

```
[1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

```
[2]: df=pd.read_csv('D:\\jupyter notebook\\heart.csv')
```

```
[3]: df.head()
```

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	52	1	0	125	212	0	1	168	0	1.0	2	2	3	0
1	53	1	0	140	203	1	0	155	1	3.1	0	0	3	0
2	70	1	0	145	174	0	1	125	1	2.6	0	0	3	0
3	61	1	0	148	203	0	1	161	0	0.0	2	1	3	0
4	62	0	0	138	294	1	1	106	0	1.9	1	3	2	0

```
[4]: df.tail()
```

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
1020	59	1	1	140	221	0	1	164	1	0.0	2	0	2	1
1021	60	1	0	125	258	0	0	141	1	2.8	1	1	3	0
1022	47	1	0	110	275	0	0	118	1	1.0	1	1	2	0
1023	50	0	0	110	254	0	0	159	0	0.0	2	0	2	1
1024	54	1	0	120	188	0	1	113	0	1.4	1	1	3	0

```
[5]: df.columns.values
```

```
[5]: array(['age', 'sex', 'cp', 'trestbps', 'chol', 'fbs', 'restecg',
        'thalach', 'exang', 'oldpeak', 'slope', 'ca', 'thal', 'target'],
        dtype=object)
```

```
[6]: df.isna().sum()
```

```
[6]: age      0
sex      0
cp       0
trestbps  0
chol     0
```

1023	50	0	0	110	254	0	0	159	0	0.0	2	0	2	1
1024	54	1	0	120	188	0	1	113	0	1.4	1	1	3	0

[5]: `df.columns.values`

[5]: array(['age', 'sex', 'cp', 'trestbps', 'chol', 'fbs', 'restecg', 'thalach', 'exang', 'oldpeak', 'slope', 'ca', 'thal', 'target'], dtype=object)

[6]: `df.isna().sum()`

[6]:

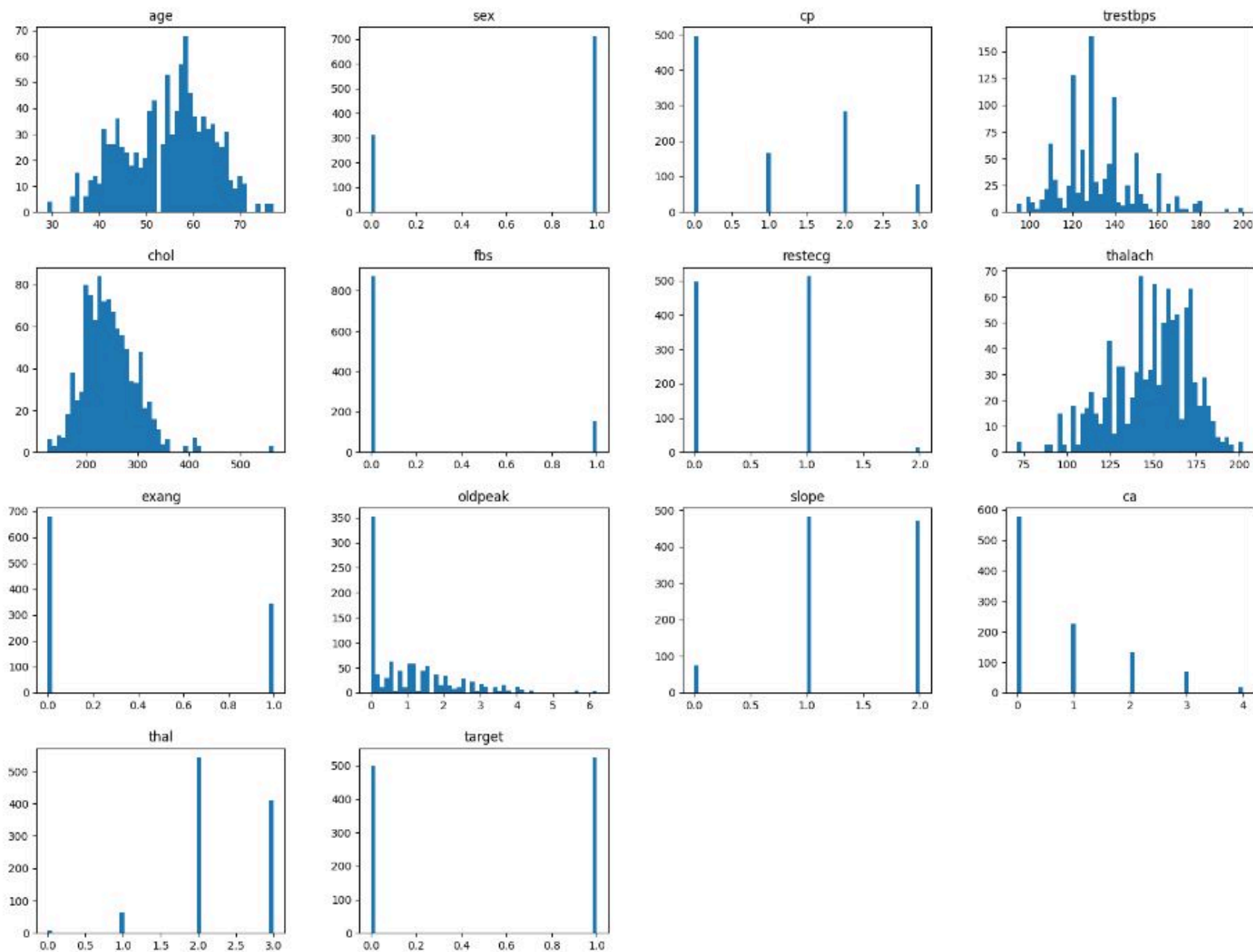
age	0
sex	0
cp	0
trestbps	0
chol	0
fbs	0
restecg	0
thalach	0
exang	0
oldpeak	0
slope	0
ca	0
thal	0
target	0

dtype: int64

[7]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1025 entries, 0 to 1024
Data columns (total 14 columns):
#   Column      Non-Null Count  Dtype
---  -
0   age         1025 non-null   int64
1   sex         1025 non-null   int64
2   cp          1025 non-null   int64
3   trestbps    1025 non-null   int64
4   chol        1025 non-null   int64
5   fbs         1025 non-null   int64
6   restecg     1025 non-null   int64
7   thalach     1025 non-null   int64
8   exang       1025 non-null   int64
9   oldpeak     1025 non-null   float64
10  slope       1025 non-null   int64
11  ca          1025 non-null   int64
12  thal        1025 non-null   int64
13  target      1025 non-null   int64
dtypes: float64(1), int64(13)
memory usage: 112.2 KB
```

```
[8]: #histogram of all numeric values  
df.hist(bins=50,grid=False,figsize=(20,15));
```



```
[10]: df.describe()
```

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca
count	1025.000000	1025.000000	1025.000000	1025.000000	1025.000000	1025.000000	1025.000000	1025.000000	1025.000000	1025.000000	1025.000000	1025.000000
mean	54.434146	0.695610	0.942439	131.611707	246.000000	0.149268	0.529756	149.114146	0.336585	1.071512	1.385366	0.754146
std	9.072290	0.460373	1.029641	17.516718	51.59251	0.356527	0.527878	23.005724	0.472772	1.175053	0.617755	1.030798
min	29.000000	0.000000	0.000000	94.000000	126.000000	0.000000	0.000000	71.000000	0.000000	0.000000	0.000000	0.000000
25%	48.000000	0.000000	0.000000	120.000000	211.000000	0.000000	0.000000	132.000000	0.000000	0.000000	1.000000	0.000000
50%	56.000000	1.000000	1.000000	130.000000	240.000000	0.000000	1.000000	152.000000	0.000000	0.800000	1.000000	0.000000
75%	61.000000	1.000000	2.000000	140.000000	275.000000	0.000000	1.000000	166.000000	1.000000	1.800000	2.000000	1.000000
max	77.000000	1.000000	3.000000	200.000000	564.000000	1.000000	2.000000	202.000000	1.000000	6.200000	2.000000	4.000000

```
[11]: ques=["1. How many people have heart disease and how many people doesn't have heart disease?",
          "2. People of which sex has more heart disease?",
          "3. People of which sex have which type of chest pain?",
          "4. People of which chest pain are mmmost pron to have heart disease?"]
ques
```

```
[11]: ["1. How many people have heart disease and how many people doesn't have heart disease?",
       '2. People of which sex has more heart disease?',
       '3. People of which sex have which type of chest pain?',
       '4. People of which chest pain are mmmost pron to have heart disease?']
```

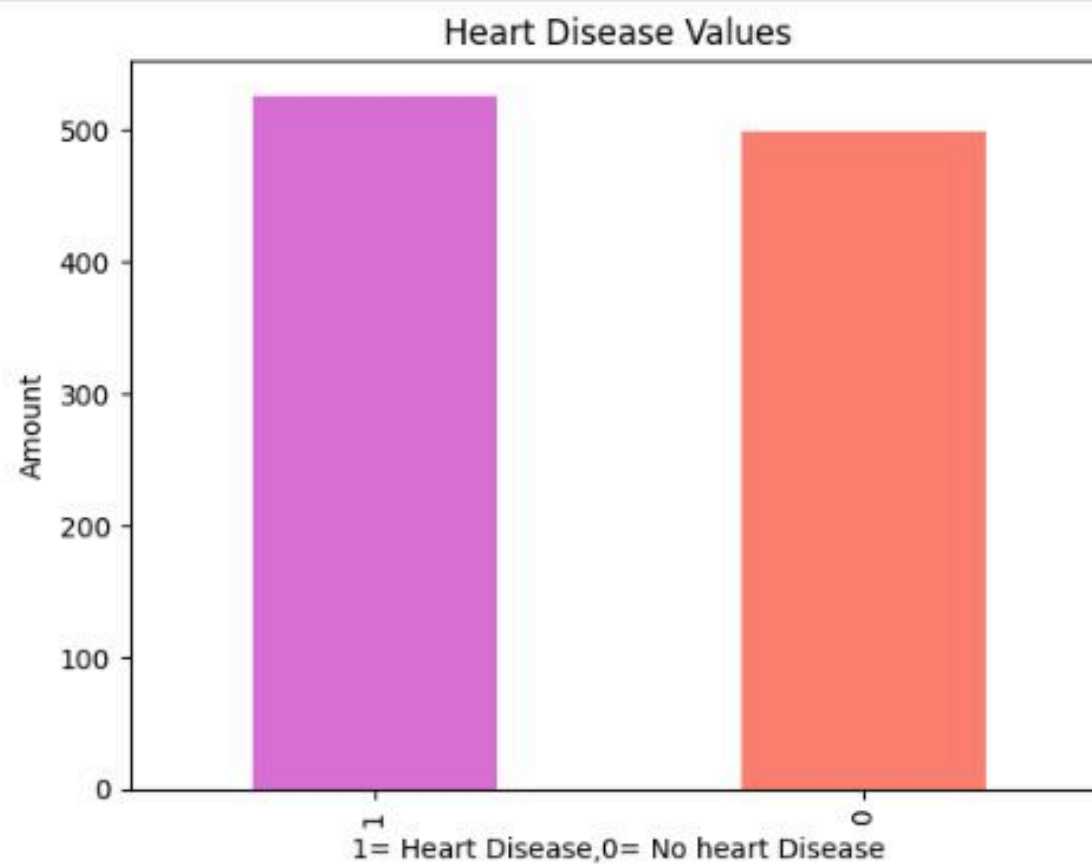
```
[12]: #to find the ans of first ques
      #getting the values
      df.target.value_counts()
```

```
[12]: target
1    526
0    499
Name: count, dtype: int64
```

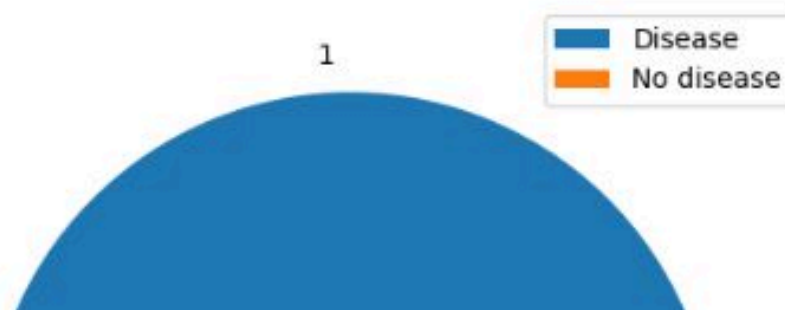
```
[13]: #plotting bar chart
      df.target.value_counts().plot(kind='bar',color=["orchid","salmon"])
      plt.title("Heart Disease Values")
      plt.xlabel("1= Heart Disease,0= No heart Disease")
      plt.ylabel("Amount");
```

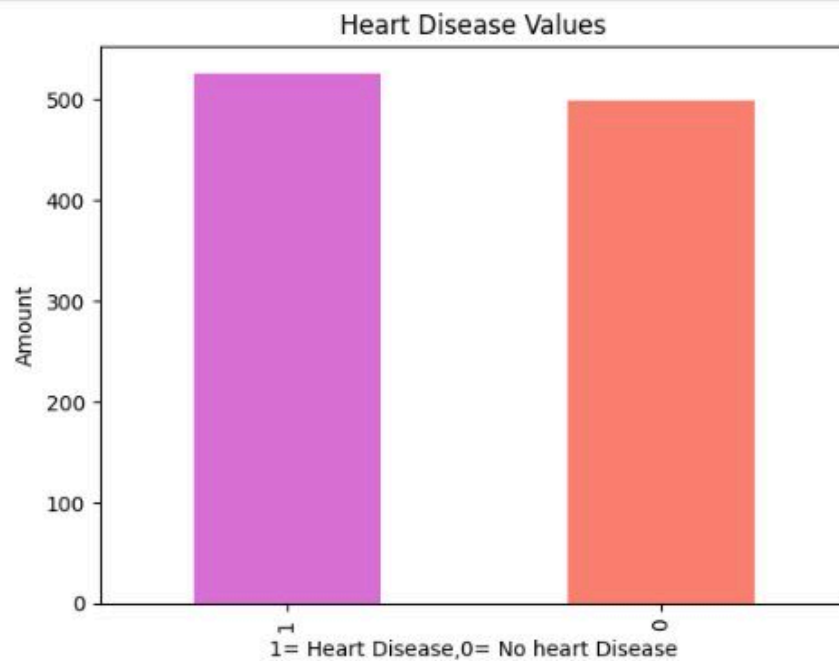
Heart Disease Values



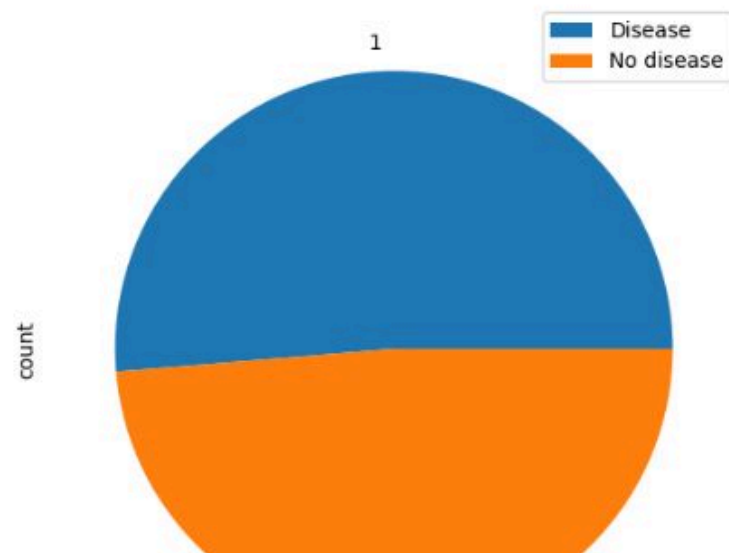


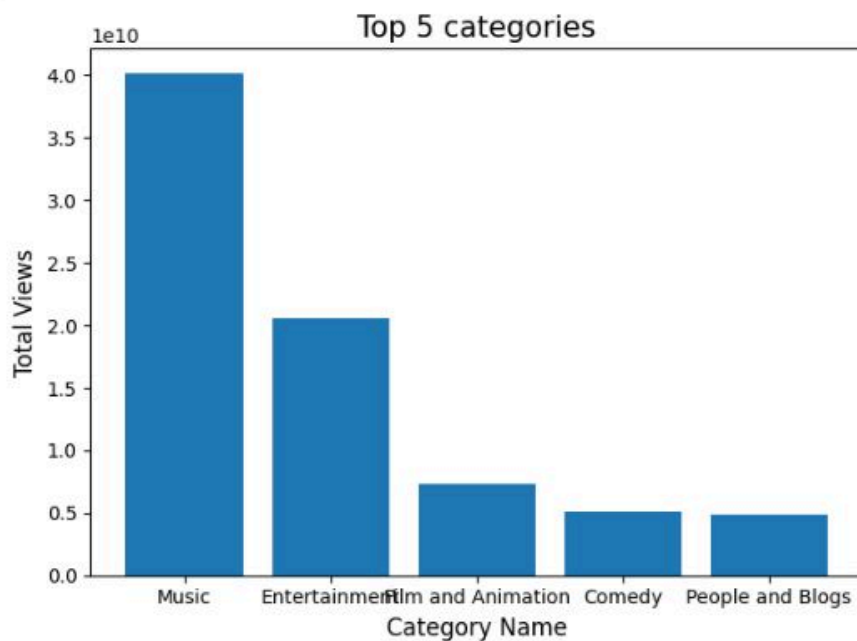
```
[14]: #plotting pie chart
df.target.value_counts().plot(kind='pie',figsize=(8,6))
plt.legend(["Disease","No disease"]);
```



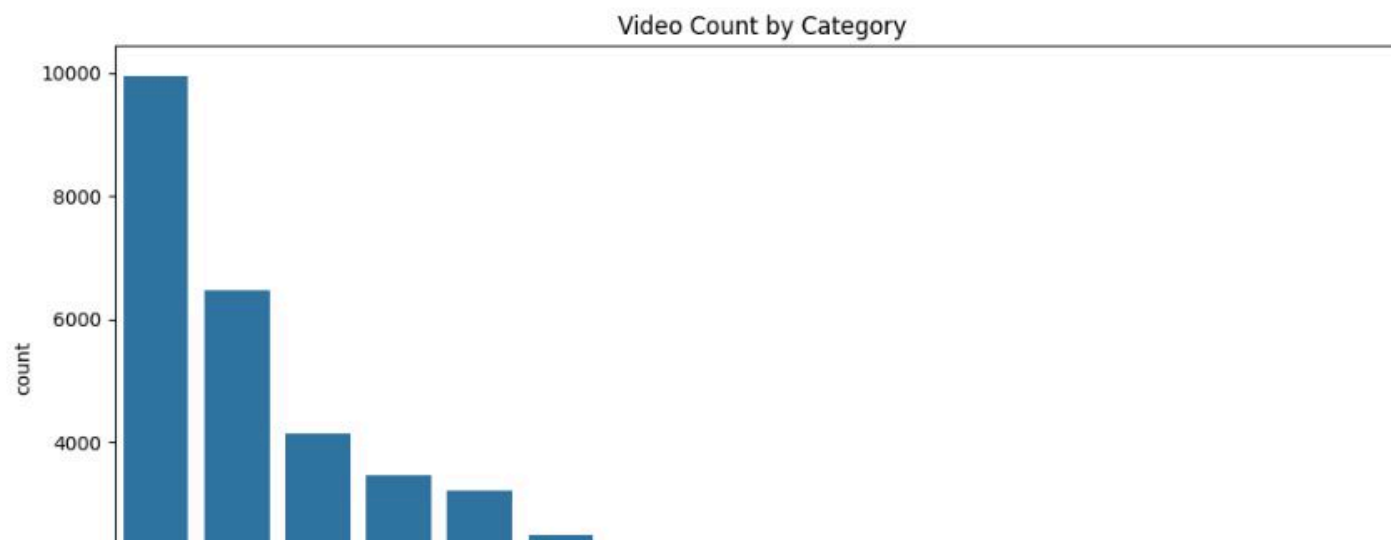


```
[14]: #plotting pie chart
df.target.value_counts().plot(kind='pie',figsize=(8,6))
plt.legend(["Disease","No disease"]);
```





```
[21]: plt.figure(figsize=(12,6))
sns.countplot(x='category_name',data=df,order=df['category_name'].value_counts().index)
plt.xticks(rotation=90)
plt.title('Video Count by Category')
plt.show()
```



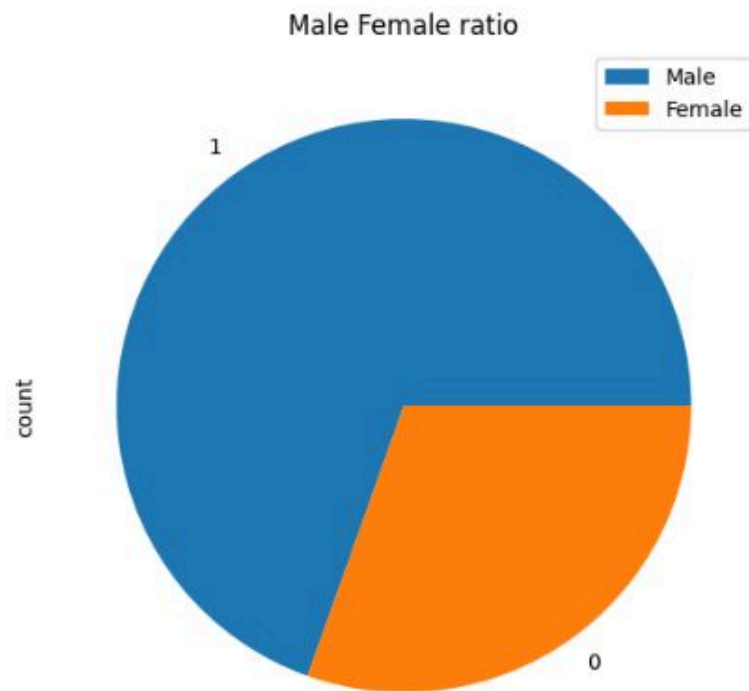
```
[15]: # 0 for Female
# 1 for male
# sex column part
# 0 represent no disease
# 1 for disease
# target column part

df.sex.value_counts()
```

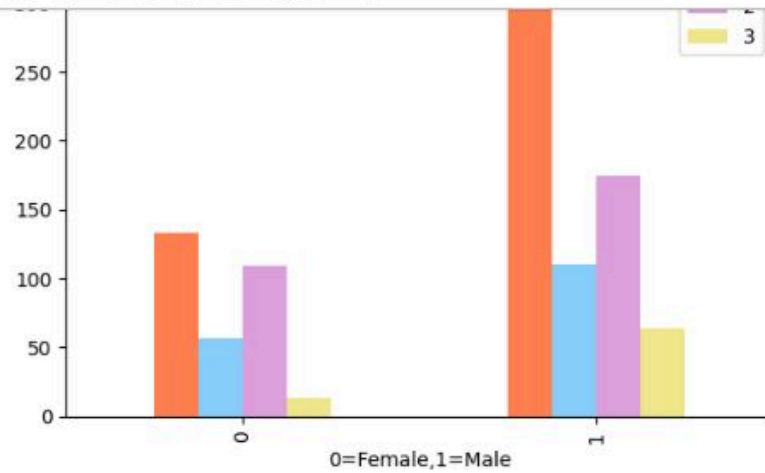
```
[15]: sex
1    713
0    312
Name: count, dtype: int64
```

```
[16]: #plotting pie chart

df.sex.value_counts().plot(kind='pie',figsize=(8,6))
plt.title('Male Female ratio')
plt.legend(['Male','Female']);
```



```
[17]: #for 2 ques
```

```
[23]: #4
pd.crosstab(df.cp, df.target)
```

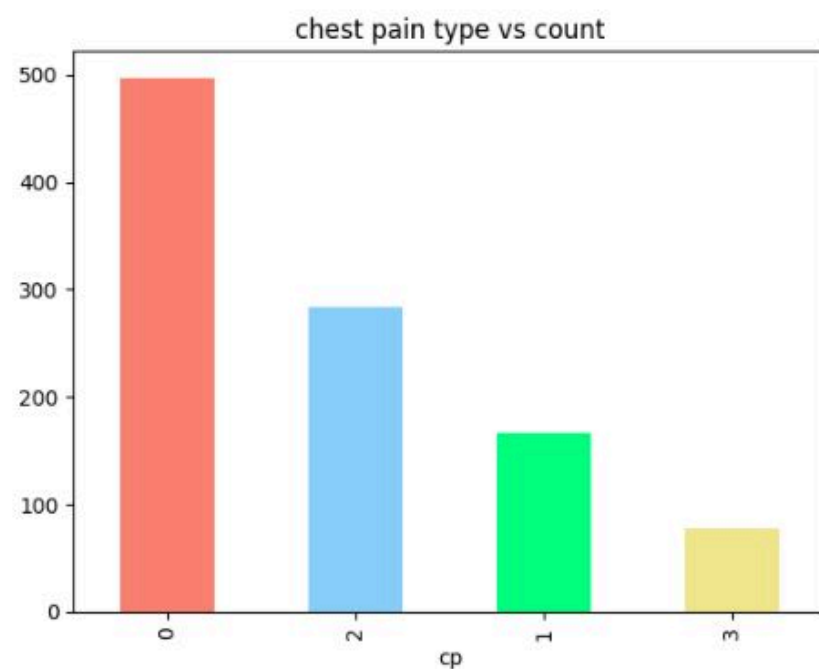
```
[23]: target  0  1
      cp
      0  375 122
      1   33 134
      2   65 219
      3   26  51
```

```
[24]: sns.countplot(x='cp', data=df, hue='target');
```



Name: count, dtype: int64

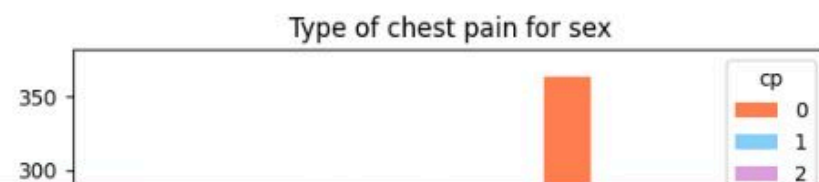
```
[20]: df.cp.value_counts().plot(kind='bar',color=['salmon','lightskyblue','springgreen','khaki'])  
plt.title('chest pain type vs count');
```

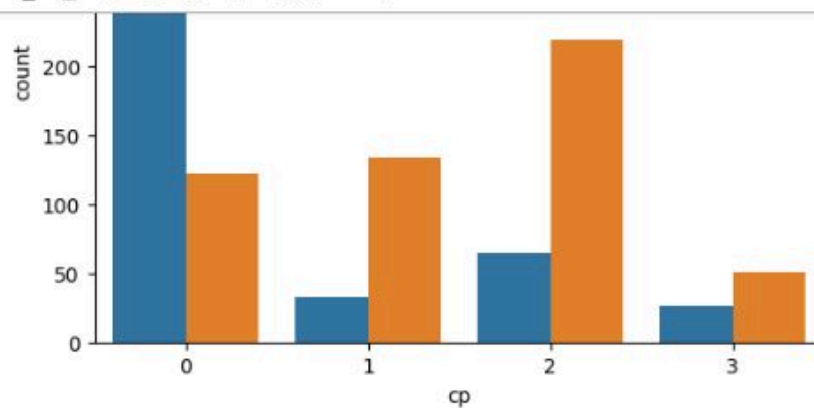


```
[21]: pd.crosstab(df.sex,df.cp)
```

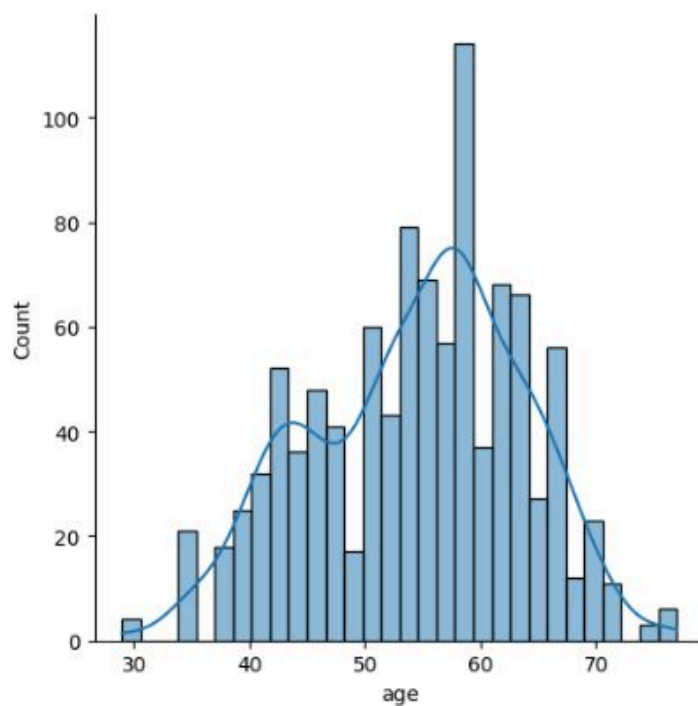
```
[21]: cp    0    1    2    3  
sex  
0  133   57  109   13  
1  364  110  175   64
```

```
[22]: pd.crosstab(df.sex,df.cp).plot(kind='bar',color=['coral','lightskyblue','plum','khaki'])  
plt.title("Type of chest pain for sex")  
plt.xlabel('0=Female,1=Male');
```

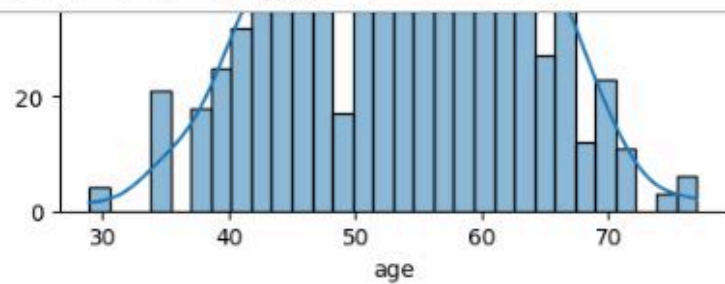




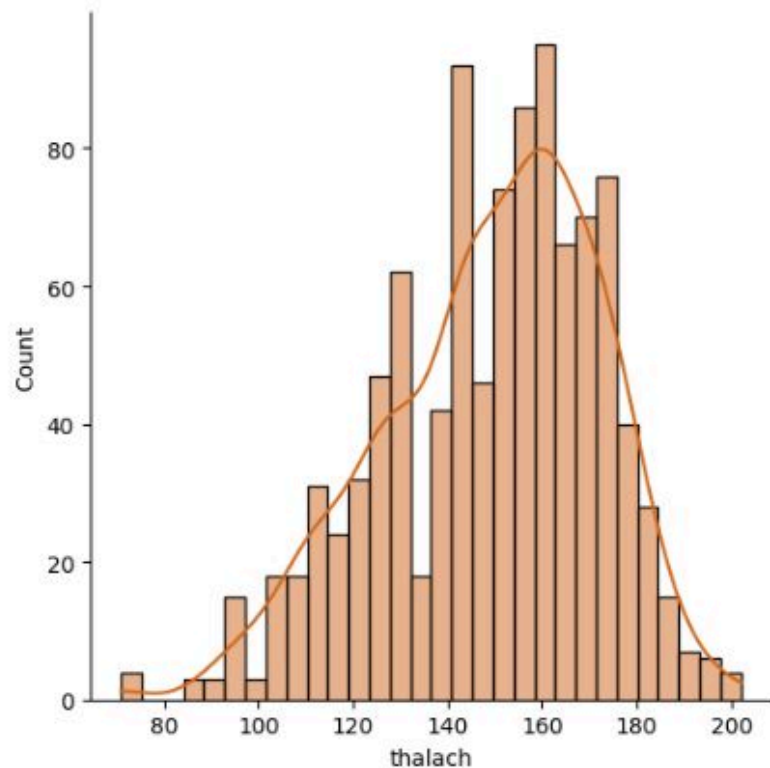
```
[25]: sns.displot(x='age',data=df,bins=30,kde=True);
```



```
[26]: sns.displot(x='thalach',data=df,bins=30,kde=True,color='chocolate');
```



```
[26]: sns.displot(x='thalach',data=df,bins=30,kde=True,color='chocolate');
```



```
[ ]:
```