```
import pandas as pd
import matplotlib.pyplot as plt
from functools import reduce


df = pd.read_csv('upload.csv')


df
```

	Id	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope
0	1	52	1	0	125	212	0	1	168	0	1.0	2
1	2	53	1	0	140	203	1	0	155	1	3.1	0
2	3	70	1	0	145	174	0	1	125	1	2.6	0
3	4	61	1	0	148	203	0	1	161	0	0.0	2
4	5	62	0	0	138	294	1	1	106	0	1.9	1
...	...	...	...	...	...	...	...	...	...	...	...	...
1020	1021	59	1	1	140	221	0	1	164	1	0.0	2
1021	1022	60	1	0	125	258	0	0	141	1	2.8	1
1022	1023	47	1	0	110	275	0	0	118	1	1.0	1
1023	1024	50	0	0	110	254	0	0	159	0	0.0	2
1024	1025	54	1	0	120	188	0	1	113	0	1.4	1

1025 rows × 15 columns

```
df.dtypes
```

```
Id            int64
age           int64
sex           int64
cp            int64
trestbps      int64
chol          int64
fbs           int64
restecg       int64
thalach       int64
exang         int64
oldpeak     float64
slope         int64
ca            int64
```

```
array([0, 1])
```

0=Nothing to note 1=ST-T Wave abnormality 2= Possible or definite left ventricular hypertrophy

```
df['restecg'].unique()
```

```
array([1, 0, 2])
```

```
df['thalach'].unique()
```

```
array([168, 155, 125, 161, 106, 122, 140, 145, 144, 116, 136, 192, 156,
       142, 109, 162, 165, 148, 172, 173, 146, 179, 152, 117, 115, 112,
       163, 147, 182, 105, 150, 151, 169, 166, 178, 132, 160, 123, 139,
       111, 180, 164, 202, 157, 159, 170, 138, 175, 158, 126, 143, 141,
       167,  95, 190, 118, 103, 181, 108, 177, 134, 120, 171, 149, 154,
       153,  88, 174, 114, 195, 133,  96, 124, 131, 185, 194, 128, 127,
       186, 184, 188, 130,  71, 137,  99, 121, 187,  97,  90, 129, 113])
```

0= no 1= yes

```
df['exang'].unique()
```

```
array([0, 1])
```

```
df['oldpeak'].unique()
```

```
array([1. , 3.1, 2.6, 0. , 1.9, 4.4, 0.8, 3.2, 1.6, 3. , 0.7, 4.2, 1.5,
       2.2, 1.1, 0.3, 0.4, 0.6, 3.4, 2.8, 1.2, 2.9, 3.6, 1.4, 0.2, 2. ,
       5.6, 0.9, 1.8, 6.2, 4. , 2.5, 0.5, 0.1, 2.1, 2.4, 3.8, 2.3, 1.3,
       3.5])
```

```
df['slope'].unique()
```

```
array([2, 0, 1])
```

```
df.isnull().sum()
```

```
Id          0
age         0
sex         0
cp          0
trestbps    0
chol        0
fbs         0
restecg     0
thalach     0
exang       0
```

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope
0	52	1	0	125	212	0	1	168	0	1.0	2
1	53	1	0	140	203	1	0	155	1	3.1	0
2	70	1	0	145	174	0	1	125	1	2.6	0
3	61	1	0	148	203	0	1	161	0	0.0	2

```
y
```

```
0       0
1       0
2       0
3       0
4       0
       ..
1020    1
1021    0
1022    0
1023    1
1024    0
Name: target, Length: 1025, dtype: int64
```

```
import numpy as np
import pandas as pd
from sklearn.tree import DecisionTreeClassifier
```

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=9)
print(X_train.shape)
print(X_test.shape)
```

```
(717, 11)
(308, 11)
```

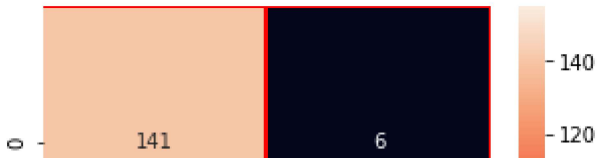
```
from sklearn.tree import DecisionTreeClassifier
clf = DecisionTreeClassifier(random_state=0)
clf.fit(X_train, y_train)
```

```
DecisionTreeClassifier(random_state=0)
```

```
clf.score(X_train,y_train)
```

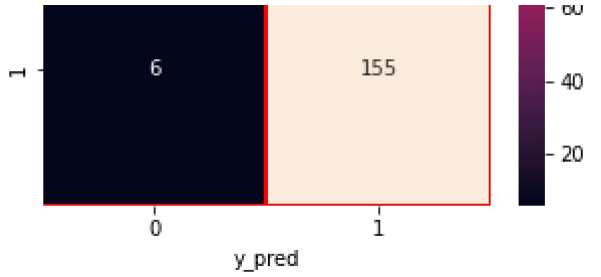
```
1.0
```

```
from sklearn.metrics import accuracy_score
y_pred = clf.predict(X_test )
accuracy_score(y_pred,y_test)
```




```
import pickle

pickle.dump(clf,open('heartd.pkl','wb'))
heart = pickle.load(open('heartd.pkl','rb'))
```



Colab paid products - Cancel contracts here

✓ 0s completed at 21:14