

```
In [63]: #import all packages
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import mean_squared_error, r2_score
import pandas as pd
from sklearn.model_selection import train_test_split
```

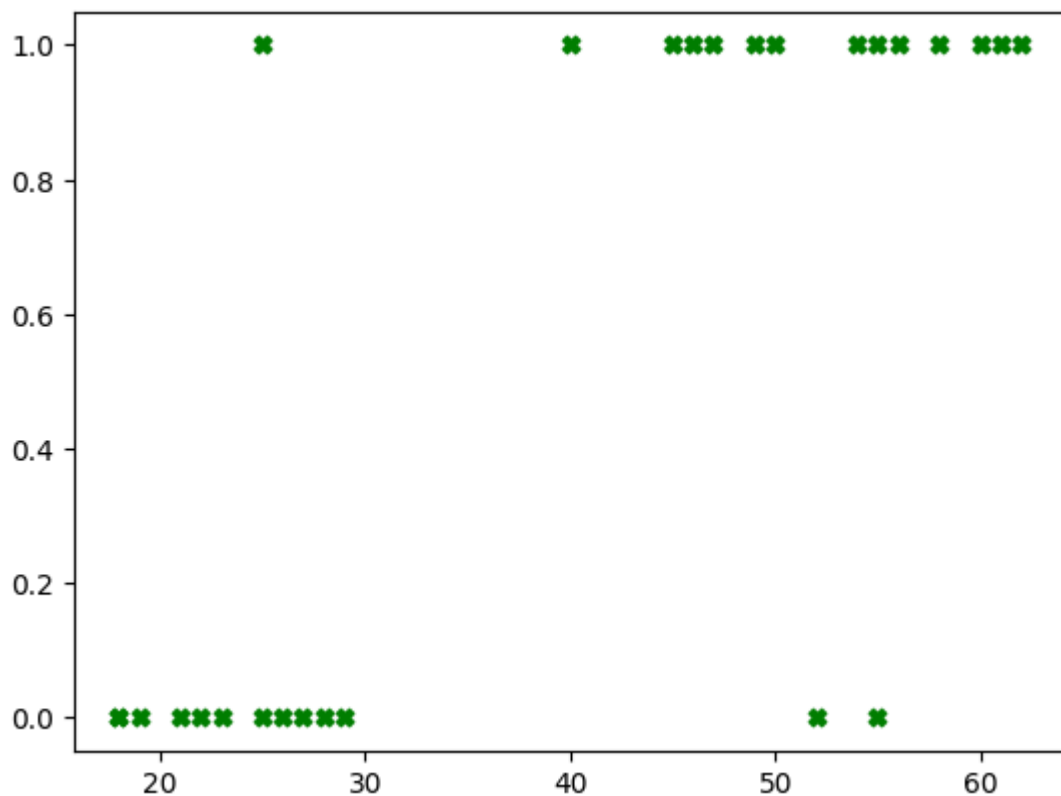
```
In [64]: #Load Data using pandas
data=pd.read_csv('insurance_data.csv')
data.head()
```

```
Out[64]:
```

	age	bought_insurance
0	22	0
1	25	0
2	47	1
3	52	0
4	46	1

```
In [65]: #plot the data
plt.scatter(data[['age']],data.bought_insurance,marker='X',color='green')
plt.show
```

```
Out[65]: <function matplotlib.pyplot.show(close=None, block=None)>
```



```
In [66]: #implement train_test_split
X_train,X_test,y_train,y_test=train_test_split(data[['age']],data.bought_insuran
```

```
In [67]: X_train
```

```
Out[67]:
```

	age
16	25
4	46
14	49
11	28
7	60
13	29
21	26
26	23
0	22
18	19
6	55
5	56
9	61
12	27
8	62
19	18
3	52
22	40
20	21
1	25
17	58
10	18
23	45
25	54

```
In [68]: y_train
```

```
Out[68]: 16    1
         4    1
         14   1
         11   0
         7    1
         13   0
         21   0
         26   0
         0    0
         18   0
         6    0
         5    1
         9    1
         12   0
         8    1
         19   0
         3    0
         22   1
         20   0
         1    0
         17   1
         10   0
         23   1
         25   1
        Name: bought_insurance, dtype: int64
```

```
In [69]: #apply Logistic Regression
        model = LogisticRegression()
        model.fit(X_train,y_train)
```

```
Out[69]: ▼ LogisticRegression ⓘ ?
        LogisticRegression()
```

```
In [70]: model.predict(X_test)
```

```
Out[70]: array([1, 1, 1])
```

```
In [71]: model.score(X_test,y_test)
```

```
Out[71]: 1.0
```

```
In [72]: model.predict_proba(X_test)
```

```
Out[72]: array([[0.30560252, 0.69439748],
                [0.23261178, 0.76738822],
                [0.14002771, 0.85997229]])
```