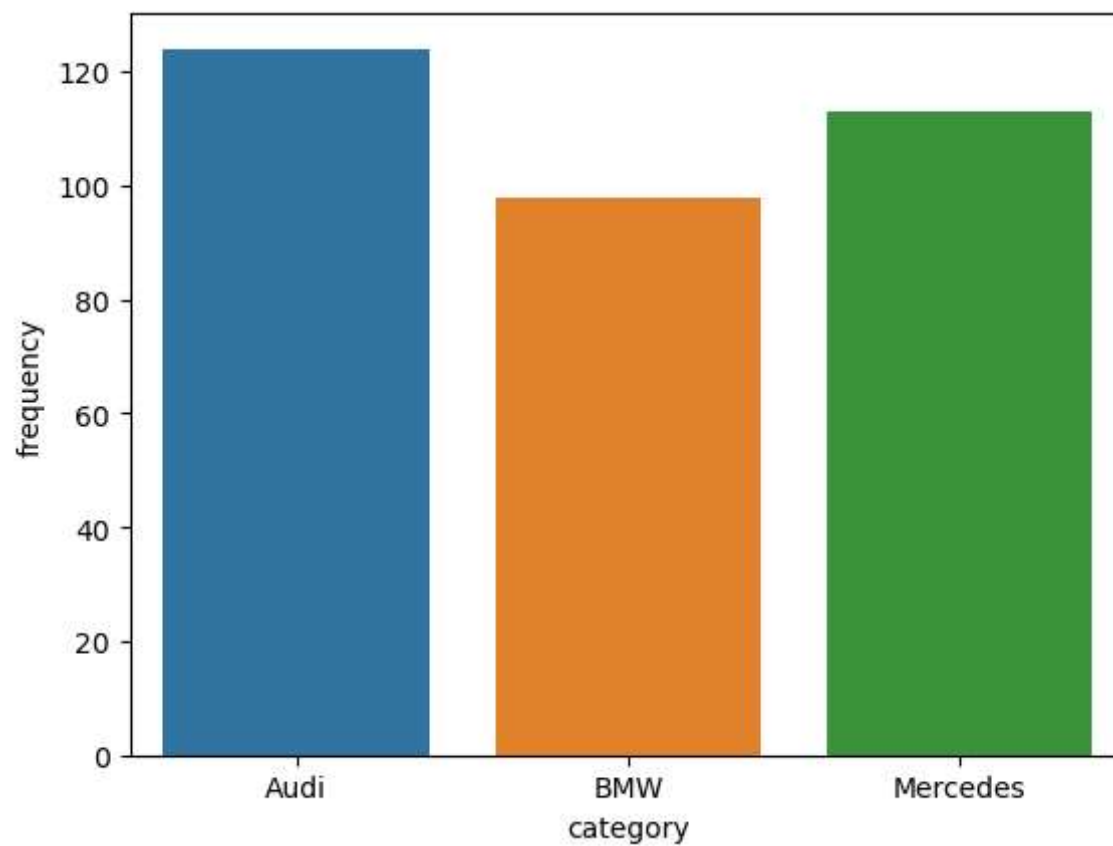
```
In [7]:  ▶ df
```

```
Out[7]:
```

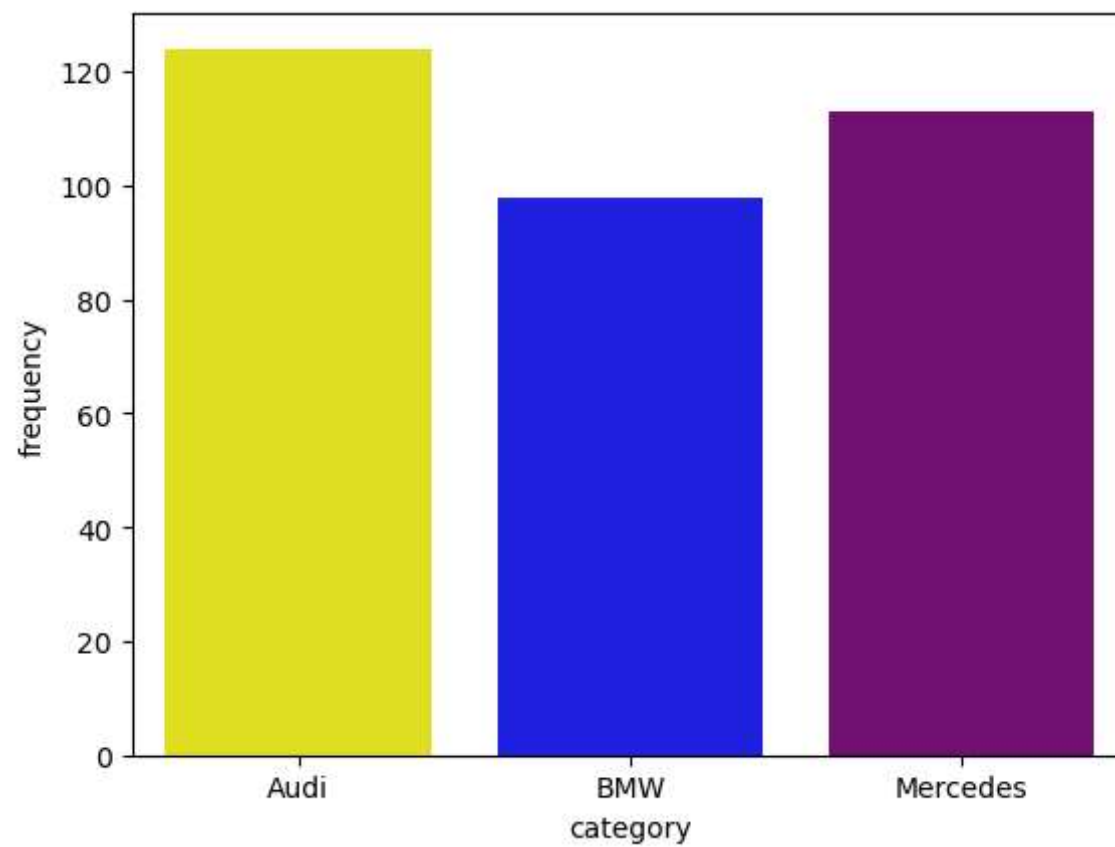
	category	frequency
0	Audi	124
1	BMW	98
2	Mercedes	113

Bar Chart

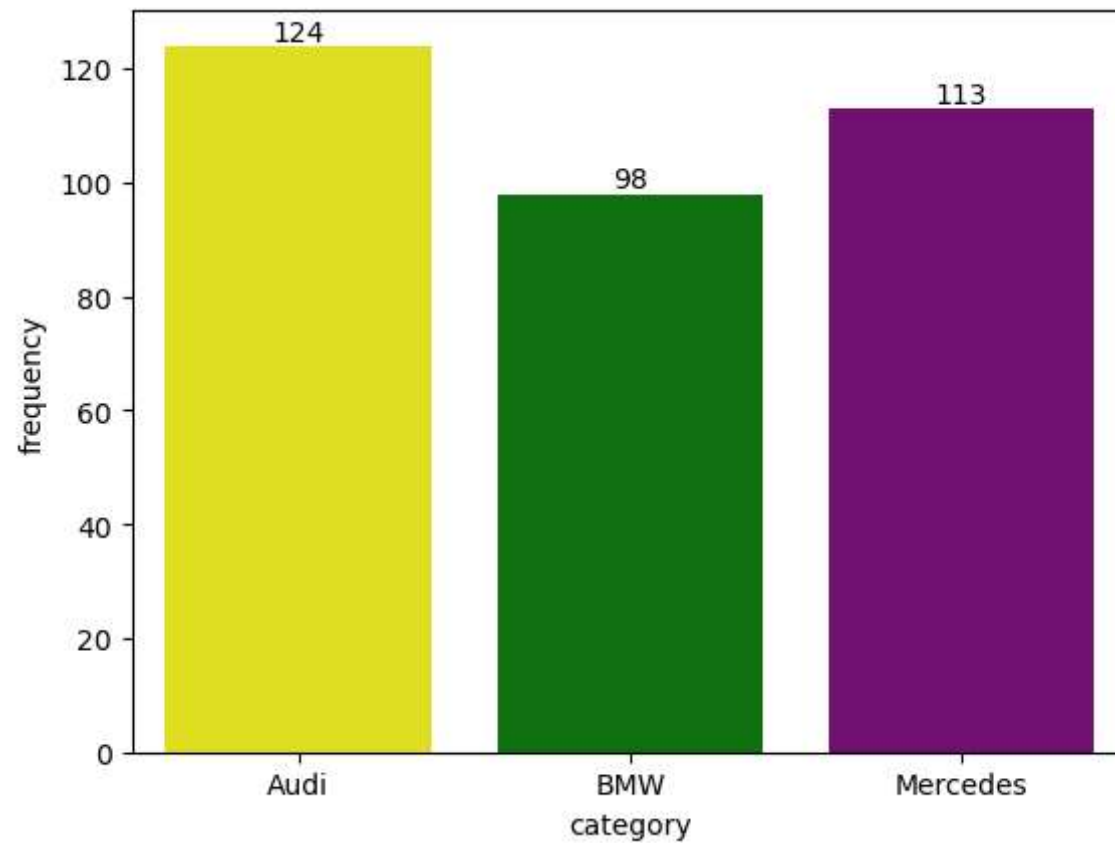
```
In [8]: bar=sns.barplot(data=df,x='category',y='frequency')
```



```
In [9]: bar=sns.barplot(data=df,x='category',y='frequency',palette=('yellow','blue','purple'))
```

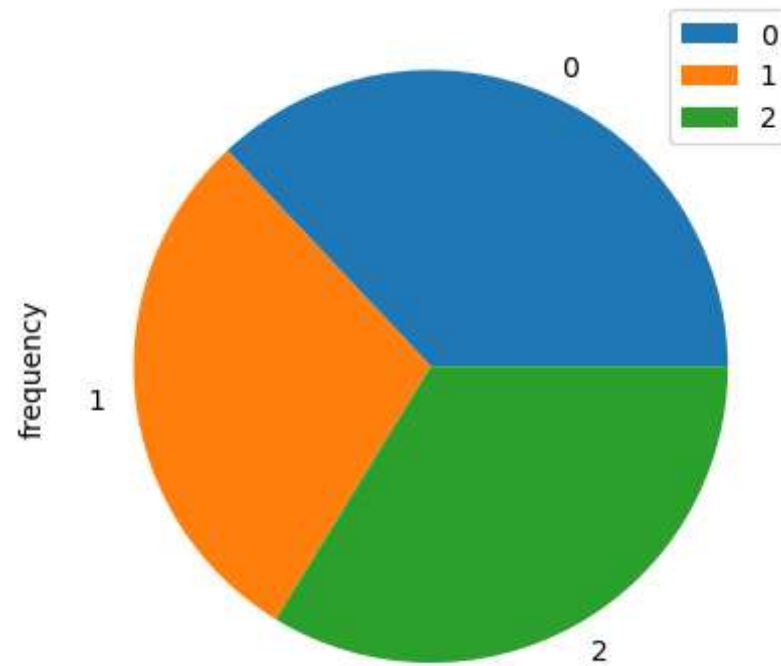


```
In [10]: ▶ bar=sns.barplot(data=df,x='category',y='frequency',palette=('yellow','green','purple'))  
for i in bar.containers:  
    bar.bar_label(i)
```

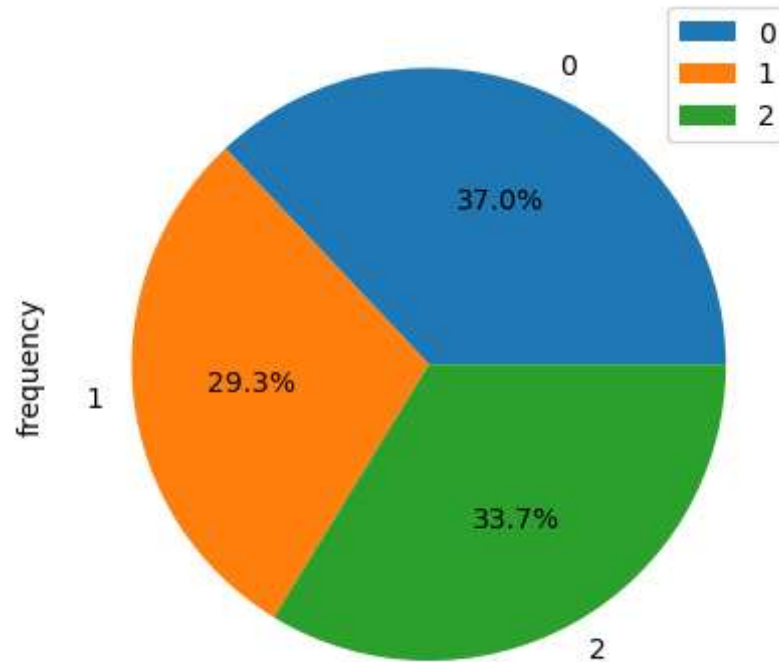


Pie Chart

```
In [11]: pie=df.plot.pie(y='frequency',legend=True)
```



```
In [12]: ▶ pie=df.plot.pie(x='category',y='frequency',autopct='%1.1f%%')
```



Pareto Diagram

```
In [13]: ▶ #we required cumulative frequency to plot pareto diagram
```

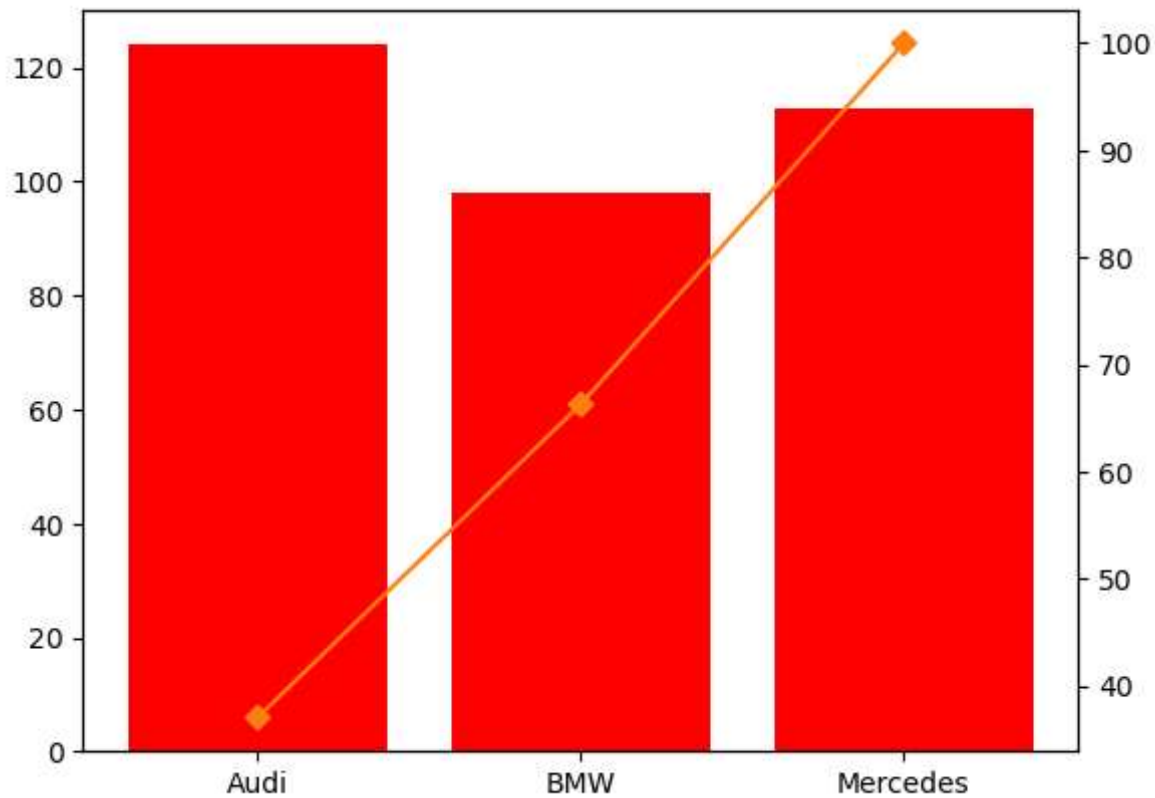
```
In [14]: ▶ df['cum_freq']=df['frequency'].cumsum()/df['frequency'].sum()*100
```

```
In [15]: df
```

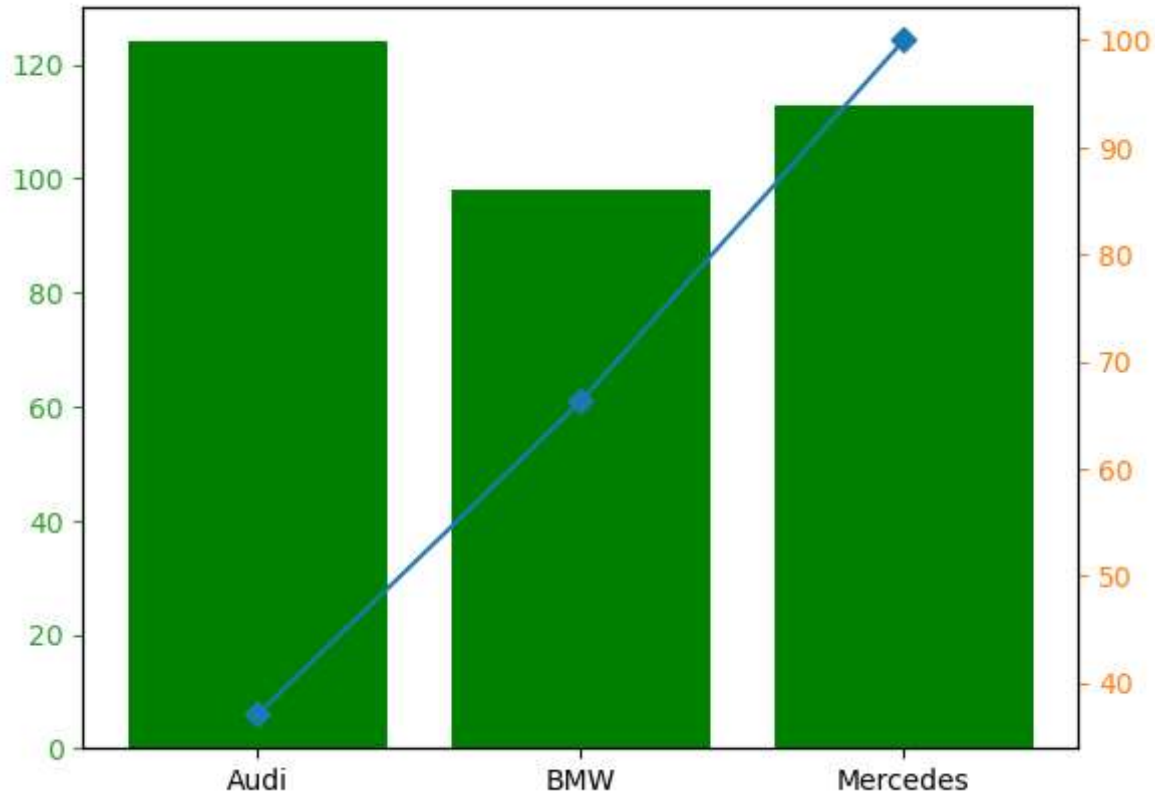
```
Out[15]:
```

	category	frequency	cum_freq
0	Audi	124	37.014925
1	BMW	98	66.268657
2	Mercedes	113	100.000000

```
In [25]: fig,ax=plt.subplots()  
ax.bar(df.category,df['frequency'],color='red')  
ax1=ax.twinx()  
ax1.plot(df.category,df['cum_freq'],marker='D',color='C1')  
plt.show()
```



```
In [27]: fig,ax=plt.subplots()
ax.bar(df.category,df['frequency'],color='green')
ax1=ax.twinx()
ax1.plot(df.category,df['cum_freq'],marker='D')
ax.tick_params(axis='y',colors='C2')
ax1.tick_params(axis='y',colors='C1')
plt.show()
```



Ice Cream Shop Dataset


```
In [2]: ▶ import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [3]: ▶ data={'area':['New York','LA','San Francisco'],'frequency':[ 12327,17129,19923]}
```

```
In [4]: ▶ data
```

```
Out[4]: {'area': ['New York', 'LA', 'San Francisco'],
'frequency': [12327, 17129, 19923]}
```

```
In [5]: ▶ df=pd.DataFrame(data)
```

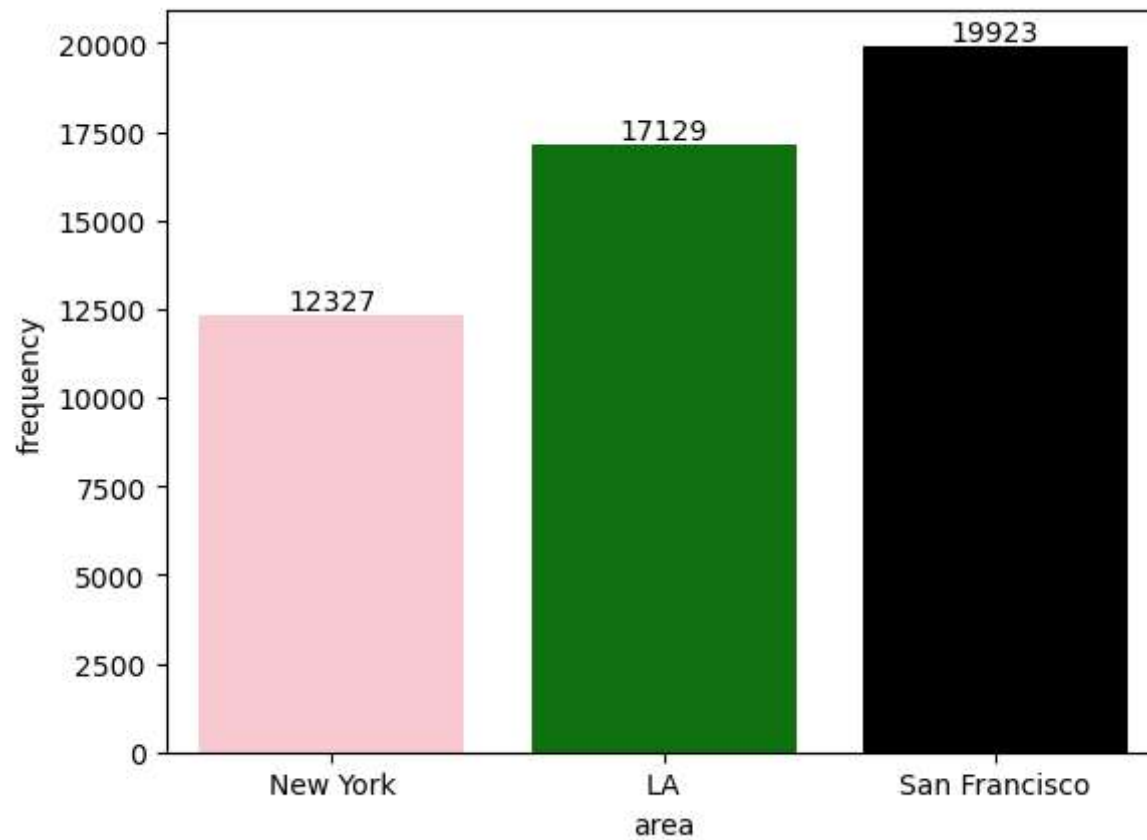
```
In [6]: ▶ df
```

```
Out[6]:
```

	area	frequency
0	New York	12327
1	LA	17129
2	San Francisco	19923

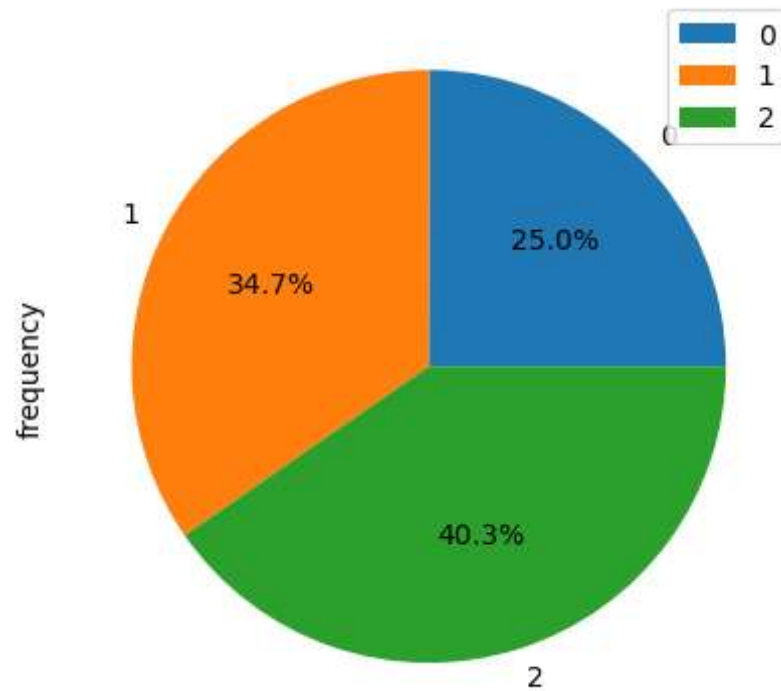
Barchart

```
In [7]: ▶ bar=sns.barplot(data=df,x='area',y='frequency',palette=('pink','green','black'))  
        for i in bar.containers:  
            bar.bar_label(i)
```



Pie Chart

```
In [8]: ▶ pie=df.plot.pie(x='area',y='frequency',autopct='%1.1f%%')
```



Pareto Diagram

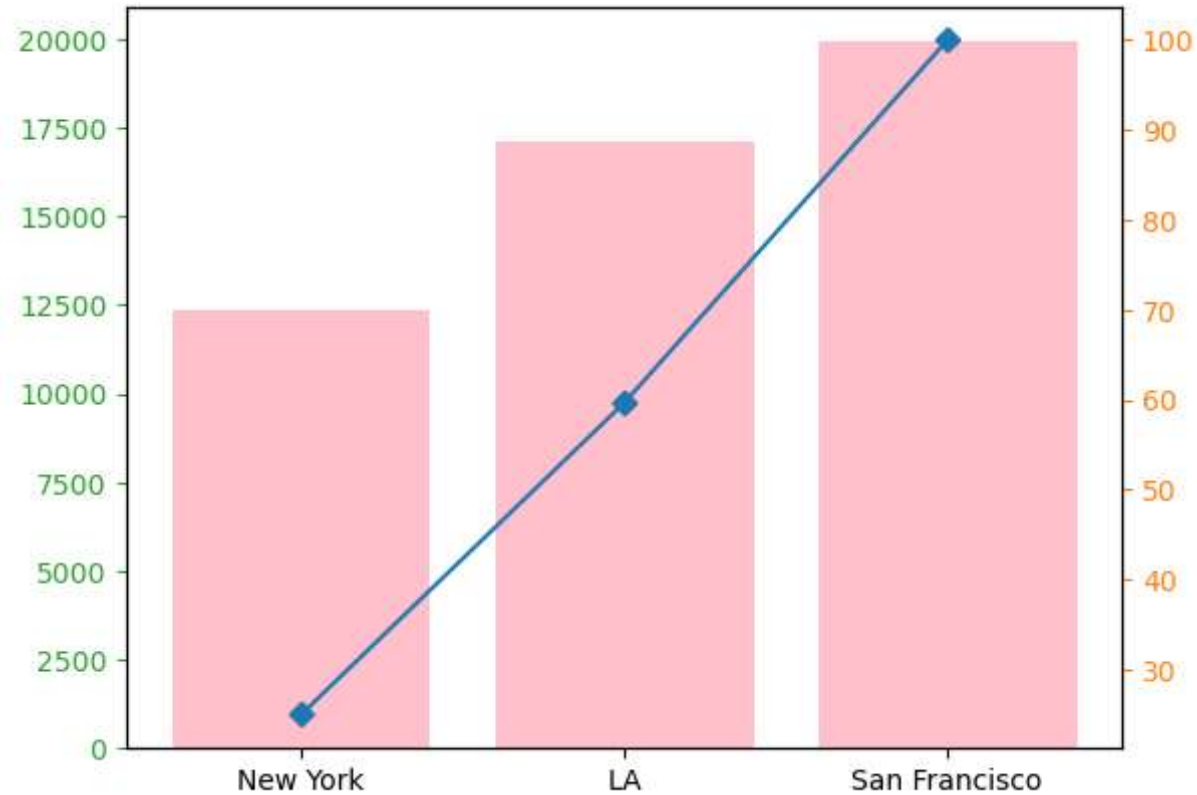
```
In [9]: ▶ df['cum_freq']=df['frequency'].cumsum()/df['frequency'].sum()*100
```

In [10]: ▶ df

Out[10]:

	area	frequency	cum_freq
0	New York	12327	24.964054
1	LA	17129	59.652889
2	San Francisco	19923	100.000000

```
In [11]: fig,ax=plt.subplots()
ax.bar(df.area,df['frequency'],color='pink')
ax1=ax.twinx()
ax1.plot(df.area,df['cum_freq'],marker='D')
ax.tick_params(axis='y',colors='C2')
ax1.tick_params(axis='y',colors='C1')
plt.show()
```



```
In [ ]: 
```

