

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

```
In [3]: car=pd.Series(["BMW","Ferrari","Porsche"])
car
```

```
Out[3]: 0      BMW
1    Ferrari
2    Porsche
dtype: object
```

```
In [4]: colour=pd.Series(["blue","pink","black"])
colour
```

```
Out[4]: 0    blue
1    pink
2    black
dtype: object
```

```
In [8]: car_data=pd.DataFrame({ "Brand" :car,
                               "Color" :colour})
car_data
```

```
Out[8]:   Brand  Color
0     BMW    blue
1  Ferrari   pink
2  Porsche   black
```

```
In [10]: hd=pd.read_csv("heart-disease.csv")
hd
```

Out[10]:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0	0
1	37	1	2	130	250	0	1	187	0	3.5	0	0	0
2	41	0	1	130	204	0	0	172	0	1.4	2	0	0
3	56	1	1	120	236	0	1	178	0	0.8	2	0	0
4	57	0	0	120	354	0	1	163	1	0.6	2	0	0
...
298	57	0	0	140	241	0	1	123	1	0.2	1	0	0
299	45	1	3	110	264	0	1	132	0	1.2	1	0	0
300	68	1	0	144	193	1	1	141	0	3.4	1	2	0
301	57	1	0	130	131	0	1	115	1	1.2	1	1	0
302	57	0	1	130	236	0	0	174	0	0.0	1	1	1

303 rows × 14 columns



Describing Data

In [13]:

hd.describe()

Out[13]:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	target
count	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000
mean	54.366337	0.683168	0.966997	131.623762	246.264026	0.148515	0.528000	154.000000	0.000000	0.000000	0.000000	0.000000	0.000000
std	9.082101	0.466011	1.032052	17.538143	51.830751	0.356198	0.525800	19.000000	0.000000	0.000000	0.000000	0.000000	0.000000
min	29.000000	0.000000	0.000000	94.000000	126.000000	0.000000	0.000000	100.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	47.500000	0.000000	0.000000	120.000000	211.000000	0.000000	0.000000	120.000000	0.000000	0.000000	0.000000	0.000000	0.000000
50%	55.000000	1.000000	1.000000	130.000000	240.000000	0.000000	0.000000	130.000000	0.000000	0.000000	0.000000	0.000000	0.000000
75%	61.000000	1.000000	2.000000	140.000000	274.500000	0.000000	0.000000	140.000000	0.000000	0.000000	0.000000	0.000000	0.000000
max	77.000000	1.000000	3.000000	200.000000	564.000000	1.000000	1.000000	200.000000	1.000000	2.000000	2.000000	4.000000	4.000000



In [15]:

hd.dtypes

```
Out[15]: age      int64
          sex      int64
          cp       int64
          trestbps int64
          chol     int64
          fbs      int64
          restecg  int64
          thalach  int64
          exang    int64
          oldpeak  float64
          slope    int64
          ca       int64
          thal     int64
          target   int64
          dtype: object
```

```
In [17]: hd.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 303 entries, 0 to 302
Data columns (total 14 columns):
 #   Column      Non-Null Count  Dtype  
 ---  --          --          --      
 0   age         303 non-null    int64  
 1   sex         303 non-null    int64  
 2   cp          303 non-null    int64  
 3   trestbps   303 non-null    int64  
 4   chol        303 non-null    int64  
 5   fbs         303 non-null    int64  
 6   restecg    303 non-null    int64  
 7   thalach    303 non-null    int64  
 8   exang      303 non-null    int64  
 9   oldpeak    303 non-null    float64 
 10  slope       303 non-null    int64  
 11  ca          303 non-null    int64  
 12  thal        303 non-null    int64  
 13  target      303 non-null    int64  
dtypes: float64(1), int64(13)
memory usage: 33.3 KB
```

```
In [19]: hd.std()
```

```
Out[19]: age      9.082101
          sex      0.466011
          cp       1.032052
          trestbps 17.538143
          chol     51.830751
          fbs      0.356198
          restecg  0.525860
          thalach  22.905161
          exang    0.469794
          oldpeak  1.161075
          slope    0.616226
          ca       1.022606
          thal     0.612277
          target   0.498835
          dtype: float64
```

```
In [21]: hd.mean()
```

```
Out[21]: age      54.366337
          sex      0.683168
          cp       0.966997
          trestbps 131.623762
          chol     246.264026
          fbs      0.148515
          restecg  0.528053
          thalach  149.646865
          exang    0.326733
          oldpeak  1.039604
          slope    1.399340
          ca       0.729373
          thal     2.313531
          target   0.544554
          dtype: float64
```

In [23]: `hd.mode()`

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	c
0	58.0	1.0	0.0	120.0	197	0.0	1.0	162.0	0.0	0.0	2.0	0.
1	NaN	NaN	NaN	NaN	204	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	234	NaN	NaN	NaN	NaN	NaN	NaN	NaN



In [25]: `hd.var()`

```
Out[25]: age      82.484558
          sex      0.217166
          cp       1.065132
          trestbps 307.586453
          chol     2686.426748
          fbs      0.126877
          restecg  0.276528
          thalach  524.646406
          exang    0.220707
          oldpeak  1.348095
          slope    0.379735
          ca       1.045724
          thal     0.374883
          target   0.248836
          dtype: float64
```

In [27]: `hd.sum()`

```
Out[27]: age      16473.0
          sex       207.0
          cp        293.0
          trestbps 39882.0
          chol     74618.0
          fbs       45.0
          restecg   160.0
          thalach   45343.0
          exang     99.0
          oldpeak   315.0
          slope     424.0
          ca        221.0
          thal      701.0
          target    165.0
          dtype: float64
```

```
In [29]: hd.median()
```

```
Out[29]: age      55.0
          sex       1.0
          cp        1.0
          trestbps 130.0
          chol     240.0
          fbs       0.0
          restecg   1.0
          thalach   153.0
          exang     0.0
          oldpeak   0.8
          slope     1.0
          ca        0.0
          thal      2.0
          target    1.0
          dtype: float64
```

```
In [31]: hd.std()
```

```
Out[31]: age      9.082101
          sex      0.466011
          cp       1.032052
          trestbps 17.538143
          chol     51.830751
          fbs      0.356198
          restecg  0.525860
          thalach   22.905161
          exang    0.469794
          oldpeak   1.161075
          slope     0.616226
          ca       1.022606
          thal     0.612277
          target    0.498835
          dtype: float64
```

```
In [33]: hd.index
```

```
Out[33]: RangeIndex(start=0, stop=303, step=1)
```

```
In [35]: len(hd)
```

```
Out[35]: 303
```

Viewing and Selection of Data

In [38]: `hd.tail()`

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	target
298	57	0	0	140	241	0	1	123	1	0.2	1	0	
299	45	1	3	110	264	0	1	132	0	1.2	1	0	
300	68	1	0	144	193	1	1	141	0	3.4	1	2	
301	57	1	0	130	131	0	1	115	1	1.2	1	1	
302	57	0	1	130	236	0	0	174	0	0.0	1	1	

In [40]: `hd.head(10)`

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2
5	57	1	0	140	192	0	1	148	0	0.4	1	0	1
6	56	0	1	140	294	0	0	153	0	1.3	1	0	2
7	44	1	1	120	263	0	1	173	0	0.0	2	0	3
8	52	1	2	172	199	1	1	162	0	0.5	2	0	3
9	57	1	2	150	168	0	1	174	0	1.6	2	0	2

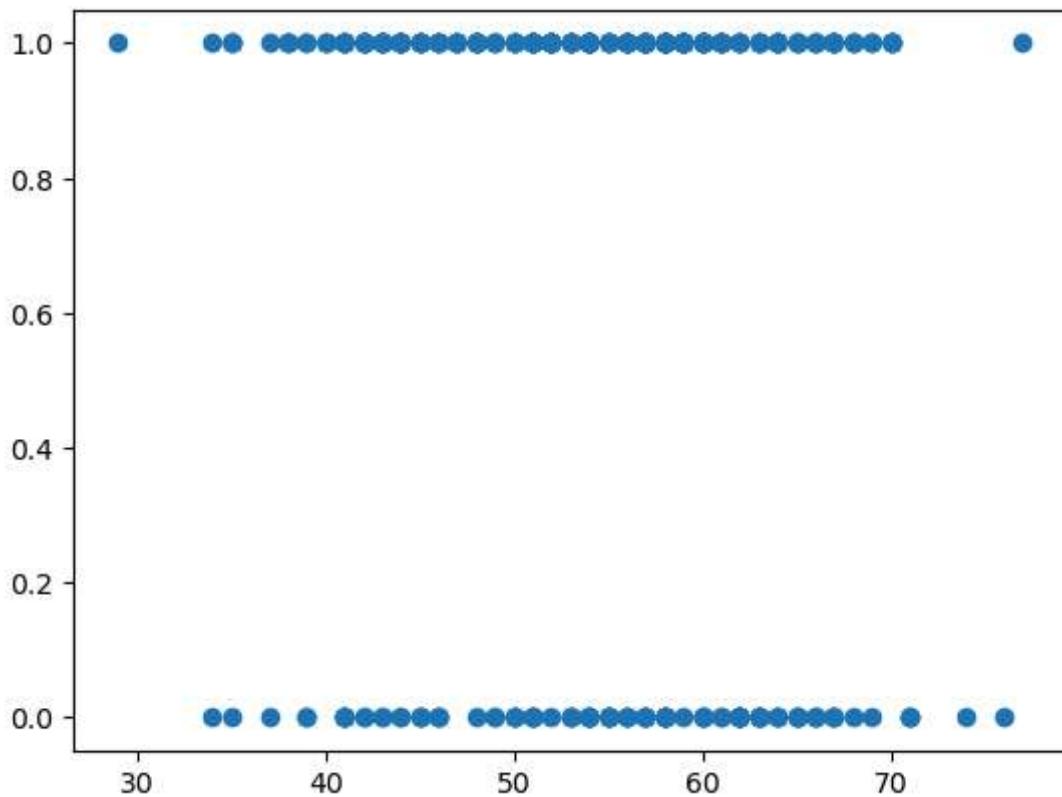
In [42]: `hd.columns`

```
Out[42]: Index(['age', 'sex', 'cp', 'trestbps', 'chol', 'fbs', 'restecg', 'thalach',
       'exang', 'oldpeak', 'slope', 'ca', 'thal', 'target'],
       dtype='object')
```

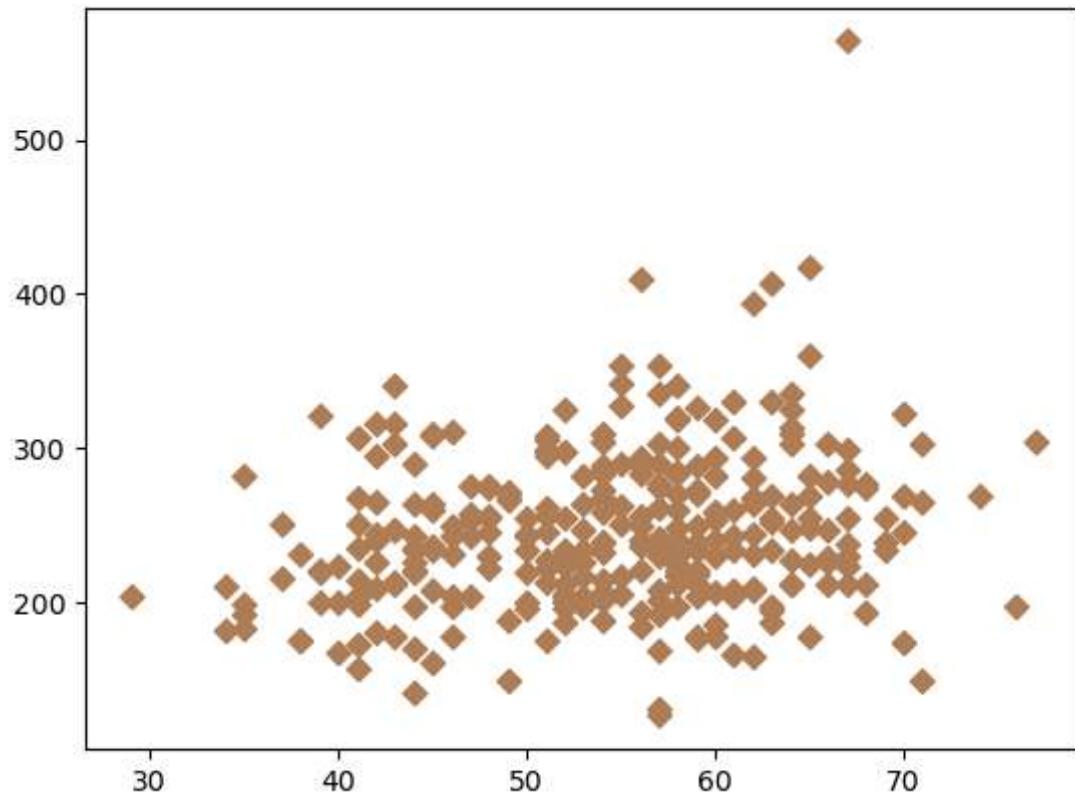
In [44]: `hd['age']`

```
Out[44]: 0      63
         1      37
         2      41
         3      56
         4      57
         ..
        298     57
        299     45
        300     68
        301     57
        302     57
Name: age, Length: 303, dtype: int64
```

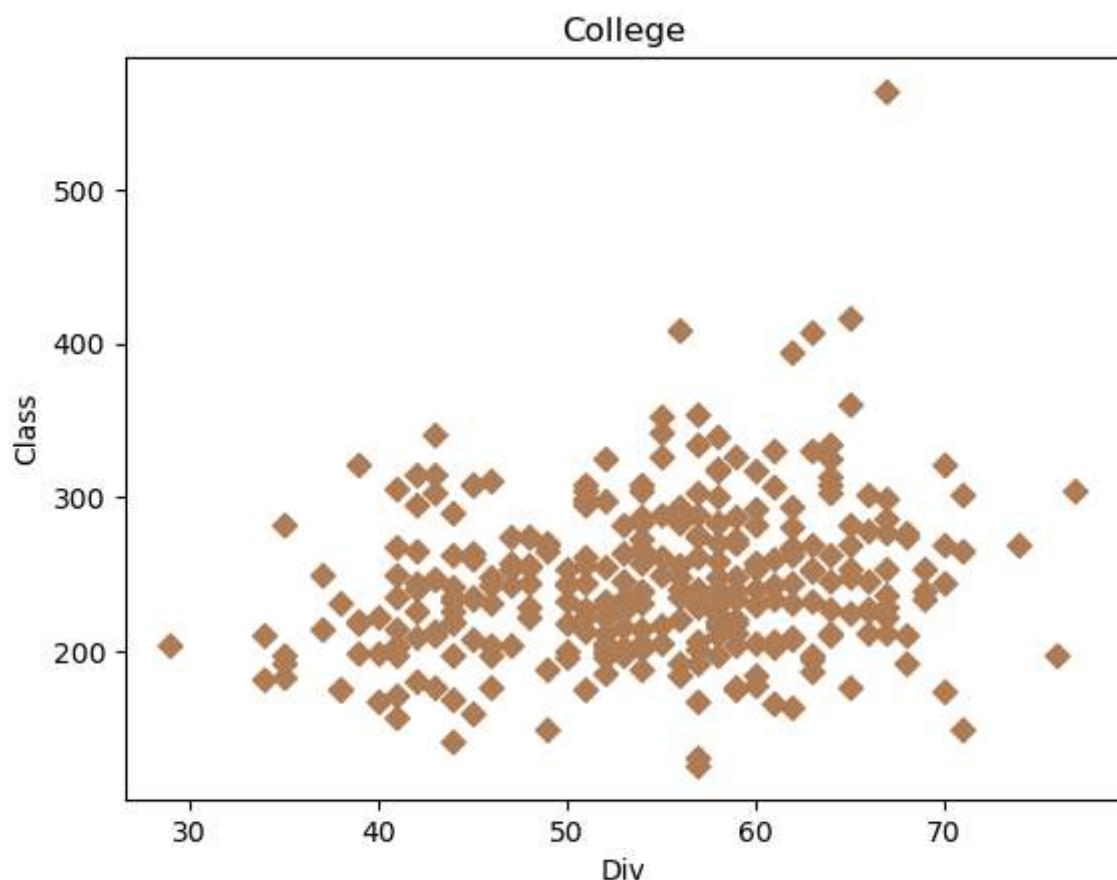
```
In [46]: x=hd['age']
y=hd['sex']
plt.scatter(x,y)
plt.show()
```



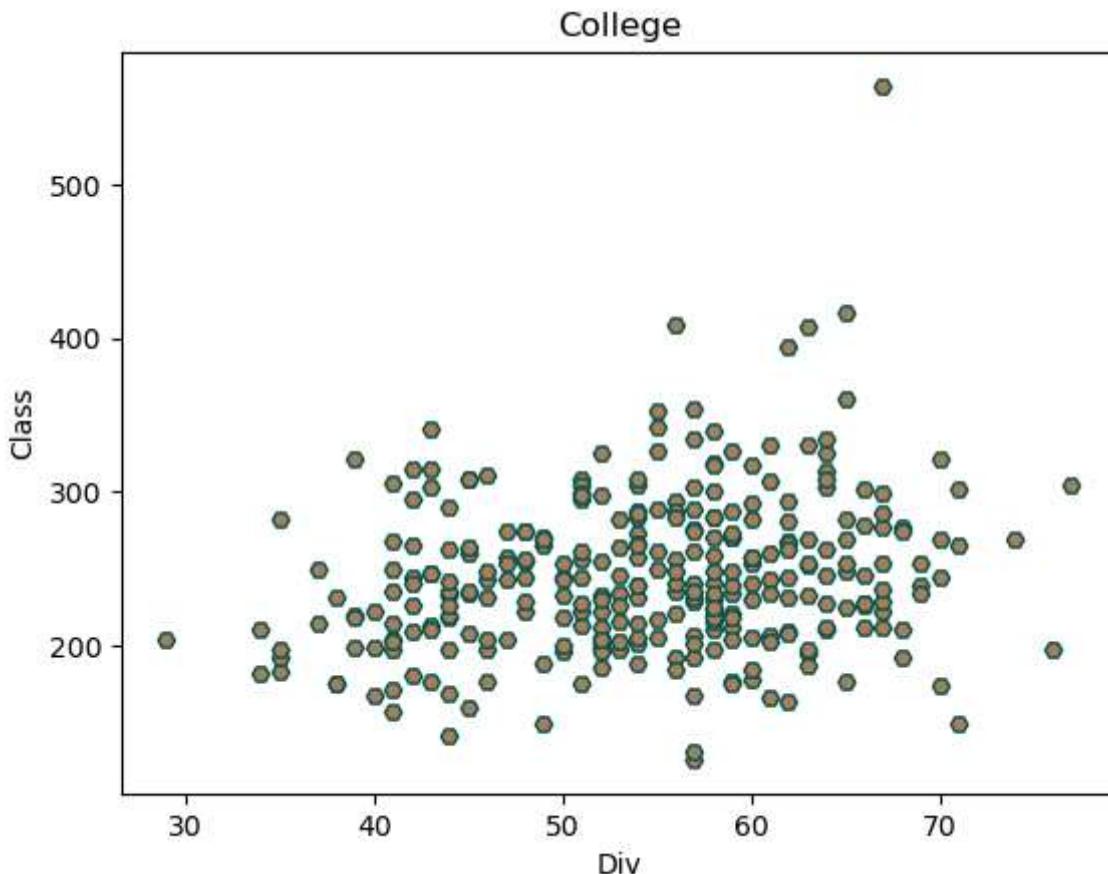
```
In [47]: x=hd['age']
y=hd['chol']
plt.scatter(x,y,marker='D',c='#B17F59')
plt.show()
```



```
In [49]: x=hd['age']
y=hd['chol']
plt.scatter(x,y,marker='D',c='#B17F59')
plt.title("College")
plt.xlabel("Div")
plt.ylabel("Class")
plt.show()
```

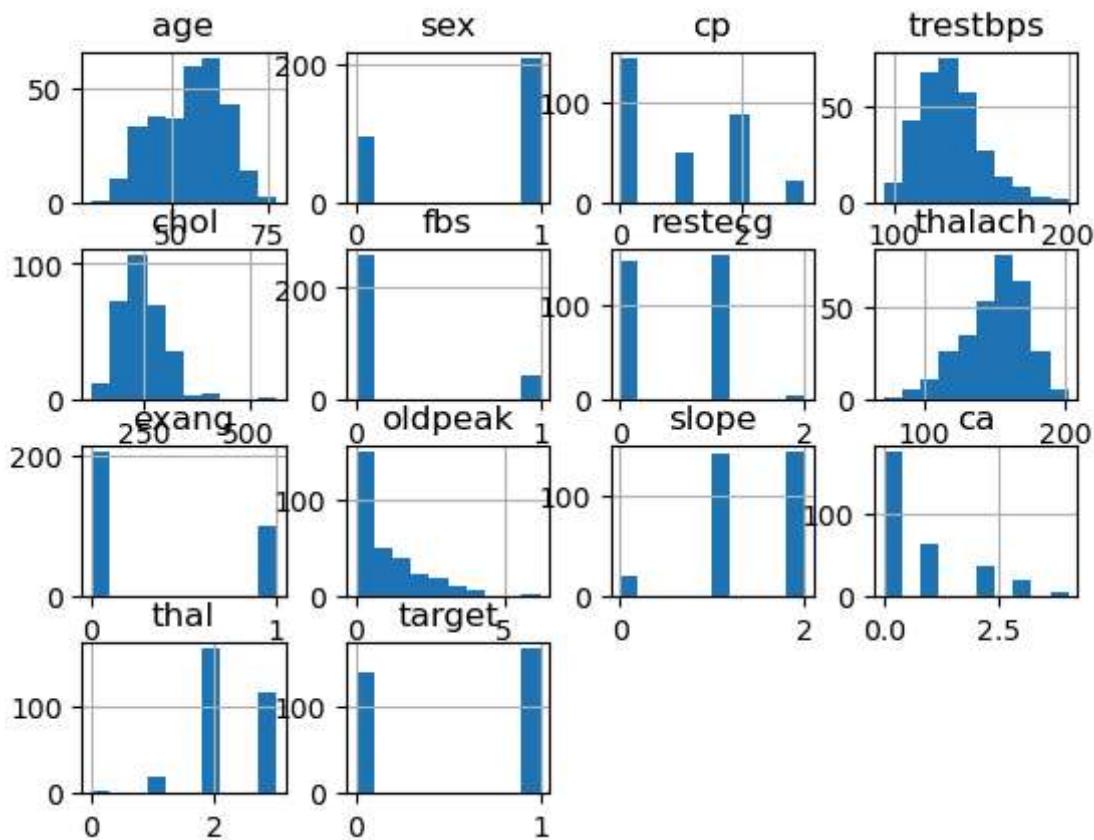


```
In [51]: x=hd['age']
y=hd['chol']
plt.scatter(x,y,marker='H',c='#B17F59',edgecolor='#015551')
plt.title("College")
plt.xlabel("Div")
plt.ylabel("Class")
plt.show()
```



```
In [53]: hd.hist()
```

```
Out[53]: array([[[<Axes: title={'center': 'age'}>, <Axes: title={'center': 'sex'}>,
   <Axes: title={'center': 'cp'}>,
   <Axes: title={'center': 'trestbps'}>],
  [<Axes: title={'center': 'chol'}>,
   <Axes: title={'center': 'fbs'}>,
   <Axes: title={'center': 'restecg'}>,
   <Axes: title={'center': 'thalach'}>],
  [<Axes: title={'center': 'exang'}>,
   <Axes: title={'center': 'oldpeak'}>,
   <Axes: title={'center': 'slope'}>,
   <Axes: title={'center': 'ca'}>],
  [<Axes: title={'center': 'thal'}>,
   <Axes: title={'center': 'target'}>, <Axes: >, <Axes: >]],  
  dtype=object)
```



```
In [55]: hd['target']
```

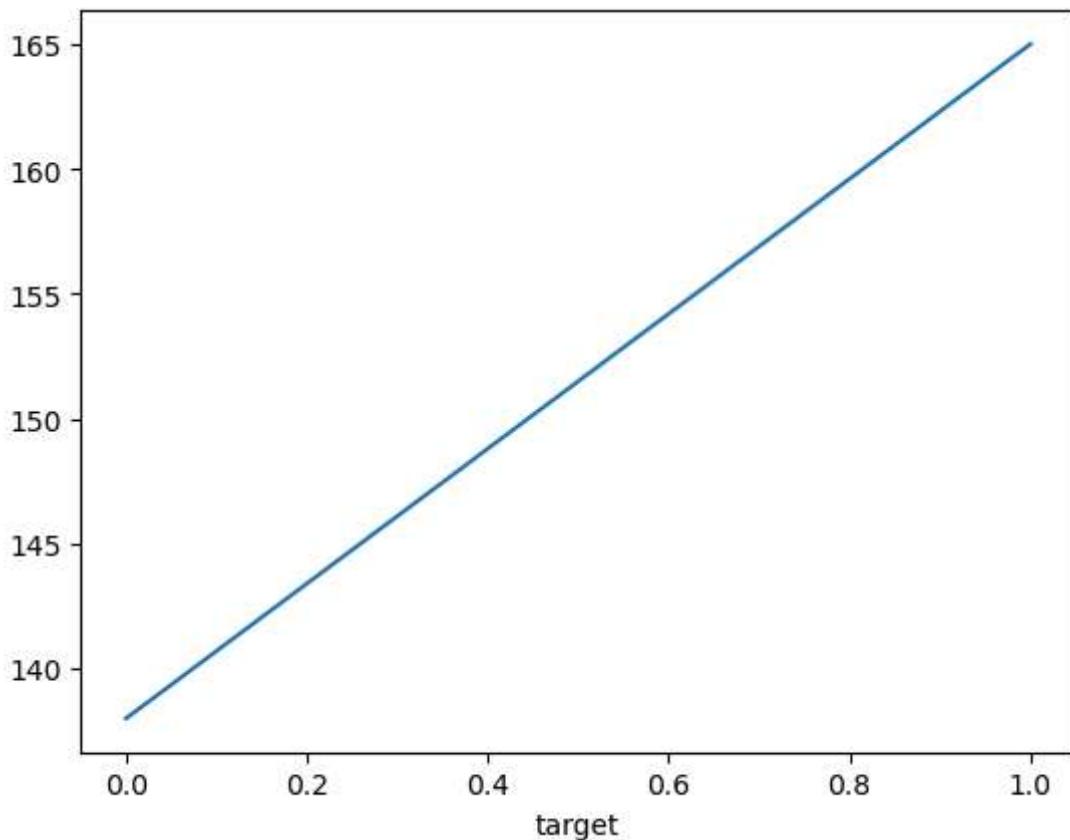
```
Out[55]: 0      1
         1      1
         2      1
         3      1
         4      1
          ..
        298     0
        299     0
        300     0
        301     0
        302     0
Name: target, Length: 303, dtype: int64
```

```
In [56]: hd['target'].value_counts()
```

```
Out[56]: target
1    165
0    138
Name: count, dtype: int64
```

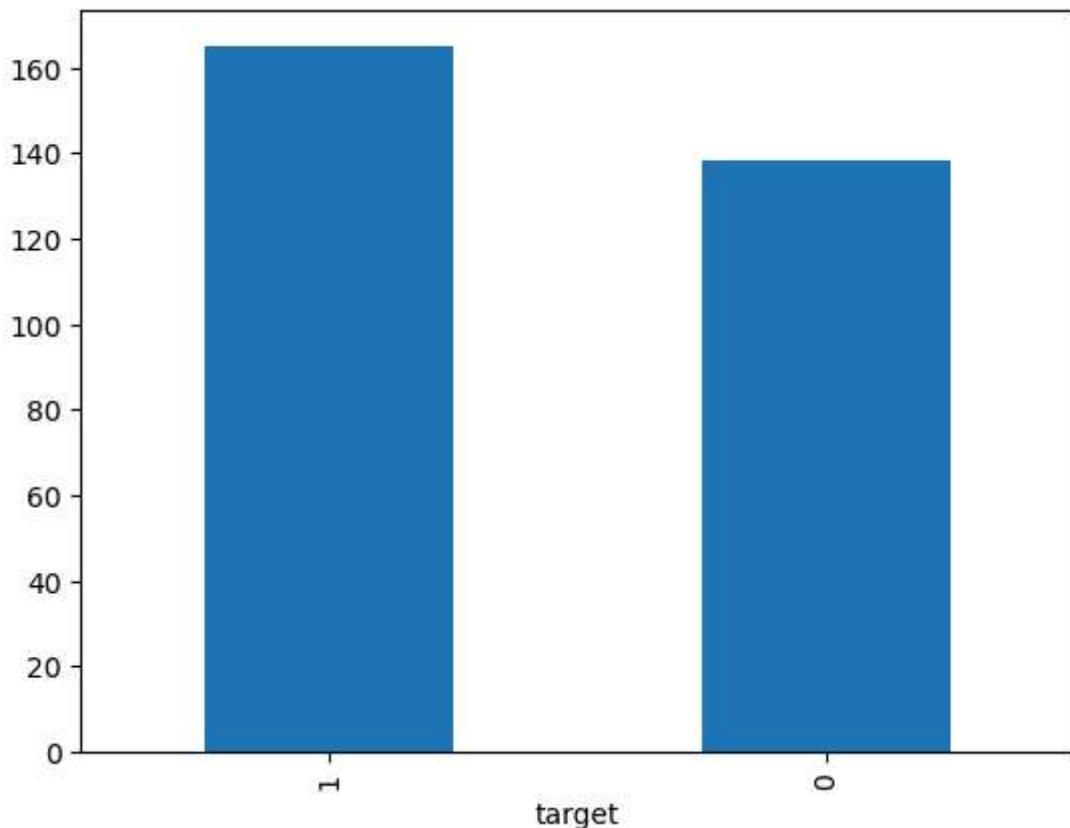
```
In [57]: hd['target'].value_counts().plot()
```

```
Out[57]: <Axes: xlabel='target'>
```



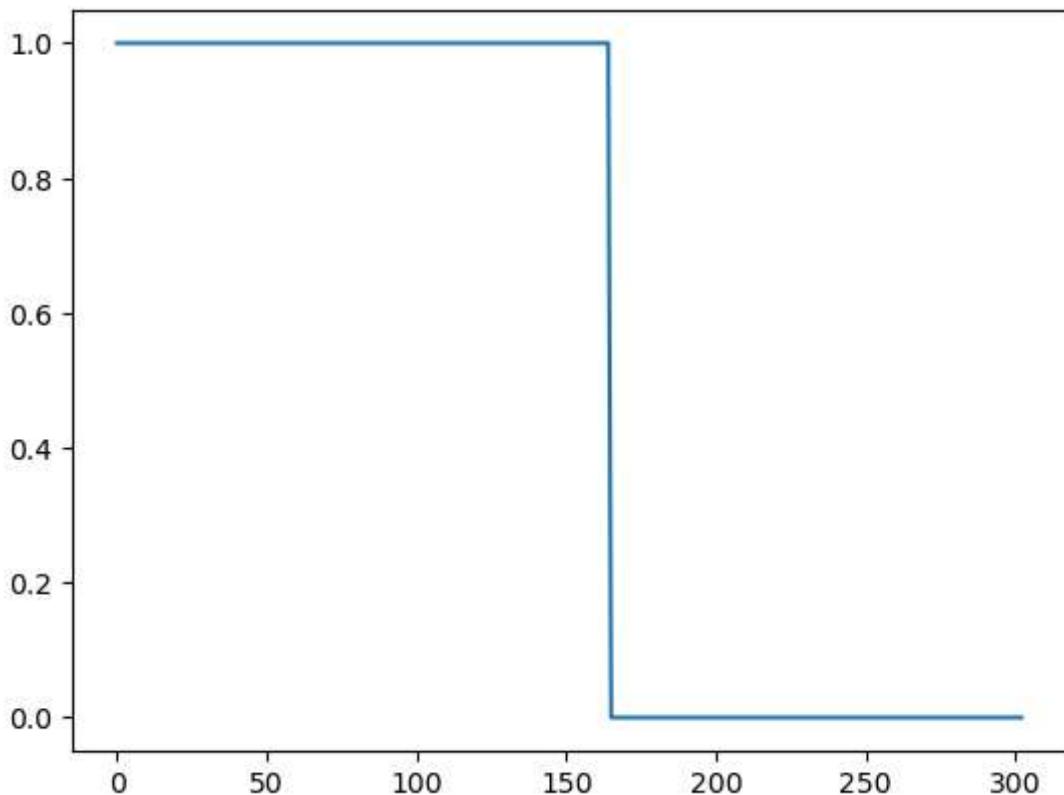
```
In [58]: hd['target'].value_counts().plot.bar()
```

```
Out[58]: <Axes: xlabel='target'>
```



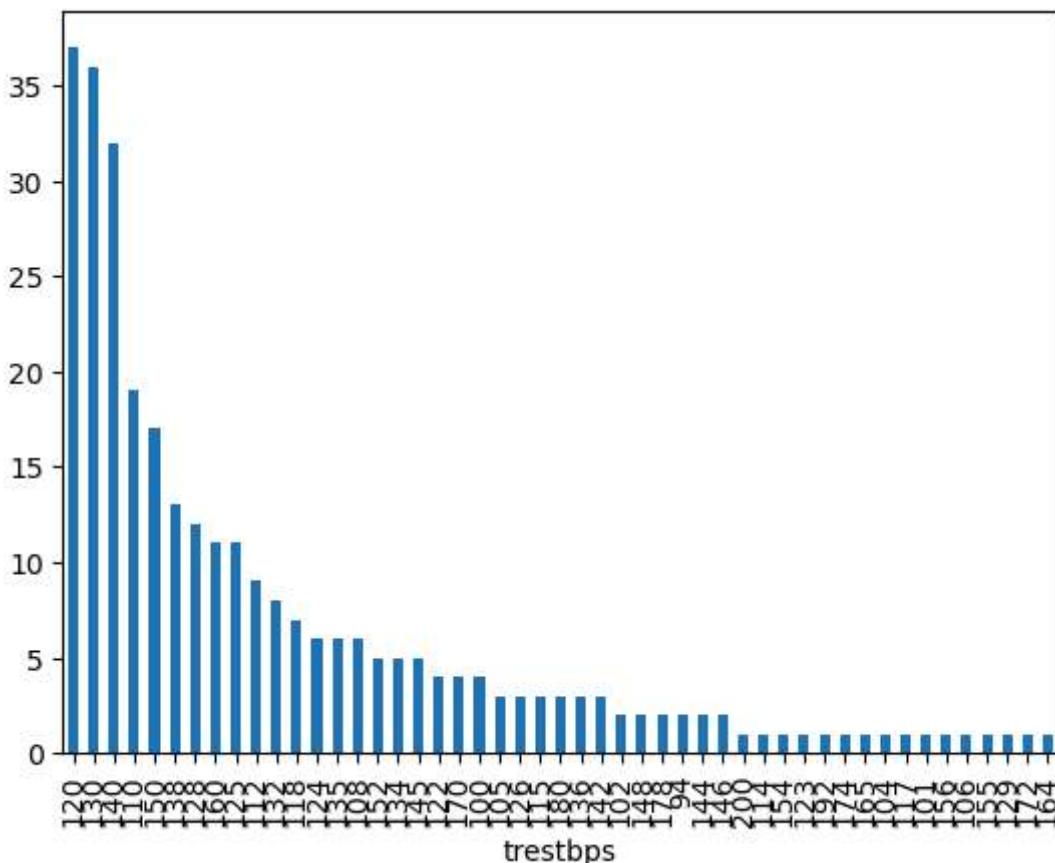
```
In [59]: hd['target'].plot()
```

```
Out[59]: <Axes: >
```



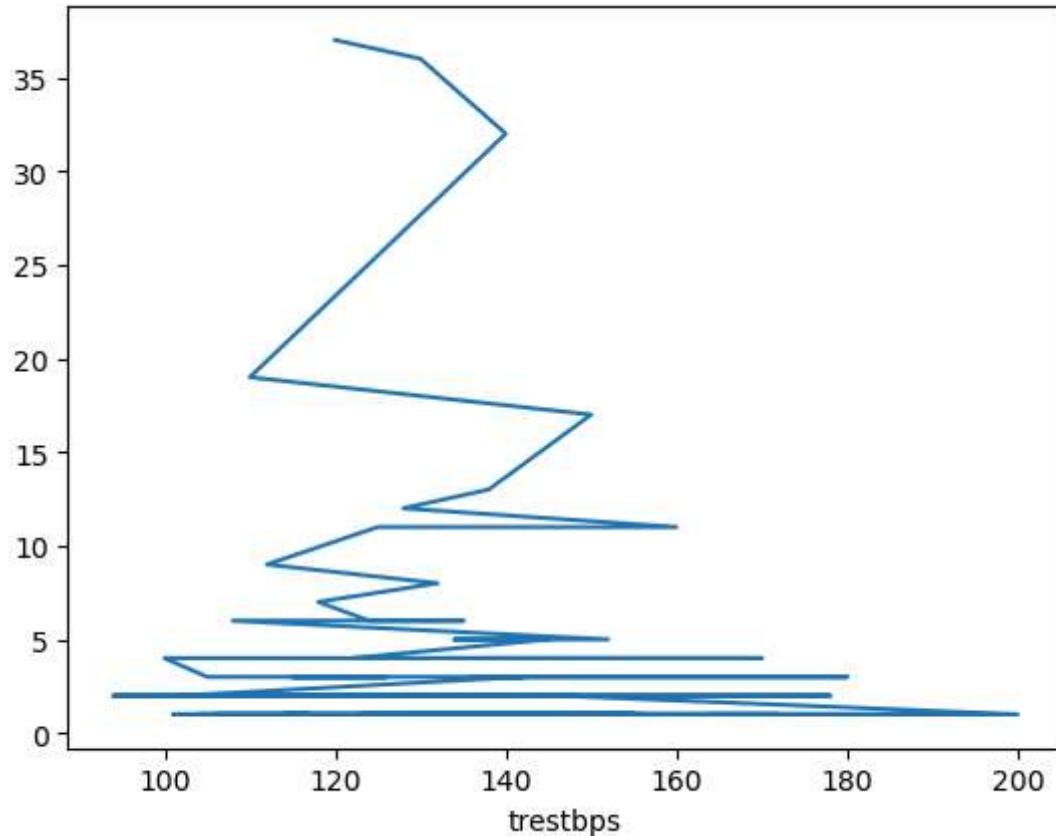
```
In [60]: hd['trestbps'].value_counts().plot.bar()
```

```
Out[60]: <Axes: xlabel='trestbps'>
```



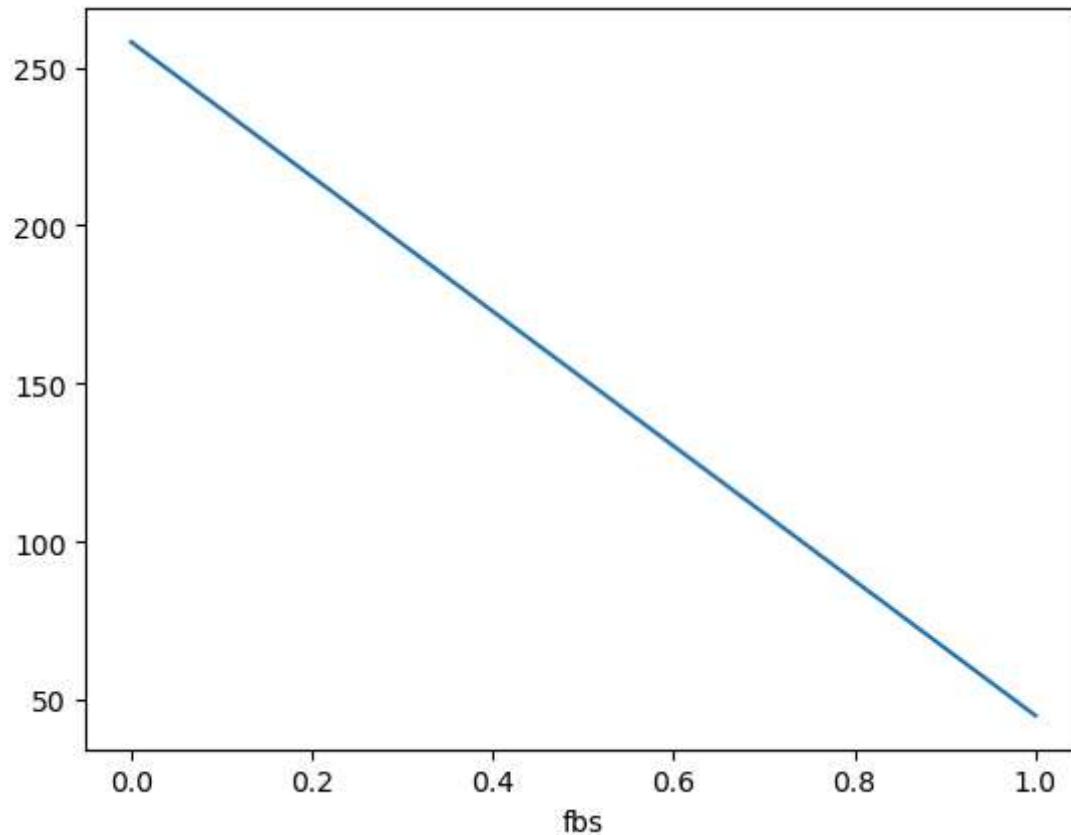
```
In [61]: hd['trestbps'].value_counts().plot.line()
```

```
Out[61]: <Axes: xlabel='trestbps'>
```



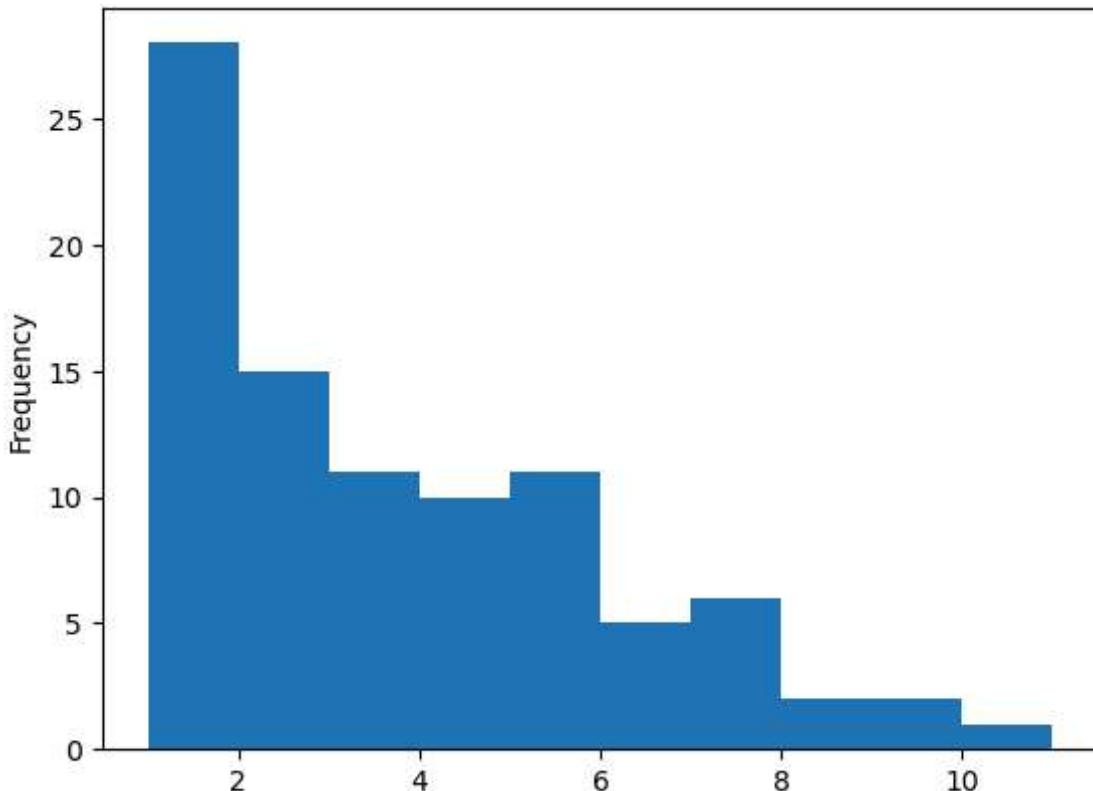
```
In [62]: hd['fbs'].value_counts().plot.line()
```

```
Out[62]: <Axes: xlabel='fbs'>
```



```
In [63]: hd['thalach'].value_counts().plot.hist()
```

```
Out[63]: <Axes: ylabel='Frequency'>
```



loc & iloc

```
In [68]: pu=pd.Series(["hii","hello","gm","gn"], index=[1,2,2,3])
pu
```

```
Out[68]: 1      hii
          2    hello
          2      gm
          3      gn
          dtype: object
```

```
In [69]: pu.loc[2]
```

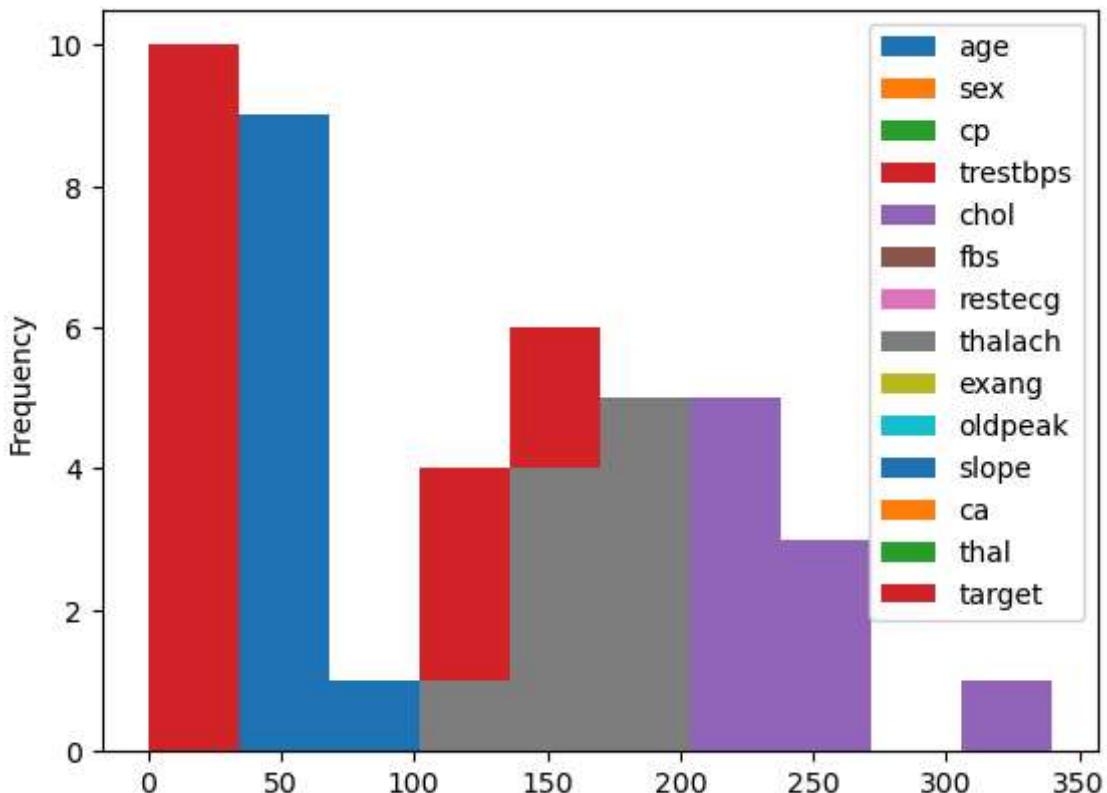
```
Out[69]: 2    hello
          2      gm
          dtype: object
```

```
In [71]: pu.iloc[0]
```

```
Out[71]: 'hii'
```

```
In [73]: hd.iloc[15:25].plot.hist()
```

```
Out[73]: <Axes: ylabel='Frequency'>
```



Conditional Filtering

`==,>,>=,<,<=,! =`

In [76]: `hd[hd["target"]==1]`

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0		
1	37	1	2	130	250	0	1	187	0	3.5	0	0		
2	41	0	1	130	204	0	0	172	0	1.4	2	0		
3	56	1	1	120	236	0	1	178	0	0.8	2	0		
4	57	0	0	120	354	0	1	163	1	0.6	2	0		
...	
160	56	1	1	120	240	0	1	169	0	0.0	0	0		
161	55	0	1	132	342	0	1	166	0	1.2	2	0		
162	41	1	1	120	157	0	1	182	0	0.0	2	0		
163	38	1	2	138	175	0	1	173	0	0.0	2	4		
164	38	1	2	138	175	0	1	173	0	0.0	2	4		

165 rows × 14 columns



In [77]: `hd[hd['age']>=40]`

Out[77]:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	tl
0	63	1	3	145	233	1	0	150	0	2.3	0	0	0
2	41	0	1	130	204	0	0	172	0	1.4	2	0	0
3	56	1	1	120	236	0	1	178	0	0.8	2	0	0
4	57	0	0	120	354	0	1	163	1	0.6	2	0	0
5	57	1	0	140	192	0	1	148	0	0.4	1	0	0
...
298	57	0	0	140	241	0	1	123	1	0.2	1	0	0
299	45	1	3	110	264	0	1	132	0	1.2	1	0	0
300	68	1	0	144	193	1	1	141	0	3.4	1	2	0
301	57	1	0	130	131	0	1	115	1	1.2	1	1	0
302	57	0	1	130	236	0	0	174	0	0.0	1	1	1

287 rows × 14 columns



In [79]:

hd[hd['chol']!=264]

Out[79]:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	tl
0	63	1	3	145	233	1	0	150	0	2.3	0	0	0
1	37	1	2	130	250	0	1	187	0	3.5	0	0	0
2	41	0	1	130	204	0	0	172	0	1.4	2	0	0
3	56	1	1	120	236	0	1	178	0	0.8	2	0	0
4	57	0	0	120	354	0	1	163	1	0.6	2	0	0
...
297	59	1	0	164	176	1	0	90	0	1.0	1	2	0
298	57	0	0	140	241	0	1	123	1	0.2	1	0	0
300	68	1	0	144	193	1	1	141	0	3.4	1	2	0
301	57	1	0	130	131	0	1	115	1	1.2	1	1	0
302	57	0	1	130	236	0	0	174	0	0.0	1	1	1

301 rows × 14 columns



In [81]:

hd[hd['oldpeak']<2.0]

Out[81]:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	tl
2	41	0	1	130	204	0	0	172	0	1.4	2	0	
3	56	1	1	120	236	0	1	178	0	0.8	2	0	
4	57	0	0	120	354	0	1	163	1	0.6	2	0	
5	57	1	0	140	192	0	1	148	0	0.4	1	0	
6	56	0	1	140	294	0	0	153	0	1.3	1	0	
...
297	59	1	0	164	176	1	0	90	0	1.0	1	2	
298	57	0	0	140	241	0	1	123	1	0.2	1	0	
299	45	1	3	110	264	0	1	132	0	1.2	1	0	
301	57	1	0	130	131	0	1	115	1	1.2	1	1	
302	57	0	1	130	236	0	0	174	0	0.0	1	1	

244 rows × 14 columns

In [83]: `hd[hd['thalach']==123]`

Out[83]:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	tl
27	51	1	2	110	175	0	1	123	0	0.6	2	0	
298	57	0	0	140	241	0	1	123	1	0.2	1	0	

Adding new col with applying some maths

In [89]: `hd['pro_data']=hd['trestbps']+hd['chol']`
`hd`

Out[89]:

	age	sex	cp	trestbps	chol	fb	restecg	thalach	exang	oldpeak	slope	ca	ti
0	63	1	3	145	233	1	0	150	0	2.3	0	0	0
1	37	1	2	130	250	0	1	187	0	3.5	0	0	0
2	41	0	1	130	204	0	0	172	0	1.4	2	0	0
3	56	1	1	120	236	0	1	178	0	0.8	2	0	0
4	57	0	0	120	354	0	1	163	1	0.6	2	0	0
...
298	57	0	0	140	241	0	1	123	1	0.2	1	0	0
299	45	1	3	110	264	0	1	132	0	1.2	1	0	0
300	68	1	0	144	193	1	1	141	0	3.4	1	2	0
301	57	1	0	130	131	0	1	115	1	1.2	1	1	0
302	57	0	1	130	236	0	0	174	0	0.0	1	1	1

303 rows × 15 columns



col drop method

In [91]:

```
hd.drop('pro_data',axis=1,inplace=True)
hd
```

Out[91]:

	age	sex	cp	trestbps	chol	fb	restecg	thalach	exang	oldpeak	slope	ca	ti
0	63	1	3	145	233	1	0	150	0	2.3	0	0	0
1	37	1	2	130	250	0	1	187	0	3.5	0	0	0
2	41	0	1	130	204	0	0	172	0	1.4	2	0	0
3	56	1	1	120	236	0	1	178	0	0.8	2	0	0
4	57	0	0	120	354	0	1	163	1	0.6	2	0	0
...
298	57	0	0	140	241	0	1	123	1	0.2	1	0	0
299	45	1	3	110	264	0	1	132	0	1.2	1	0	0
300	68	1	0	144	193	1	1	141	0	3.4	1	2	0
301	57	1	0	130	131	0	1	115	1	1.2	1	1	0
302	57	0	1	130	236	0	0	174	0	0.0	1	1	1

303 rows × 14 columns



In [92]:

```
hd.drop('thal',axis=1,inplace=True)
hd
```

Out[92]:

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0	0
1	37	1	2	130	250	0	1	187	0	3.5	0	0	0
2	41	0	1	130	204	0	0	172	0	1.4	2	0	0
3	56	1	1	120	236	0	1	178	0	0.8	2	0	0
4	57	0	0	120	354	0	1	163	1	0.6	2	0	0
...
298	57	0	0	140	241	0	1	123	1	0.2	1	0	0
299	45	1	3	110	264	0	1	132	0	1.2	1	0	0
300	68	1	0	144	193	1	1	141	0	3.4	1	2	0
301	57	1	0	130	131	0	1	115	1	1.2	1	1	0
302	57	0	1	130	236	0	0	174	0	0.0	1	1	1

303 rows × 13 columns



cross tab

In [99]: `hd.nunique()`

Out[99]:

age	41
sex	2
cp	4
trestbps	49
chol	152
fbs	2
restecg	3
thalach	91
exang	2
oldpeak	40
slope	3
ca	5
target	2
dtype: int64	

In [102...]: `pd.crosstab(hd['sex'], hd['cp'])`

Out[102...]:

	cp	0	1	2	3
sex					
0	39	18	35	4	
1	104	32	52	19	

In [106...]: `pd.crosstab(hd['sex'], hd['target'])`

```
Out[106... target  0  1
```

sex

0	24	72
1	114	93

```
In [109... pd.crosstab(hd['sex'],hd['restecg'])
```

```
Out[109... restecg  0  1  2
```

sex

0	44	49	3
1	103	103	1

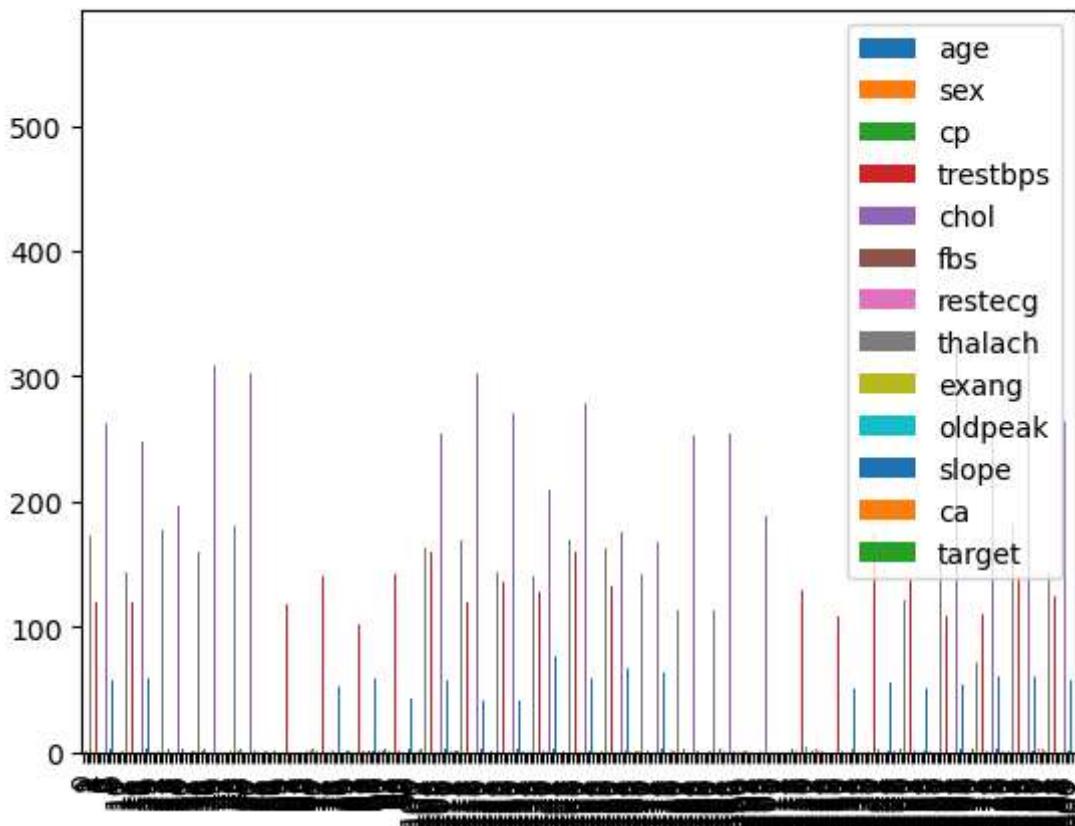
Sp. Graphs

```
In [112... hd.plot(kind='bar')  
hd
```

```
Out[112...      age  sex  cp  trestbps  chol  fbs  restecg  thalach  exang  oldpeak  slope  ca  t:  
          0    63    1    3      145   233    1      0     150      0     2.3      0    0  
          1    37    1    2      130   250    0      1     187      0     3.5      0    0  
          2    41    0    1      130   204    0      0     172      0     1.4      2    0  
          3    56    1    1      120   236    0      1     178      0     0.8      2    0  
          4    57    0    0      120   354    0      1     163      1     0.6      2    0  
          ...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ...  
          298   57    0    0      140   241    0      1     123      1     0.2      1    0  
          299   45    1    3      110   264    0      1     132      0     1.2      1    0  
          300   68    1    0      144   193    1      1     141      0     3.4      1    2  
          301   57    1    0      130   131    0      1     115      1     1.2      1    1  
          302   57    0    1      130   236    0      0     174      0     0.0      1    1
```

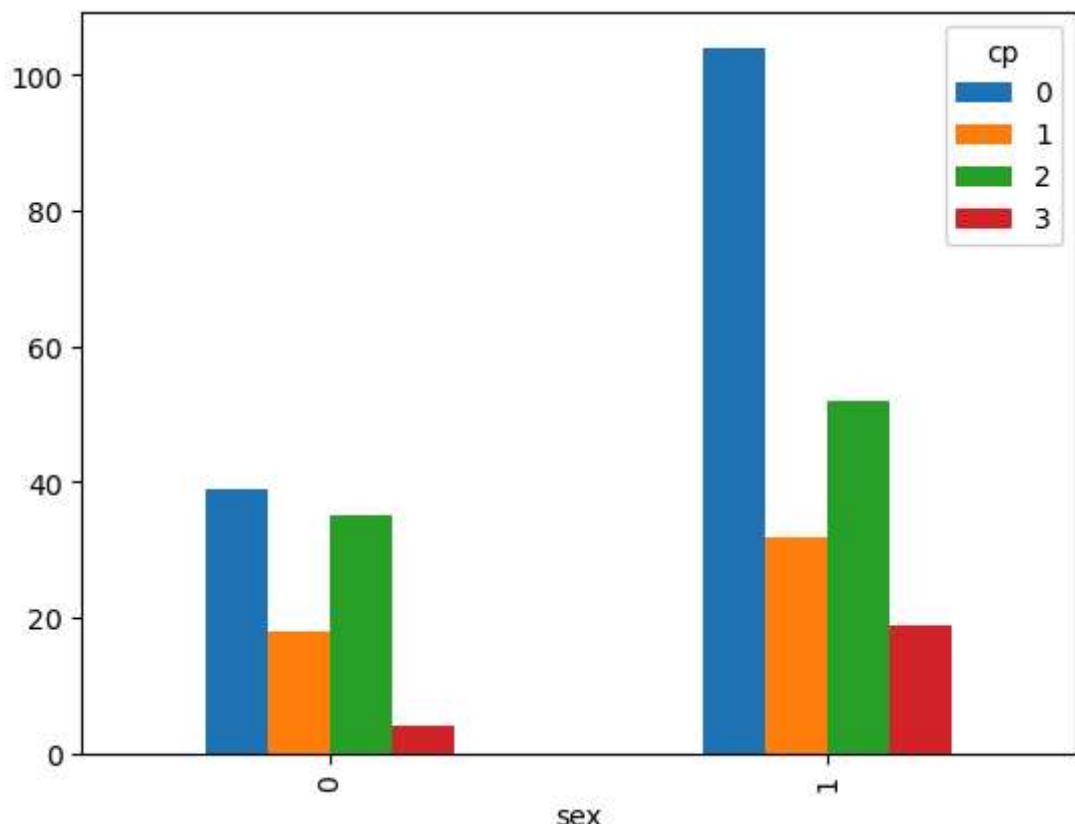
303 rows × 13 columns





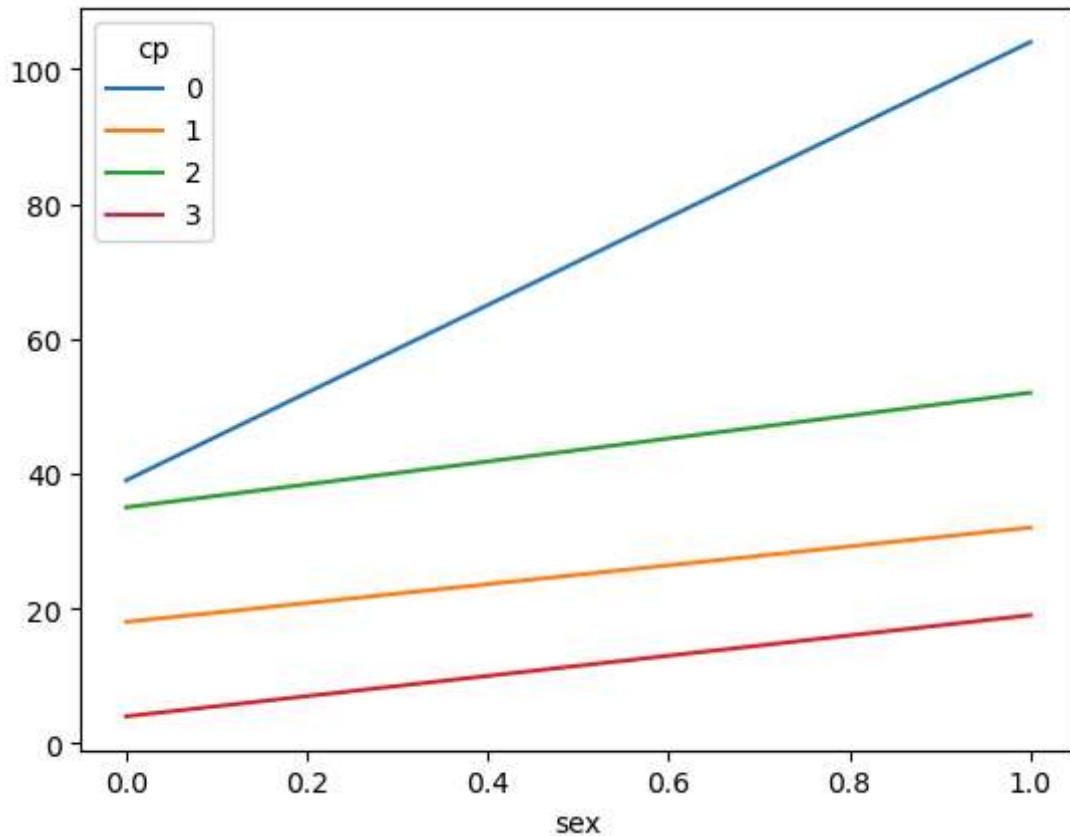
```
In [113... pd.crosstab(hd['sex'],hd['cp']).plot.bar()
```

```
Out[113... <Axes: xlabel='sex'>
```



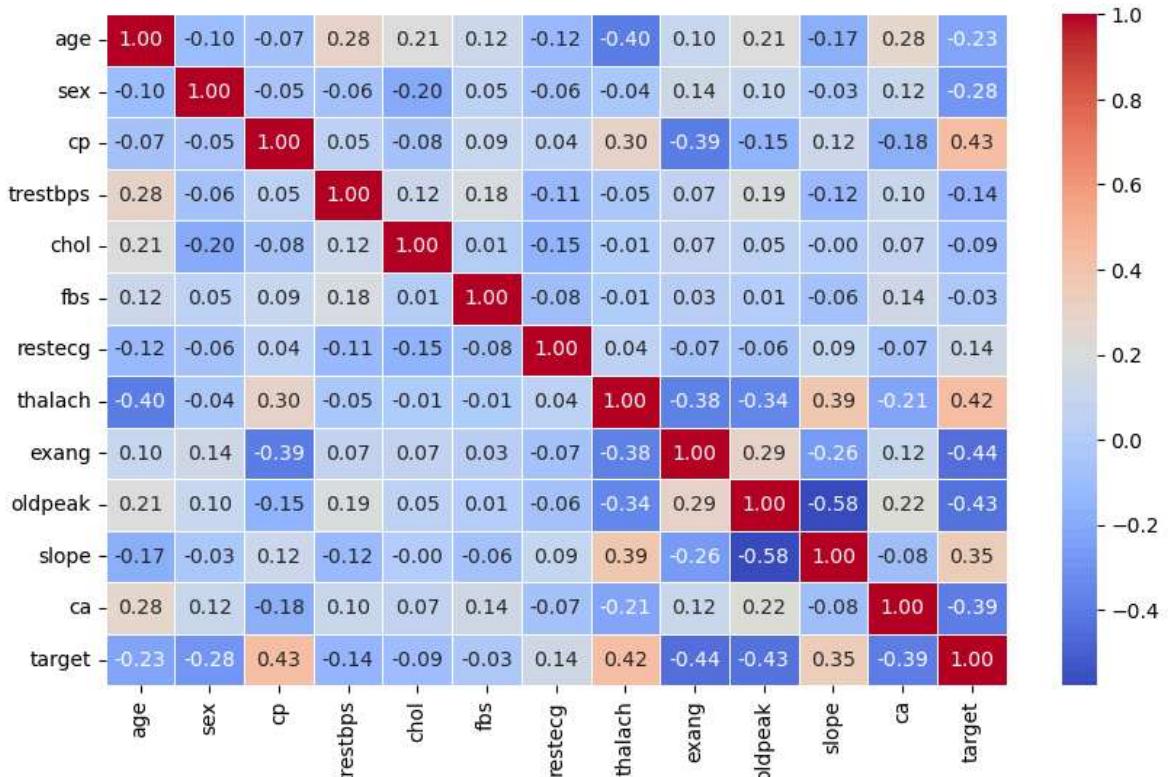
```
In [114... pd.crosstab(hd['sex'],hd['cp']).plot.line()
```

```
Out[114... <Axes: xlabel='sex'>
```



In [115...]

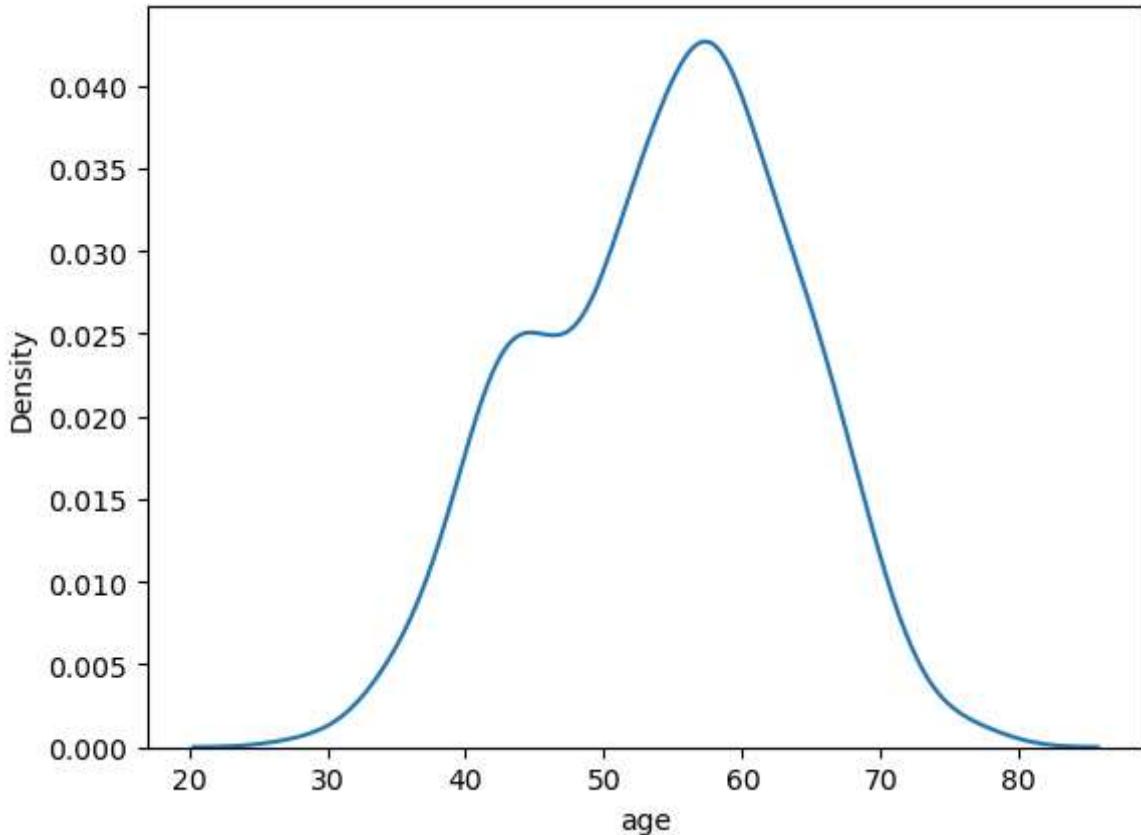
```
import seaborn as sns
correlation=hd.corr()
plt.figure(figsize=(10,6))
sns.heatmap(correlation, annot=True, cmap='coolwarm',
            fmt='.2f', linewidths=0.5, cbar=True, )
plt.show()
```



In [116...]

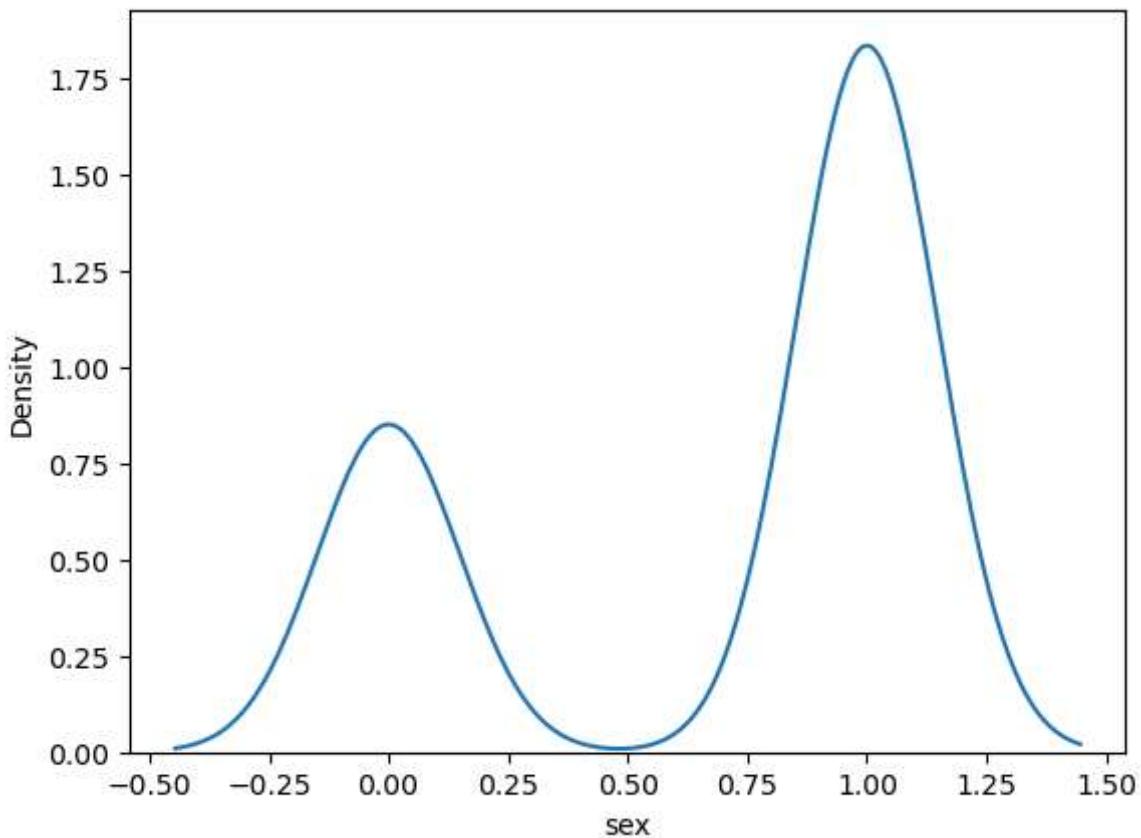
```
sns.kdeplot(hd["age"])
```

```
Out[116... <Axes: xlabel='age', ylabel='Density'>
```



```
In [117... sns.kdeplot(hd["sex"])
```

```
Out[117... <Axes: xlabel='sex', ylabel='Density'>
```



```
In [118... plt.subplot(2,2,1)  
sns.kdeplot(hd["oldpeak"])]
```

```
plt.subplot(2,2,2)
sns.kdeplot(hd["slope"])

plt.subplot(2,2,3)
sns.kdeplot(hd["sex"])

plt.subplot(2,2,4)
sns.kdeplot(hd["pro_data"])
plt.show()
```

```

-----
KeyError                                                 Traceback (most recent call last)
File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3805, in Index.get_
_loc(self, key)
    3804     try:
-> 3805         return self._engine.get_loc(casted_key)
    3806     except KeyError as err:

```

File index.pyx:167, in pandas._libs.index.IndexEngine.get_loc()

File index.pyx:196, in pandas._libs.index.IndexEngine.get_loc()

File pandas_libs\\hashtable_class_helper.pxi:7081, in pandas._libs.hashtable.Py
ObjectHashTable.get_item()

File pandas_libs\\hashtable_class_helper.pxi:7089, in pandas._libs.hashtable.Py
ObjectHashTable.get_item()

KeyError: 'pro_data'

The above exception was the direct cause of the following exception:

```

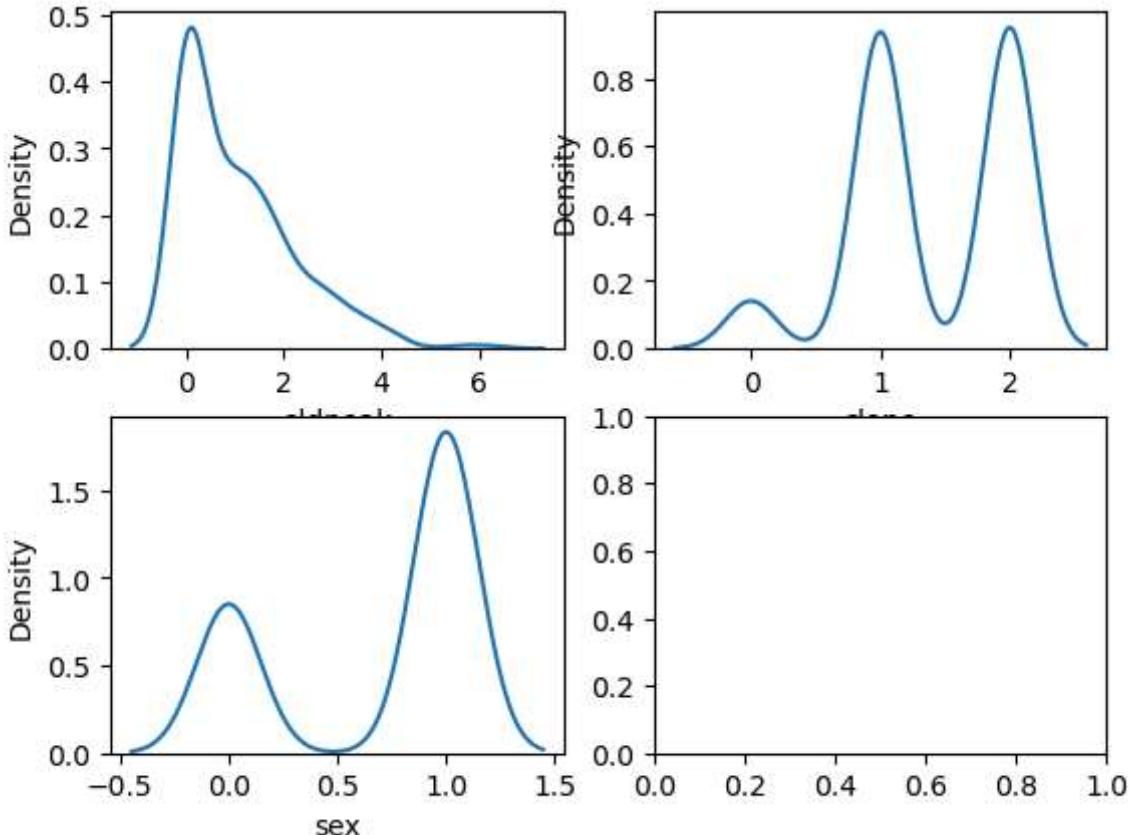
KeyError                                                 Traceback (most recent call last)
Cell In[118], line 11
    8 sns.kdeplot(hd["sex"])
    9 plt.subplot(2,2,4)
-> 11 sns.kdeplot(hd["pro_data"])
   12 plt.show()

```

File ~\anaconda3\Lib\site-packages\pandas\core\frame.py:4102, in DataFrame.__geti
tem__(self, key)
 4100 if self.columns.nlevels > 1:
 4101 return self._getitem_multilevel(key)
-> 4102 indexer = self.columns.get_loc(key)
 4103 if is_integer(indexer):
 4104 indexer = [indexer]

File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3812, in Index.get_
_loc(self, key)
 3807 if isinstance(casted_key, slice) or (
 3808 isinstance(casted_key, abc.Iterable)
 3809 and any(isinstance(x, slice) for x in casted_key)
 3810):
 3811 raise InvalidIndexError(key)
-> 3812 raise KeyError(key) from err
 3813 except TypeError:
 3814 # If we have a listlike key, _check_indexing_error will raise
 3815 # InvalidIndexError. Otherwise we fall through and re-raise
 3816 # the TypeError.
 3817 self._check_indexing_error(key)

KeyError: 'pro_data'



Violin Chart

```
In [ ]: plt.figure(figsize=(10,6))
sns.violinplot (x=hd["target"],y=hd["sex"],data=hd)
```

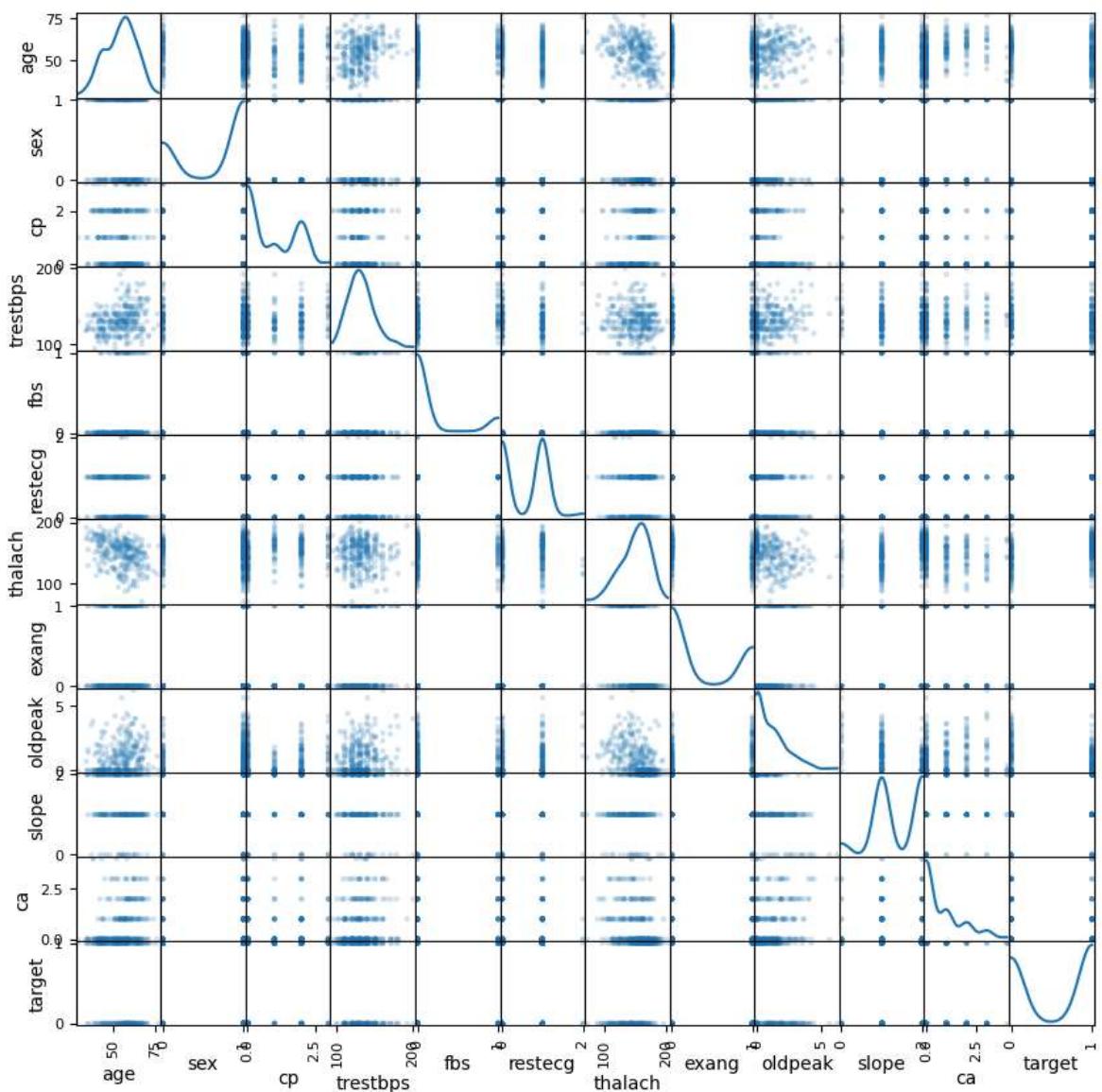
```
In [ ]: hd.drop('pro_data',axis=1,inplace=True)
```

```
In [185...]: from pandas.plotting import scatter_matrix
scatter_matrix(hd,alpha=0.2,figsize=(10,10),diagonal='kde')
```

```
Out[185... array([[<Axes: xlabel='age', ylabel='age'>,
   <Axes: xlabel='sex', ylabel='age'>,
   <Axes: xlabel='cp', ylabel='age'>,
   <Axes: xlabel='trestbps', ylabel='age'>,
   <Axes: xlabel='fbs', ylabel='age'>,
   <Axes: xlabel='restecg', ylabel='age'>,
   <Axes: xlabel='thalach', ylabel='age'>,
   <Axes: xlabel='exang', ylabel='age'>,
   <Axes: xlabel='oldpeak', ylabel='age'>,
   <Axes: xlabel='slope', ylabel='age'>,
   <Axes: xlabel='ca', ylabel='age'>,
   <Axes: xlabel='target', ylabel='age'>],
[<Axes: xlabel='age', ylabel='sex'>,
   <Axes: xlabel='sex', ylabel='sex'>,
   <Axes: xlabel='cp', ylabel='sex'>,
   <Axes: xlabel='trestbps', ylabel='sex'>,
   <Axes: xlabel='fbs', ylabel='sex'>,
   <Axes: xlabel='restecg', ylabel='sex'>,
   <Axes: xlabel='thalach', ylabel='sex'>,
   <Axes: xlabel='exang', ylabel='sex'>,
   <Axes: xlabel='oldpeak', ylabel='sex'>,
   <Axes: xlabel='slope', ylabel='sex'>,
   <Axes: xlabel='ca', ylabel='sex'>,
   <Axes: xlabel='target', ylabel='sex'>],
[<Axes: xlabel='age', ylabel='cp'>,
   <Axes: xlabel='sex', ylabel='cp'>,
   <Axes: xlabel='cp', ylabel='cp'>,
   <Axes: xlabel='trestbps', ylabel='cp'>,
   <Axes: xlabel='fbs', ylabel='cp'>,
   <Axes: xlabel='restecg', ylabel='cp'>,
   <Axes: xlabel='thalach', ylabel='cp'>,
   <Axes: xlabel='exang', ylabel='cp'>,
   <Axes: xlabel='oldpeak', ylabel='cp'>,
   <Axes: xlabel='slope', ylabel='cp'>,
   <Axes: xlabel='ca', ylabel='cp'>,
   <Axes: xlabel='target', ylabel='cp'>],
[<Axes: xlabel='age', ylabel='trestbps'>,
   <Axes: xlabel='sex', ylabel='trestbps'>,
   <Axes: xlabel='cp', ylabel='trestbps'>,
   <Axes: xlabel='trestbps', ylabel='trestbps'>,
   <Axes: xlabel='fbs', ylabel='trestbps'>,
   <Axes: xlabel='restecg', ylabel='trestbps'>,
   <Axes: xlabel='thalach', ylabel='trestbps'>,
   <Axes: xlabel='exang', ylabel='trestbps'>,
   <Axes: xlabel='oldpeak', ylabel='trestbps'>,
   <Axes: xlabel='slope', ylabel='trestbps'>,
   <Axes: xlabel='ca', ylabel='trestbps'>,
   <Axes: xlabel='target', ylabel='trestbps'>],
[<Axes: xlabel='age', ylabel='fbs'>,
   <Axes: xlabel='sex', ylabel='fbs'>,
   <Axes: xlabel='cp', ylabel='fbs'>,
   <Axes: xlabel='trestbps', ylabel='fbs'>,
   <Axes: xlabel='fbs', ylabel='fbs'>,
   <Axes: xlabel='restecg', ylabel='fbs'>,
   <Axes: xlabel='thalach', ylabel='fbs'>,
   <Axes: xlabel='exang', ylabel='fbs'>,
   <Axes: xlabel='oldpeak', ylabel='fbs'>,
   <Axes: xlabel='slope', ylabel='fbs'>,
   <Axes: xlabel='ca', ylabel='fbs'>,
   <Axes: xlabel='target', ylabel='fbs'>],
```

```
[<Axes: xlabel='age', ylabel='restecg'>,
 <Axes: xlabel='sex', ylabel='restecg'>,
 <Axes: xlabel='cp', ylabel='restecg'>,
 <Axes: xlabel='trestbps', ylabel='restecg'>,
 <Axes: xlabel='fbs', ylabel='restecg'>,
 <Axes: xlabel='restecg', ylabel='restecg'>,
 <Axes: xlabel='thalach', ylabel='restecg'>,
 <Axes: xlabel='exang', ylabel='restecg'>,
 <Axes: xlabel='oldpeak', ylabel='restecg'>,
 <Axes: xlabel='slope', ylabel='restecg'>,
 <Axes: xlabel='ca', ylabel='restecg'>,
 <Axes: xlabel='target', ylabel='restecg'>],
 [<Axes: xlabel='age', ylabel='thalach'>,
 <Axes: xlabel='sex', ylabel='thalach'>,
 <Axes: xlabel='cp', ylabel='thalach'>,
 <Axes: xlabel='trestbps', ylabel='thalach'>,
 <Axes: xlabel='fbs', ylabel='thalach'>,
 <Axes: xlabel='restecg', ylabel='thalach'>,
 <Axes: xlabel='thalach', ylabel='thalach'>,
 <Axes: xlabel='exang', ylabel='thalach'>,
 <Axes: xlabel='oldpeak', ylabel='thalach'>,
 <Axes: xlabel='slope', ylabel='thalach'>,
 <Axes: xlabel='ca', ylabel='thalach'>,
 <Axes: xlabel='target', ylabel='thalach'>],
 [<Axes: xlabel='age', ylabel='exang'>,
 <Axes: xlabel='sex', ylabel='exang'>,
 <Axes: xlabel='cp', ylabel='exang'>,
 <Axes: xlabel='trestbps', ylabel='exang'>,
 <Axes: xlabel='fbs', ylabel='exang'>,
 <Axes: xlabel='restecg', ylabel='exang'>,
 <Axes: xlabel='thalach', ylabel='exang'>,
 <Axes: xlabel='exang', ylabel='exang'>,
 <Axes: xlabel='oldpeak', ylabel='exang'>,
 <Axes: xlabel='slope', ylabel='exang'>,
 <Axes: xlabel='ca', ylabel='exang'>,
 <Axes: xlabel='target', ylabel='exang'>],
 [<Axes: xlabel='age', ylabel='oldpeak'>,
 <Axes: xlabel='sex', ylabel='oldpeak'>,
 <Axes: xlabel='cp', ylabel='oldpeak'>,
 <Axes: xlabel='trestbps', ylabel='oldpeak'>,
 <Axes: xlabel='fbs', ylabel='oldpeak'>,
 <Axes: xlabel='restecg', ylabel='oldpeak'>,
 <Axes: xlabel='thalach', ylabel='oldpeak'>,
 <Axes: xlabel='exang', ylabel='oldpeak'>,
 <Axes: xlabel='oldpeak', ylabel='oldpeak'>,
 <Axes: xlabel='slope', ylabel='oldpeak'>,
 <Axes: xlabel='ca', ylabel='oldpeak'>,
 <Axes: xlabel='target', ylabel='oldpeak'>],
 [<Axes: xlabel='age', ylabel='slope'>,
 <Axes: xlabel='sex', ylabel='slope'>,
 <Axes: xlabel='cp', ylabel='slope'>,
 <Axes: xlabel='trestbps', ylabel='slope'>,
 <Axes: xlabel='fbs', ylabel='slope'>,
 <Axes: xlabel='restecg', ylabel='slope'>,
 <Axes: xlabel='thalach', ylabel='slope'>,
 <Axes: xlabel='exang', ylabel='slope'>,
 <Axes: xlabel='oldpeak', ylabel='slope'>,
 <Axes: xlabel='slope', ylabel='slope'>,
 <Axes: xlabel='ca', ylabel='slope'>,
 <Axes: xlabel='target', ylabel='slope'>],
```

```
[<Axes: xlabel='age', ylabel='ca'>,
 <Axes: xlabel='sex', ylabel='ca'>,
 <Axes: xlabel='cp', ylabel='ca'>,
 <Axes: xlabel='trestbps', ylabel='ca'>,
 <Axes: xlabel='fbs', ylabel='ca'>,
 <Axes: xlabel='restecg', ylabel='ca'>,
 <Axes: xlabel='thalach', ylabel='ca'>,
 <Axes: xlabel='exang', ylabel='ca'>,
 <Axes: xlabel='oldpeak', ylabel='ca'>,
 <Axes: xlabel='slope', ylabel='ca'>,
 <Axes: xlabel='ca', ylabel='ca'>,
 <Axes: xlabel='target', ylabel='ca'>],
 [<Axes: xlabel='age', ylabel='target'>,
 <Axes: xlabel='sex', ylabel='target'>,
 <Axes: xlabel='cp', ylabel='target'>,
 <Axes: xlabel='trestbps', ylabel='target'>,
 <Axes: xlabel='fbs', ylabel='target'>,
 <Axes: xlabel='restecg', ylabel='target'>,
 <Axes: xlabel='thalach', ylabel='target'>,
 <Axes: xlabel='exang', ylabel='target'>,
 <Axes: xlabel='oldpeak', ylabel='target'>,
 <Axes: xlabel='slope', ylabel='target'>,
 <Axes: xlabel='ca', ylabel='target'>,
 <Axes: xlabel='target', ylabel='target'>]], dtype=object)
```



In []: