In [1]:

```
import os
os.getcwd()
```

Out[1]:

```
'/Users/shivraj/Desktop/ML LABS'
```

In [2]:

```
import pandas as pd
```

In [3]:

```
#import the dataset
df = pd.read_csv('Heart 2.csv')
```

In [6]:

```
df.head(5)
```

Out[6]:

| | Unnamed: 0 | Age | Sex | ChestPain | RestBP | Chol | Fbs | RestECG | MaxHR | ExAng | Oldpeak |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 63 | 1 | typical | 145 | 233 | 1 | 2 | 150 | 0 | 2.3 |
| **1** | 2 | 67 | 1 | asymptomatic | 160 | 286 | 0 | 2 | 108 | 1 | 1.5 |
| **2** | 3 | 67 | 1 | asymptomatic | 120 | 229 | 0 | 2 | 129 | 1 | 2.6 |
| **3** | 4 | 37 | 1 | nonanginal | 130 | 250 | 0 | 0 | 187 | 0 | 3.5 |
| **4** | 5 | 41 | 0 | nontypical | 130 | 204 | 0 | 2 | 172 | 0 | 1.4 |

In [7]:

```
#shape of the Data
df.shape
```

Out[7]:

```
(303, 15)
```

In [9]:

```
df.isnull()
```

Out[9]:

| | Unnamed: 0 | Age | Sex | ChestPain | RestBP | Chol | Fbs | RestECG | MaxHR | ExAng | Oldpe |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | False | False | False | False | False | False | False | False | False | False | Fal |
| **1** | False | False | False | False | False | False | False | False | False | False | Fal |
| **2** | False | False | False | False | False | False | False | False | False | False | Fal |
| **3** | False | False | False | False | False | False | False | False | False | False | Fal |
| **4** | False | False | False | False | False | False | False | False | False | False | Fal |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| **298** | False | False | False | False | False | False | False | False | False | False | Fal |
| **299** | False | False | False | False | False | False | False | False | False | False | Fal |
| **300** | False | False | False | False | False | False | False | False | False | False | Fal |
| **301** | False | False | False | False | False | False | False | False | False | False | Fal |
| **302** | False | False | False | False | False | False | False | False | False | False | Fal |

303 rows × 15 columns

In [10]:

```
#find Null Values
df.isnull().sum()
```

Out[10]:

```
Unnamed: 0    0
Age           0
Sex           0
ChestPain     0
RestBP        0
Chol          0
Fbs           0
RestECG       0
MaxHR         0
ExAng         0
Oldpeak       0
Slope         0
Ca            4
Thal          2
AHD           0
dtype: int64
```

In [11]:

```
df.count()
```

Out[11]:

```
Unnamed: 0    303
Age           303
Sex           303
ChestPain     303
RestBP        303
Chol          303
Fbs           303
RestECG       303
MaxHR         303
ExAng         303
Oldpeak       303
Slope         303
Ca            299
Thal          301
AHD           303
dtype: int64
```

In [15]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 303 entries, 0 to 302
Data columns (total 15 columns):
 #   Column      Non-Null Count  Dtype  
---  ------      --------------  -----  
 0   Unnamed: 0  303 non-null    int64  
 1   Age         303 non-null    int64  
 2   Sex         303 non-null    int64  
 3   ChestPain   303 non-null    object 
 4   RestBP      303 non-null    int64  
 5   Chol        303 non-null    int64  
 6   Fbs         303 non-null    int64  
 7   RestECG     303 non-null    int64  
 8   MaxHR       303 non-null    int64  
 9   ExAng       303 non-null    int64  
 10  Oldpeak     303 non-null    float64
 11  Slope       303 non-null    int64  
 12  Ca          299 non-null    float64
 13  Thal        301 non-null    object 
 14  AHD         303 non-null    object 
dtypes: float64(2), int64(10), object(3)
memory usage: 35.6+ KB
```

In [16]:

```
#find data types of Each columns
df.dtypes
```

Out[16]:

```
Unnamed: 0      int64
Age             int64
Sex             int64
ChestPain      object
RestBP          int64
Chol            int64
Fbs             int64
RestECG         int64
MaxHR           int64
ExAng           int64
Oldpeak       float64
Slope           int64
Ca            float64
Thal           object
AHD            object
dtype: object
```

In [17]:

```
#Find out Zero's
df == 0
```

Out[17]:

| | Unnamed: 0 | Age | Sex | ChestPain | RestBP | Chol | Fbs | RestECG | MaxHR | ExAng | Oldpe |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | False | False | False | False | False | False | False | False | False | True | Fal |
| **1** | False | False | False | False | False | False | True | False | False | False | Fal |
| **2** | False | False | False | False | False | False | True | False | False | False | Fal |
| **3** | False | False | False | False | False | False | True | True | False | True | Fal |
| **4** | False | False | True | False | False | False | True | False | False | True | Fal |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| **298** | False | False | False | False | False | False | True | True | False | True | Fal |
| **299** | False | False | False | False | False | False | False | True | False | True | Fal |
| **300** | False | False | False | False | False | False | True | True | False | False | Fal |
| **301** | False | False | True | False | False | False | True | False | False | True | Tr |
| **302** | False | False | False | False | False | False | True | True | False | True | Tr |

303 rows × 15 columns

In [18]:

```
df[df == 0]
```

Out[18]:

| | Unnamed: 0 | Age | Sex | ChestPain | RestBP | Chol | Fbs | RestECG | MaxHR | ExAng | Oldpeak |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | 0.0 | NaN |
| **1** | NaN | NaN | NaN | NaN | NaN | NaN | 0.0 | NaN | NaN | NaN | NaN |
| **2** | NaN | NaN | NaN | NaN | NaN | NaN | 0.0 | NaN | NaN | NaN | NaN |
| **3** | NaN | NaN | NaN | NaN | NaN | NaN | 0.0 | 0.0 | NaN | 0.0 | NaN |
| **4** | NaN | NaN | 0.0 | NaN | NaN | NaN | 0.0 | NaN | NaN | 0.0 | NaN |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| **298** | NaN | NaN | NaN | NaN | NaN | NaN | 0.0 | 0.0 | NaN | 0.0 | NaN |
| **299** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | 0.0 | NaN | 0.0 | NaN |
| **300** | NaN | NaN | NaN | NaN | NaN | NaN | 0.0 | 0.0 | NaN | NaN | NaN |
| **301** | NaN | NaN | 0.0 | NaN | NaN | NaN | 0.0 | NaN | NaN | 0.0 | 0.0 |
| **302** | NaN | NaN | NaN | NaN | NaN | NaN | 0.0 | 0.0 | NaN | 0.0 | 0.0 |

303 rows × 15 columns

In [20]:

```
df[df == 0].count()
```

Out[20]:

```
Unnamed: 0      0
Age             0
Sex            97
ChestPain       0
RestBP          0
Chol            0
Fbs           258
RestECG       151
MaxHR           0
ExAng         204
Oldpeak        99
Slope           0
Ca            176
Thal            0
AHD             0
dtype: int64
```

In [22]:

```
df.columns
```

Out[22]:

```
Index(['Unnamed: 0', 'Age', 'Sex', 'ChestPain', 'RestBP', 'Chol', 'Fb
s',
       'RestECG', 'MaxHR', 'ExAng', 'Oldpeak', 'Slope', 'Ca', 'Thal',
'AHD'],
      dtype='object')
```

In [23]:

```
#Find Mean of Age
df['Age'].mean()
```

Out[23]:

```
54.43894389438944
```

In [24]:

```
df['Age'].median()
```

Out[24]:

```
56.0
```

In [25]:

```
newdf = df[['Age','Sex','ChestPain','RestBP','Chol']]
```

In [26]:

```
newdf
```

Out[26]:

| | Age | Sex | ChestPain | RestBP | Chol |
|---|---|---|---|---|---|
| **0** | 63 | 1 | typical | 145 | 233 |
| **1** | 67 | 1 | asymptomatic | 160 | 286 |
| **2** | 67 | 1 | asymptomatic | 120 | 229 |
| **3** | 37 | 1 | nonanginal | 130 | 250 |
| **4** | 41 | 0 | nontypical | 130 | 204 |
| **...** | ... | ... | ... | ... | ... |
| **298** | 45 | 1 | typical | 110 | 264 |
| **299** | 68 | 1 | asymptomatic | 144 | 193 |
| **300** | 57 | 1 | asymptomatic | 130 | 131 |
| **301** | 57 | 0 | nontypical | 130 | 236 |
| **302** | 38 | 1 | nonanginal | 138 | 175 |

303 rows × 5 columns

In [27]:

```
#cross Validation

from sklearn.model_selection import train_test_split

train , test = train_test_split(df,random_state=0,test_size=0.25)
```

In [29]:

```
train.shape
```

Out[29]:

```
(227, 15)
```

In [31]:

```
test.shape
```

Out[31]:

```
(76, 15)
```

In [32]:

```
import numpy as np
```

In [33]:

```
actual = list(np.ones(45)) + list(np.zeros(55))
```

In [34]:

```
np.array(actual)
```

Out[34]:

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

In [39]:

```
predicted = list(np.ones(40)) + list(np.zeros(52)) + list(np.ones(8))
```

In [40]:

```
np.array(predicted)
```

Out[40]:

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

In [41]:

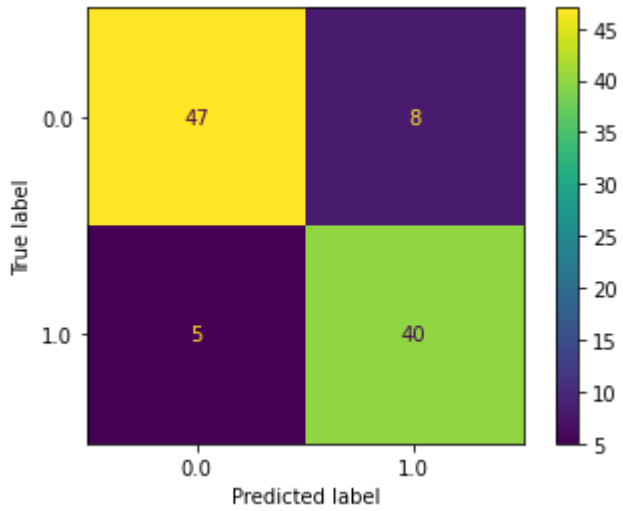
```
from sklearn.metrics import ConfusionMatrixDisplay
```

In [42]:

```
ConfusionMatrixDisplay.from_predictions(actual,predicted)
```

Out[42]:

```
<sklearn.metrics._plot.confusion_matrix.ConfusionMatrixDisplay at 0x7f
7c03c91220>
```




In [46]:

```
from sklearn.metrics import classification_report,accuracy_score
```

In [47]:

```
print(classification_report(actual,predicted))
```

```
              precision    recall  f1-score   support

         0.0       0.90      0.85      0.88        55
         1.0       0.83      0.89      0.86        45

    accuracy                           0.87       100
   macro avg       0.87      0.87      0.87       100
weighted avg       0.87      0.87      0.87       100
```

In [48]:

```
accuracy_score(actual,predicted)
```

Out[48]:

```
0.87
```

In [ ]: