Name: Atharva Suryakant Mali | Class: TY BCA | Roll No.: 04 | DS LAB

Experiment 7

Create a logistic regression model for insurance data set and CET score data set

1. Logistic regression model for insurance data set.

In [1]:

```
import pandas as pd
from matplotlib import pyplot as pt
```

In [2]:

```
df =pd.read_csv("insurance.csv")
```

In [3]:

```
df.head()
```

Out[3]:

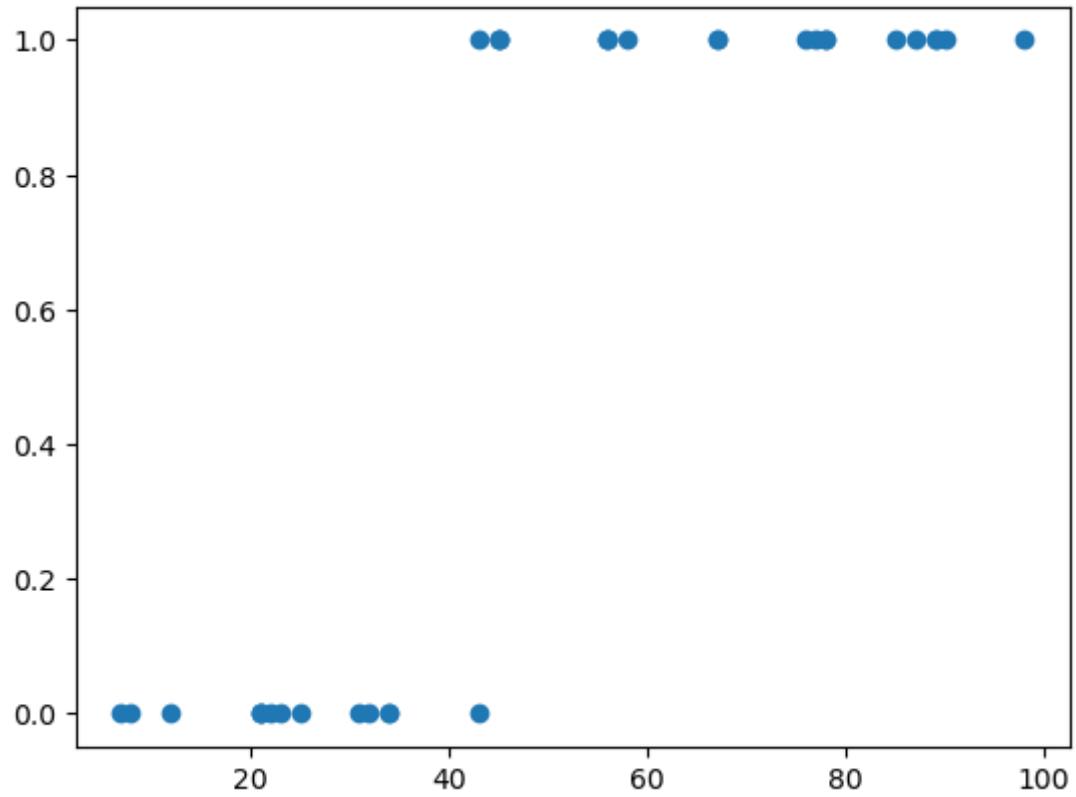
	age	insurance
0	12	0
1	34	0
2	56	1
3	23	0
4	56	1

In [4]:

```
pt.scatter(df.age,df.insurance)
```

Out[4]:

```
<matplotlib.collections.PathCollection at 0x23a1ec4d9f0>
```



In [5]:

```
from sklearn.model_selection import train_test_split
```

In [6]:

```
x_train, x_test, y_train, y_test = train_test_split(df[['age']],df.insurance,test_size=0.
```

In [7]:

```
len(x_train)
```

Out[7]:

```
30
```

In [8]:

```
len(x_test)
```

Out[8]:

```
4
```

In [9]:

```
len(x_test)
```

Out[9]:

```
4
```

In [10]:

```
from sklearn.linear_model import LogisticRegression
```

In [11]:

```
model = LogisticRegression()
```

In [12]:

```
model.fit(x_train, y_train)
```

Out[12]:

```
LogisticRegression
LogisticRegression()
```

In [13]:

```
x_test
```

Out[13]:

	age
33	58
5	78
8	90
19	45

In [14]:

```
model.predict(x_test)
```

Out[14]:

```
array([1, 1, 1, 1], dtype=int64)
```

In [15]:

```
model.predict([[5]])
```

```
C:\Users\surya\anaconda3\lib\site-packages\sklearn\base.py:420: UserWarnin
g: X does not have valid feature names, but LogisticRegression was fitted
with feature names
  warnings.warn(
```

Out[15]:

```
array([0], dtype=int64)
```

2. Logistic regression model for CET score data set.

In [16]:

```
import pandas as pd
from matplotlib import pyplot as pt
```

In [17]:

```
df =pd.read_csv("cet score.csv")
```

In [18]:

```
df.head()
```

Out[18]:

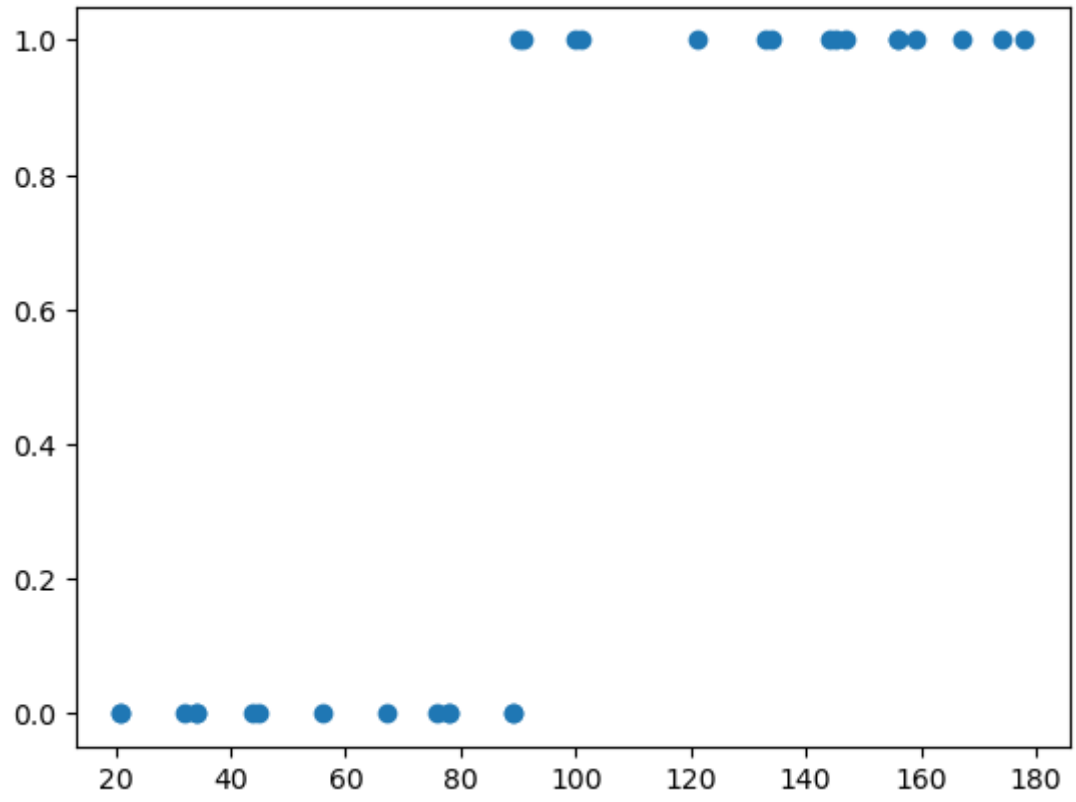
	CET_score	get_admission
0	121	1
1	100	1
2	56	0
3	45	0
4	78	0

In [19]:

```
pt.scatter(df.CET_score,df.get_admission)
```

Out[19]:

```
<matplotlib.collections.PathCollection at 0x23a21e8bfa0>
```



In [20]:

```
from sklearn.model_selection import train_test_split
```

In [21]:

```
x_train, x_test, y_train, y_test = train_test_split(df[['CET_score']],df.get_admission,te
```

In [22]:

```
len(x_train)
```

Out[22]:

```
27
```

In [23]:

```
len(x_test)
```

Out[23]:

```
3
```

In [24]:

```
len(x_test)
```

Out[24]:

```
3
```

In [25]:

```
from sklearn.linear_model import LogisticRegression
```

In [26]:

```
model = LogisticRegression()
```

In [27]:

```
model.fit(x_train, y_train)
```

Out[27]:

```
LogisticRegression
LogisticRegression()
```

In [28]:

```
x_test
```

Out[28]:

	CET_score
20	156
26	44
14	159

In [29]:

```
model.predict(x_test)
```

Out[29]:

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

In [30]:

```
model.predict([[100]])
```

```
C:\Users\surya\anaconda3\lib\site-packages\sklearn\base.py:420: UserWarnin
g: X does not have valid feature names, but LogisticRegression was fitted
with feature names
  warnings.warn(
```

Out[30]:

```
array([1], dtype=int64)
```