ML LAB ASSIGNMENT

SUPRATIM NAG -- CSE-AIML/22/057 -- GROUP-B

Q-4:Implementation of Multivariate Linear Regression

(a)Using my own dataset containing information of BMI, Age,Cholesterol Level,Blood Pressure and Heart Rate. Spliting the dataset into training and test dataset in 80:20 ratio. Then training the Linear Regression model on the training dataset and predict the Age for test dataset. (Multivariate Linear Regression)

In [ ]:
```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
```

In [ ]:
```
file_path="C:\\Users\SUPRATIM NAG\OneDrive\Documents\ML\Personal_Datasets\Dataset.csv"
df=pd.read_csv(file_path)
```

In [ ]:
```
X = df[[ 'Blood Pressure', 'Cholesterol Levels', 'Heart Rate', 'BMI']]
Y = df['Age']
```

In [ ]:
```
X.head(5)
```

Out[ ]:

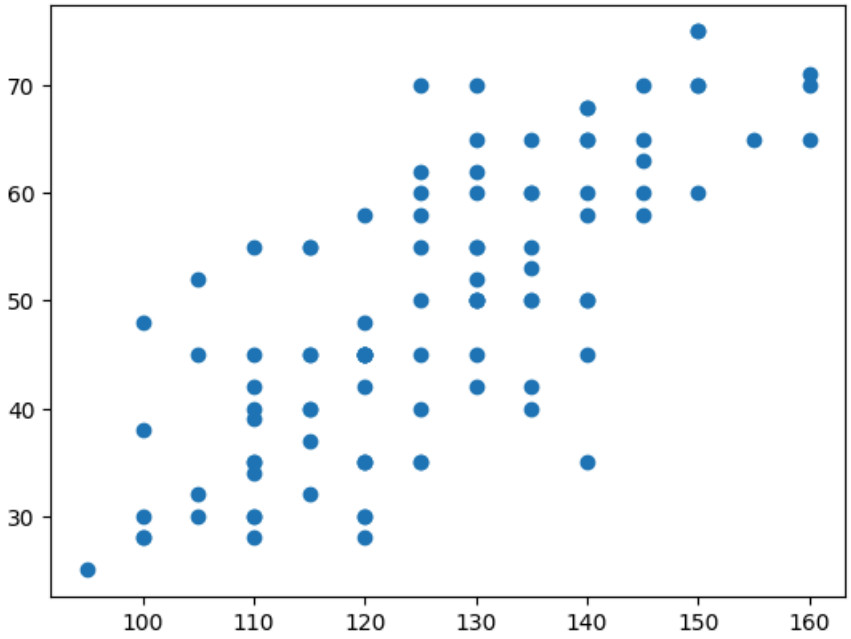
	Blood Pressure	Cholesterol Levels	Heart Rate	BMI
0	130	250	72	28.0
1	110	150	76	24.0
2	140	200	80	30.0
3	160	220	88	32.0
4	120	180	74	27.0

In [ ]:
```
Y.head(5)
```

Out[ ]:
```
0    65
1    42
2    58
3    71
4    35
Name: Age, dtype: int64
```

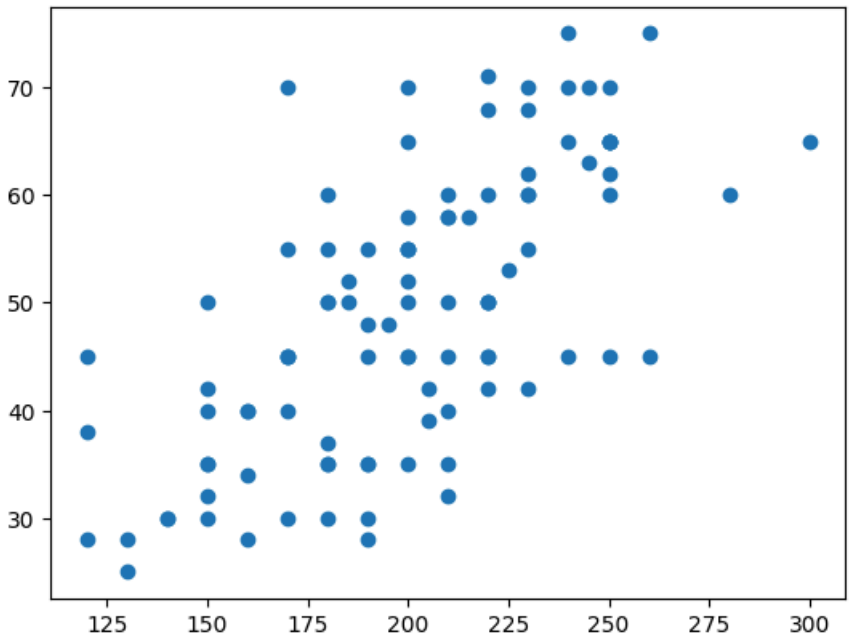
In [ ]:
```
plt.scatter(df['Blood Pressure'],df['Age'])
```

Out[ ]:
```
<matplotlib.collections.PathCollection at 0x1a1b2a80b50>
```



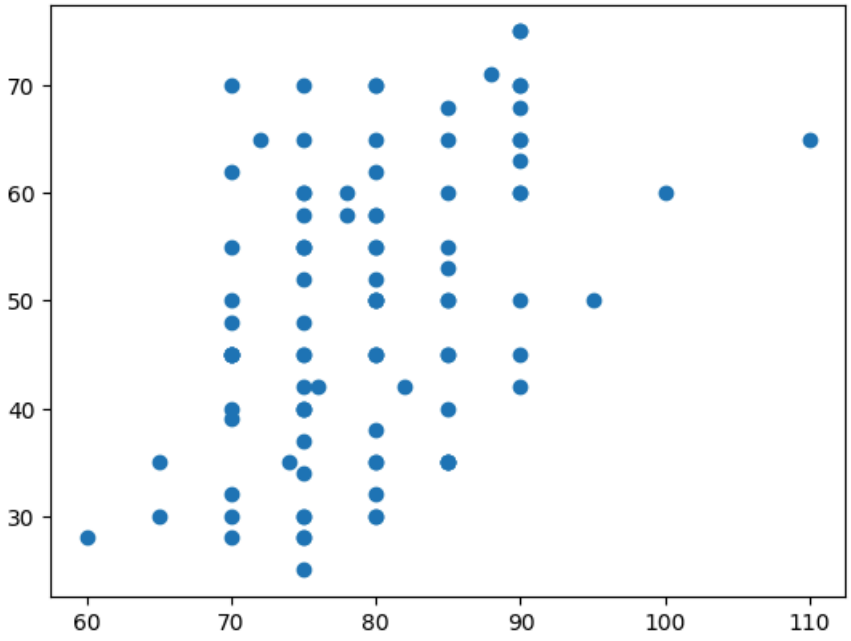
In [ ]:
```
plt.scatter(df['Cholesterol Levels'],df['Age'])
```

Out[ ]:
```
<matplotlib.collections.PathCollection at 0x1a1b2b12810>
```



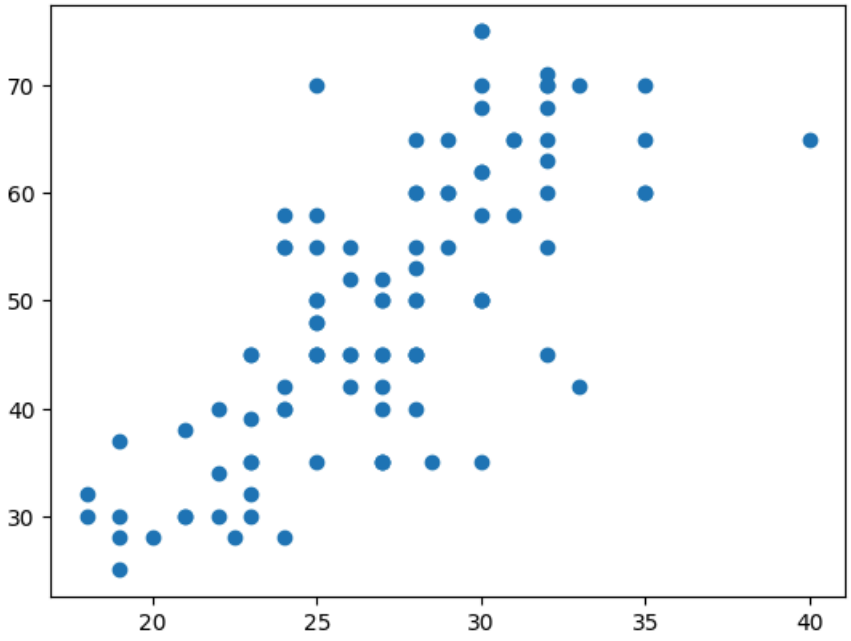
In [ ]:
```
plt.scatter(df['Heart Rate'],df['Age'])
```

Out[ ]:
```
<matplotlib.collections.PathCollection at 0x1a1b2a8f510>
```



In [ ]:
```
plt.scatter(df['BMI'],df['Age'])
```

Out[ ]:
```
<matplotlib.collections.PathCollection at 0x1a1b4d36490>
```



In [ ]:
```
x_train,x_test,y_train,y_test = train_test_split(X,Y,test_size=0.20)
len(x_train)
```

Out[ ]:
```
80
```

In [ ]:
```
x_train
```

Out[ ]:

	Blood Pressure	Cholesterol Levels	Heart Rate	BMI
79	130	205	82	26.0
85	145	245	90	32.0
98	120	190	75	23.0
50	130	220	85	28.0
78	120	180	80	19.0
...	...	...	...	...
29	125	180	80	25.0
71	130	220	85	28.0
25	140	220	85	32.0
66	100	190	70	25.0
34	100	120	80	21.0

80 rows × 4 columns

In [ ]:
```
y_train
```

Out[ ]:
```
79    42
85    63
98    30
50    50
78    30
      ..
29    50
71    60
25    68
66    48
34    38
Name: Age, Length: 80, dtype: int64
```

In [ ]:
```
multi_var = LinearRegression()
multi_var.fit(x_train,y_train)
```

Out[ ]:
```
LinearRegression()
```

In [ ]:
```
y_pred=multi_var.predict(x_test)
y_pred
```

Out[ ]:
```
array([58.01322814, 55.66418093, 57.0144665 , 38.60591063, 56.51354561,
       48.19670038, 63.9552295 , 52.07757406, 31.18817915, 44.62714621,
       61.18601235, 33.26635952, 34.82613107, 51.87202483, 54.26950593,
       47.11051299, 59.33642522, 39.59605936, 49.24595048, 57.98066282])
```

In [ ]:
```
y_test
```

Out[ ]:
```
88    42
8     35
5     60
60    52
0     65
74    35
42    60
23    52
86    32
69    45
12    60
70    30
93    30
94    50
17    70
4     35
62    65
44    35
96    70
56    62
Name: Age, dtype: int64
```

In [ ]:
```
x_test
```

Out[ ]:

	Blood Pressure	Cholesterol Levels	Heart Rate	BMI
88	135	230	90	33.0
8	140	180	85	30.0
5	135	230	78	29.0
60	105	185	80	26.0
0	130	250	72	28.0
74	125	190	80	27.0
42	140	250	75	32.0
23	130	200	75	27.0
86	105	150	70	18.0
69	120	200	80	25.0
12	135	180	75	35.0
70	110	150	70	18.0
93	110	170	80	21.0
94	130	220	80	27.0
17	130	170	80	32.0
4	120	180	74	27.0
62	135	250	85	32.0
44	110	150	65	23.0
96	125	200	70	25.0
56	130	230	70	30.0

In [ ]:
```
r_sq=multi_var.score(X,Y)
print(r_sq)
```

```
0.6303574846154316
```