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

In [ ]:
```
df = pd.read_csv('heart.csv')
```

In [3]:
```
df
```

Out[3]:

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

303 rows × 14 columns

In [4]:
```
df.dtypes
```

Out[4]:
```
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 [28]:
```
# df['sex']=df['sex'].astype('object')
df['sex']=df['sex'].astype('int64')
```

In [9]:
```
df.dtypes
```

Out[9]:
```
age           int64
sex          object
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 [25]:
```
df.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    object 
 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(12), object(1)
memory usage: 33.3+ KB
```

In [27]:
```
df.describe()
```

Out[27]:

	age	cp	trestbps	chol	fbs	restecg	thalach	
count	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303
mean	54.366337	0.966997	131.623762	246.264026	0.148515	0.528053	149.646865	0
std	9.082101	1.032052	17.538143	51.830751	0.356198	0.525860	22.905161	0
min	29.000000	0.000000	94.000000	126.000000	0.000000	0.000000	71.000000	0
25%	47.500000	0.000000	120.000000	211.000000	0.000000	0.000000	133.500000	0
50%	55.000000	1.000000	130.000000	240.000000	0.000000	1.000000	153.000000	0
75%	61.000000	2.000000	140.000000	274.500000	0.000000	1.000000	166.000000	1
max	77.000000	3.000000	200.000000	564.000000	1.000000	2.000000	202.000000	1

In [30]:
```
df['sex'] = df.sex.replace({1:'Male',0:'Female'})
```

In [32]:
```
df['target'] = df.target.replace({1:'Disease',0:'No_disease'})
```

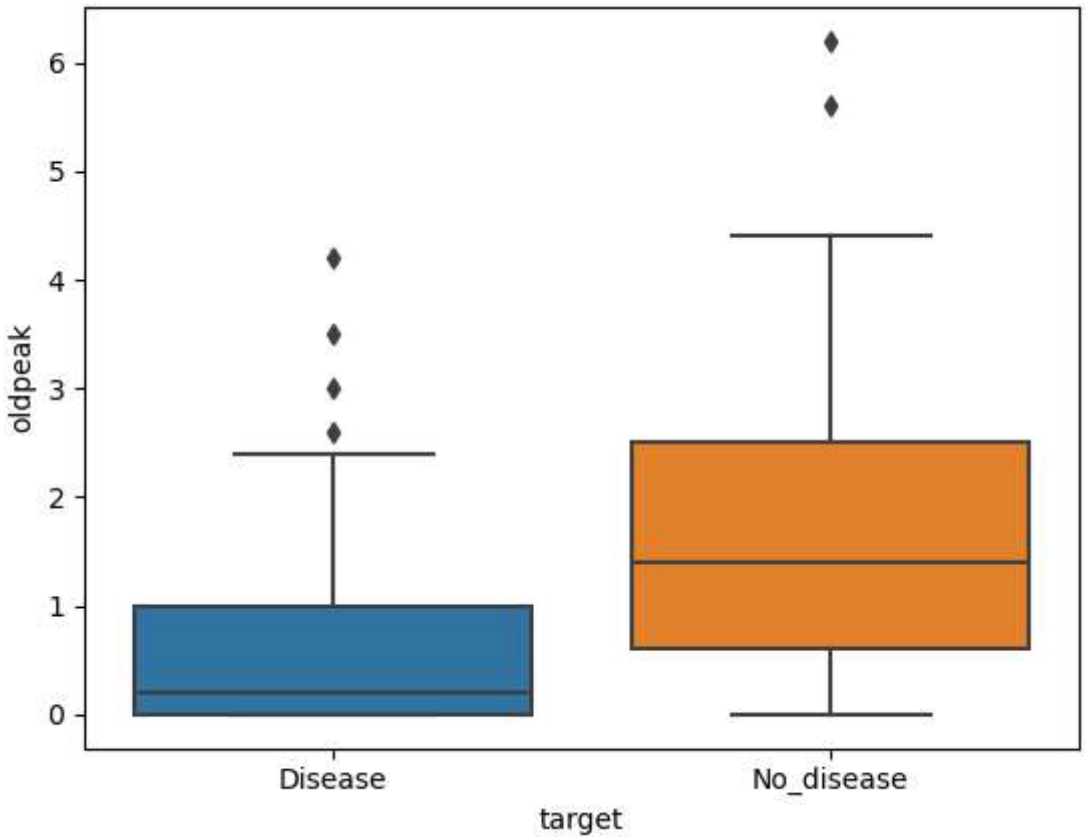
In [1]:
```
sns.heatmap(df.corr(),annot=True,linewidths=2)
plt.tight_layout()
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
Cell In[1], line 1
----> 1 sns.heatmap(df.corr(),annot=True,linewidths=2)
      2 plt.tight_layout()

NameError: name 'sns' is not defined
```

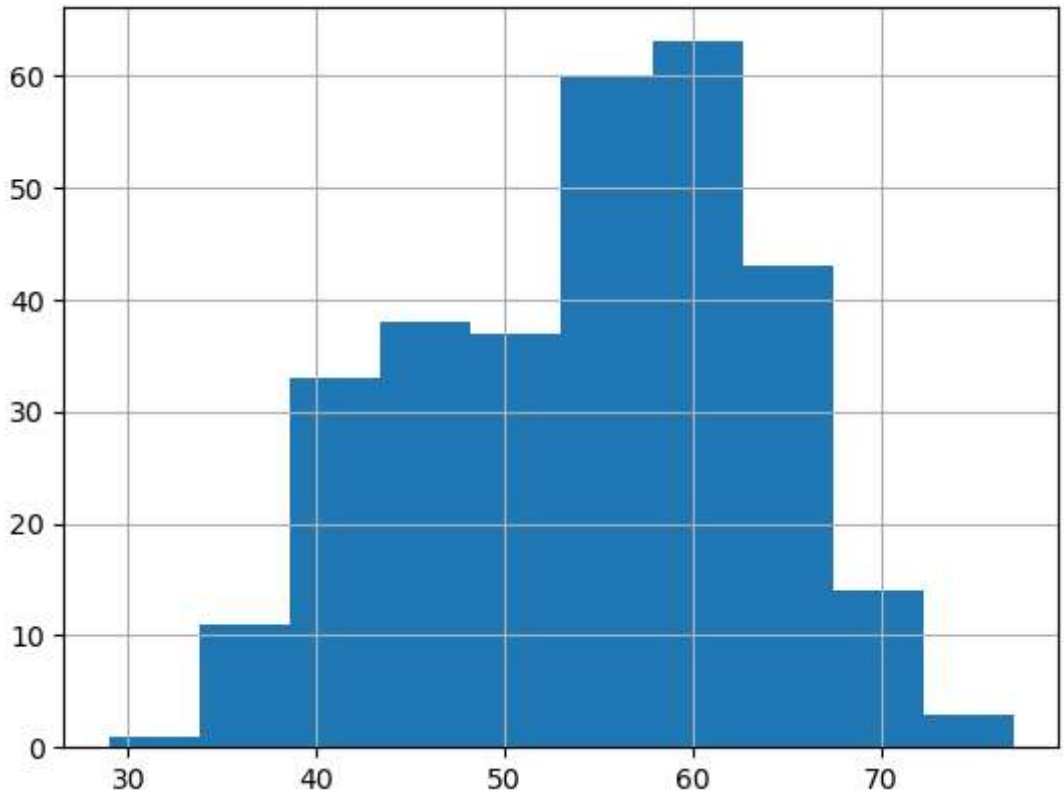
In [34]:
```
sns.boxplot(x='target',y='oldpeak',data=df)
```

Out[34]: <AxesSubplot:xlabel='target', ylabel='oldpeak'>



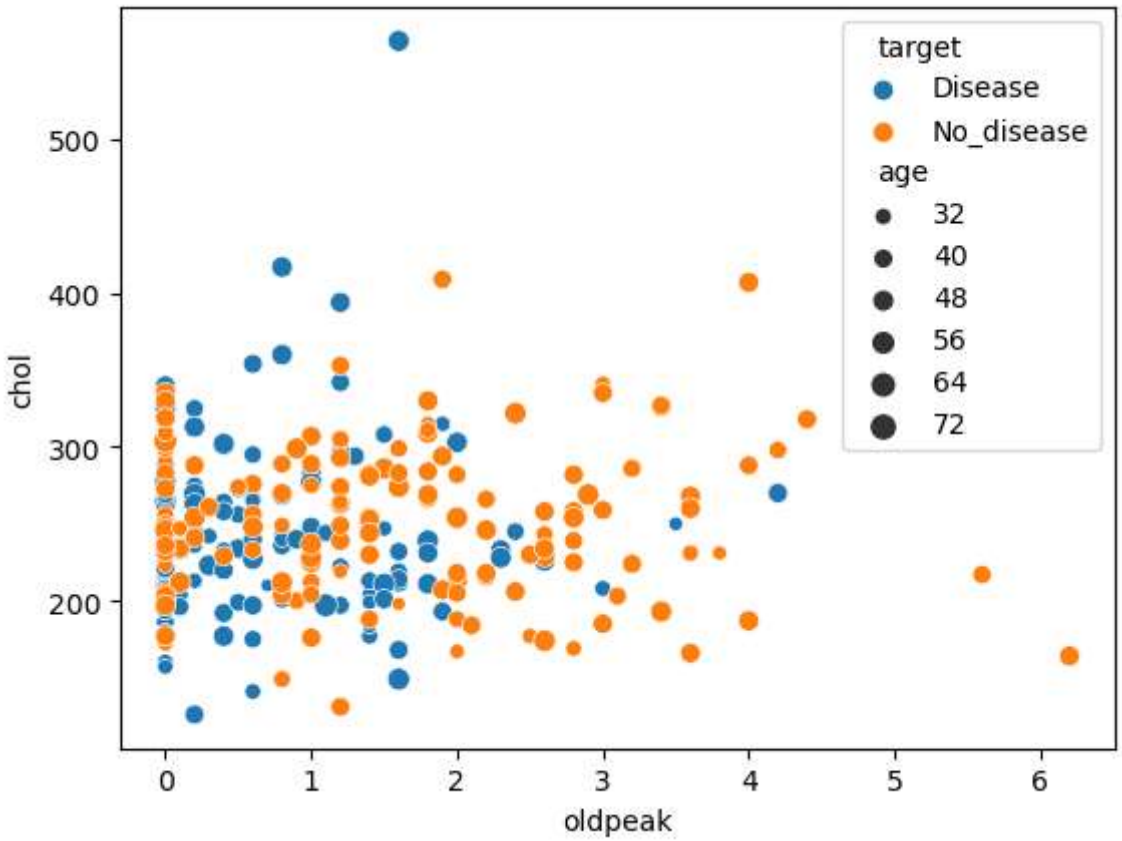
In [35]:
```
df['age'].hist().plot(kind='bar')
```

Out[35]: []



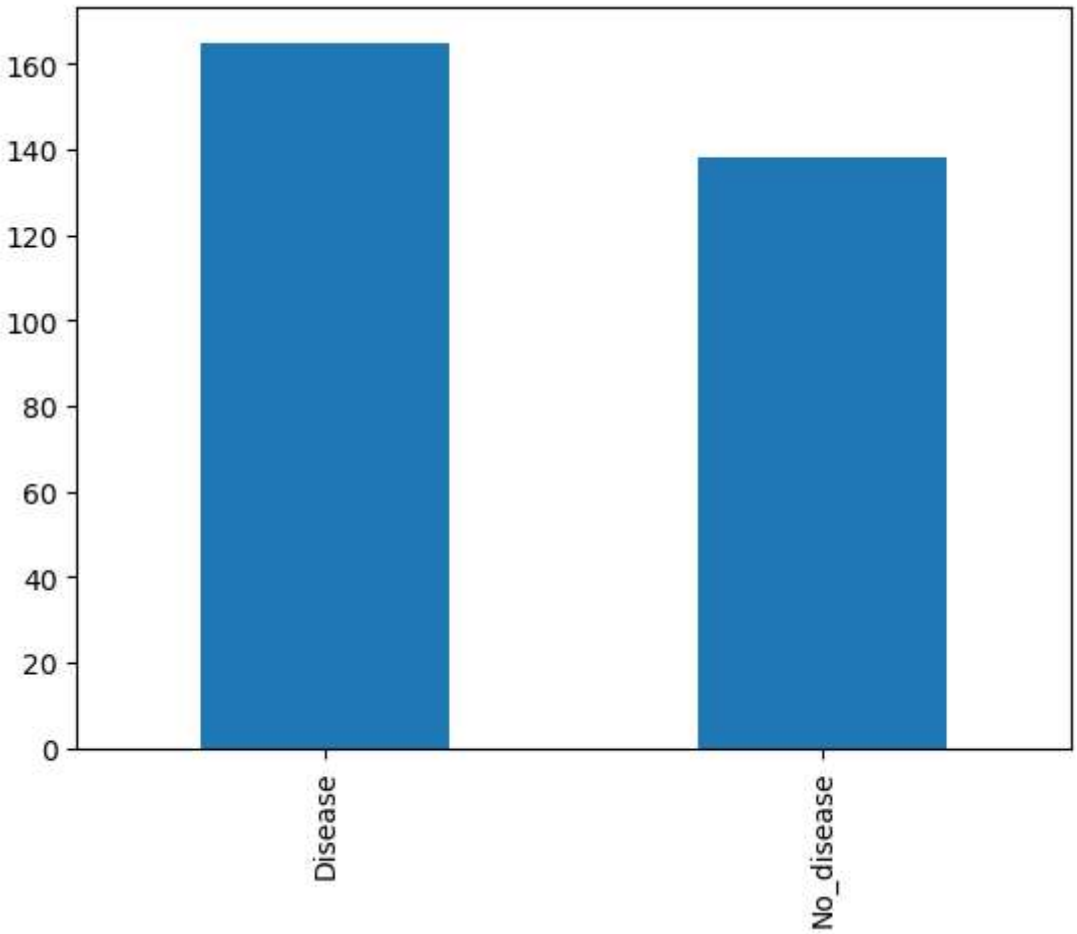
In [37]:
```
sns.scatterplot(x=df['oldpeak'],y=df['chol'],hue=df['target'],size=df['age'])
```

Out[37]:
```
<AxesSubplot:xlabel='oldpeak', ylabel='chol'>
```



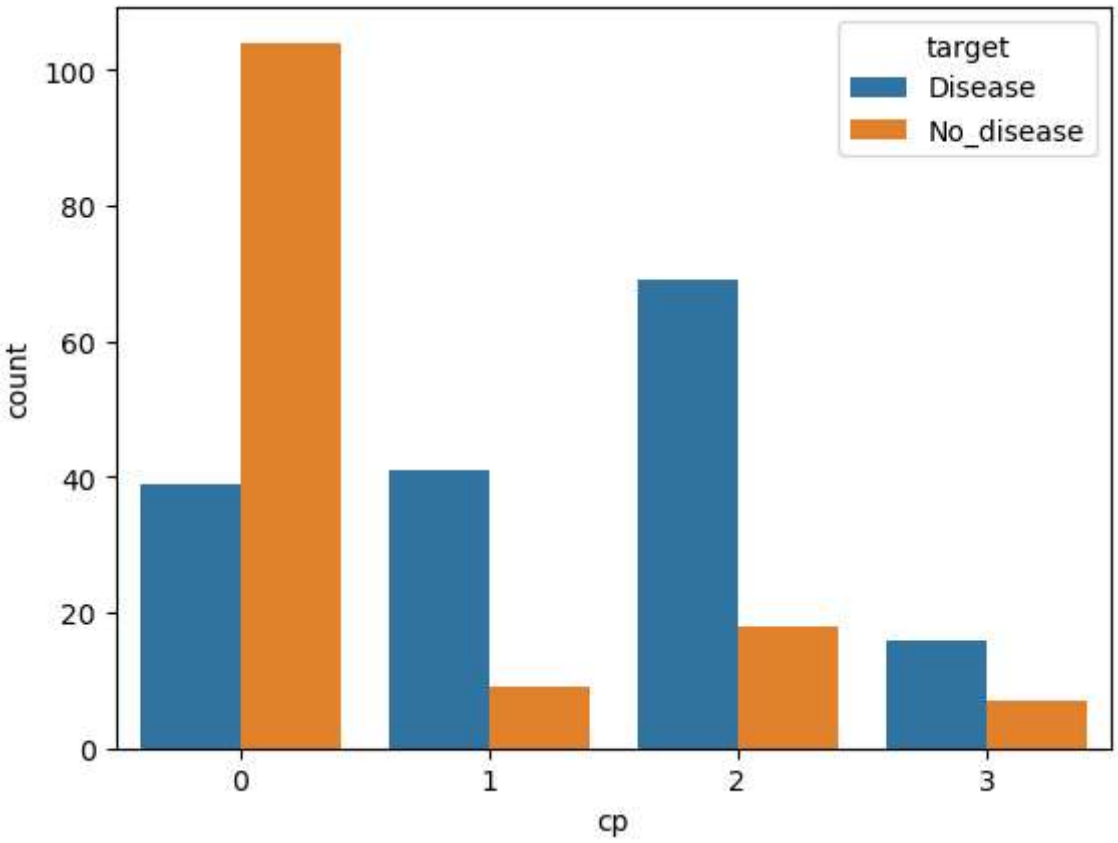
In [52]:
```
df.target.value_counts().plot(kind='bar')
```

Out[52]: <AxesSubplot:>



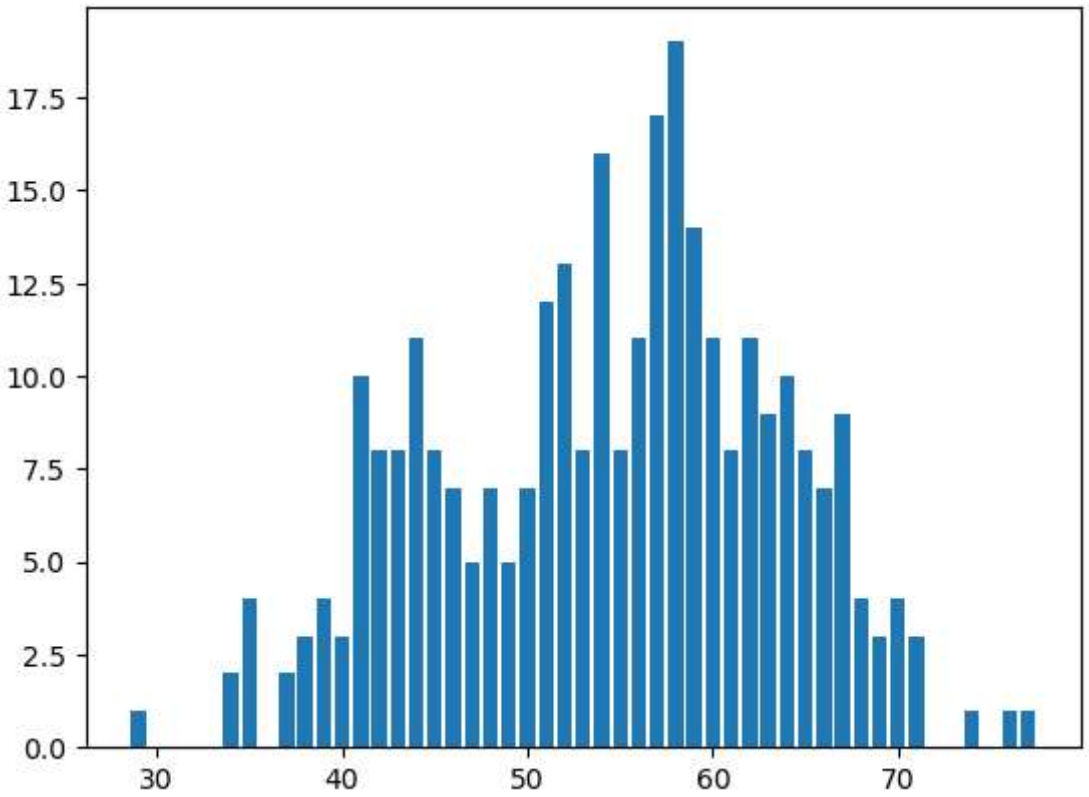
In [57]:
```
sns.countplot(x='cp',hue='target',data=df)
```

Out[57]:
```
<AxesSubplot:xlabel='cp', ylabel='count'>
```



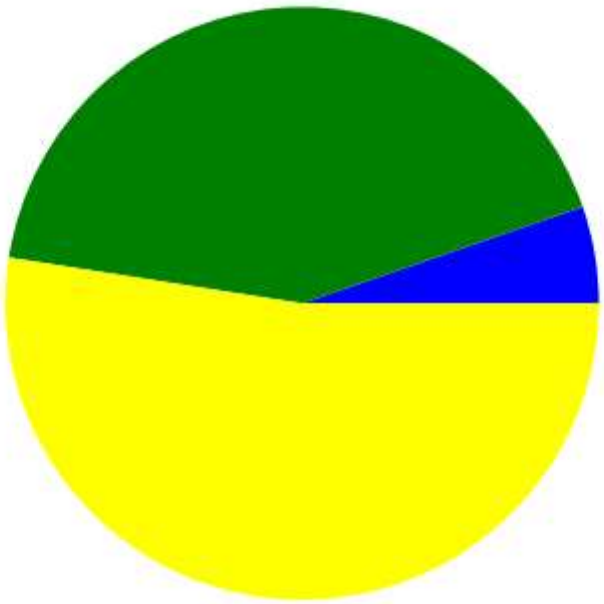
In [67]:
```
plt.bar(df['age'].value_counts().index,df['age'].value_counts())
```

Out[67]:
```
<BarContainer object of 41 artists>
```



In [79]:
```
young = df[(df['age'] >=29) & (df['age'] < 40)]
middle = df[(df['age'] >=40) & (df['age'] < 55)]
old = df[(df['age'] >=55)]
```

In [83]:
```
colors = ['blue','green','yellow']
plt.pie([len(young),len(middle),len(old)],colors=colors);
```



In [89]:
```
df['age'].valu
```

```
---------------------------------------------------------------------------
AttributeError                            Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_5032\2875843194.py in <module>
----> 1 df['age'].valuse_count()

C:\ProgramData\Anaconda3\lib\site-packages\pandas\core\generic.py in __getatt
r__(self, name)
   5573         ):
   5574             return self[name]
-> 5575         return object.__getattribute__(self, name)
   5576
   5577     def __setattr__(self, name: str, value) -> None:

AttributeError: 'Series' object has no attribute 'valuse_count'
```

In [ ]:
```

```