In [2]:
```
import pandas as pd
from matplotlib import pyplot as plt
%matplotlib inline
```

In [3]:
```
df = pd.read_csv("C:/Users/prasa/Desktop/py codes/ds projects/ML/7 Logi
stic Regression/insurance_data.csv")
df
```

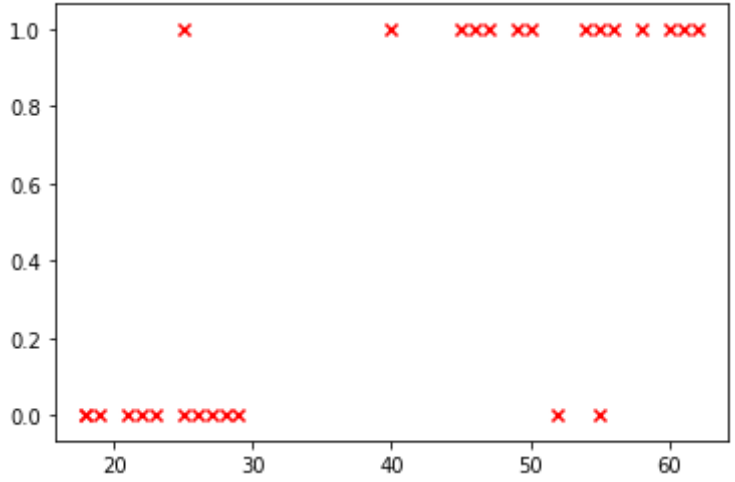
Out[3]:

	age	bought_insurance
0	22	0
1	25	0
2	47	1
3	52	0
4	46	1
5	56	1
6	55	0
7	60	1
8	62	1
9	61	1
10	18	0
11	28	0
12	27	0
13	29	0
14	49	1
15	55	1
16	25	1

	age	bought_insurance
17	58	1
18	19	0
19	18	0
20	21	0
21	26	0
22	40	1
23	45	1
24	50	1
25	54	1
26	23	0

In [4]:
```
plt.scatter(df.age,df.bought_insurance,marker='x',color='red')
```

Out[4]: <matplotlib.collections.PathCollection at 0x247025fc988>



In [5]:
```
df.shape
```

Out[5]: (27, 2)

In [7]:
```
from sklearn.model_selection import train_test_split
```

In [9]:
```
x_train, x_test, y_train, y_test = train_test_split(df[['age']],df.boug
ht_insurance,test_size= 0.1)
```

In [10]:
```
x_test
```

Out[10]:

	age
26	23
20	21
23	45

In [11]:
```
from sklearn.linear_model import LogisticRegression
```

In [12]:
```
model = LogisticRegression()
```

In [13]:
```
model.fit(x_train,y_train)
```

Out[13]:
```
LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=
True,
                   intercept_scaling=1, l1_ratio=None, max_iter=100,
                   multi_class='auto', n_jobs=None, penalty='l2',
                   random_state=None, solver='lbfgs', tol=0.0001, verbo
se=0,
                   warm_start=False)
```

In [14]:
```
model.predict(x_test)
```

Out[14]:
```
array([0, 0, 1], dtype=int64)
```

In [15]:
```
model.score(x_test,y_test)
```

Out[15]:
```
1.0
```

In [16]:
```
model.predict_proba(x_test)
```

Out[16]:
```
array([[0.88308844, 0.11691156],
       [0.90667966, 0.09332034],
       [0.32143809, 0.67856191]])
```

In [20]:
```
model.predict([[40]])
```

Out[20]:
```
array([1], dtype=int64)
```