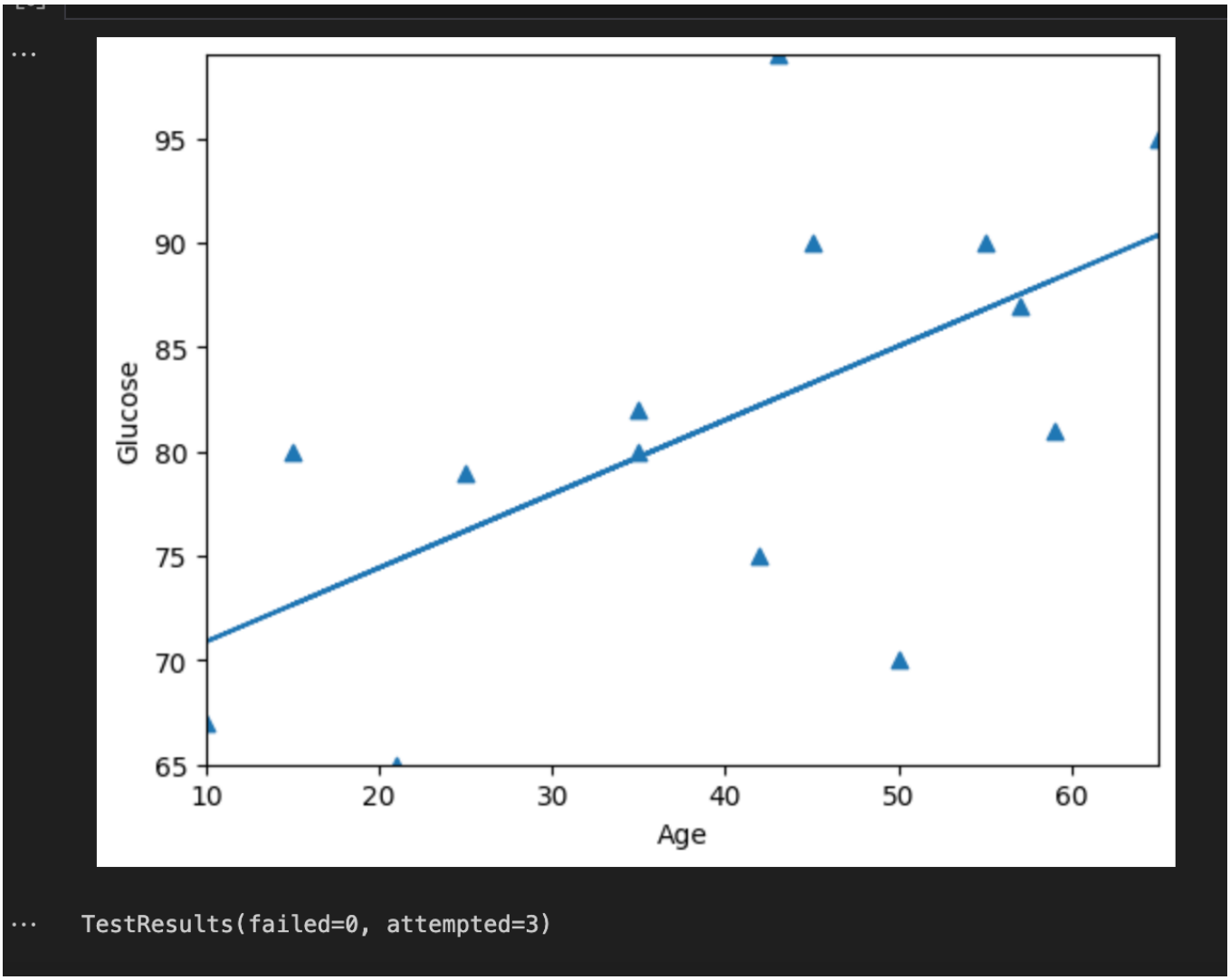
LINEAR REGRESSION REPORT

Yuhuan Huang

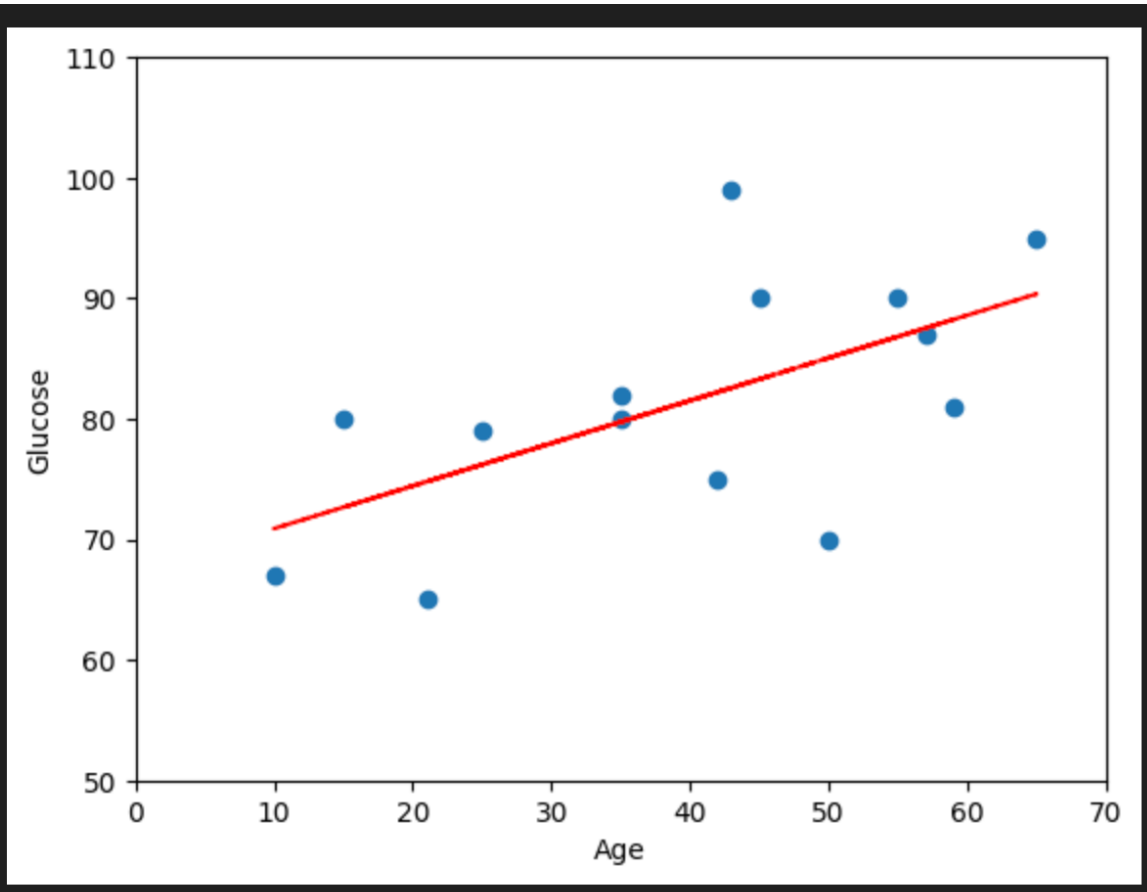
Part 1

Console Output:



My reflection: Age might not be a very good predictor of glucose.

Part 2




```
**********************************************************************
File "__main__", line 3, in __main__
Failed example:
    print('b1:', simple_LR_coeffs_sklearn(Age_reshaped, Glucose_reshaped)[0][0])
Expected:
    b1: [0.35391241]
Got:
    b1: 0.3539124057393233
**********************************************************************
File "__main__", line 5, in __main__
Failed example:
    print('b0:', simple_LR_coeffs_sklearn(Age_reshaped, Glucose_reshaped)[1][0])
Expected:
    b0: 67.34791357165692
Got:
    b0: 67.34791357165693
**********************************************************************
1 items had failures:
   2 of   3 in __main__
***Test Failed*** 2 failures.
```

Although the tests failed, I think the result is okay(?).

Part 3

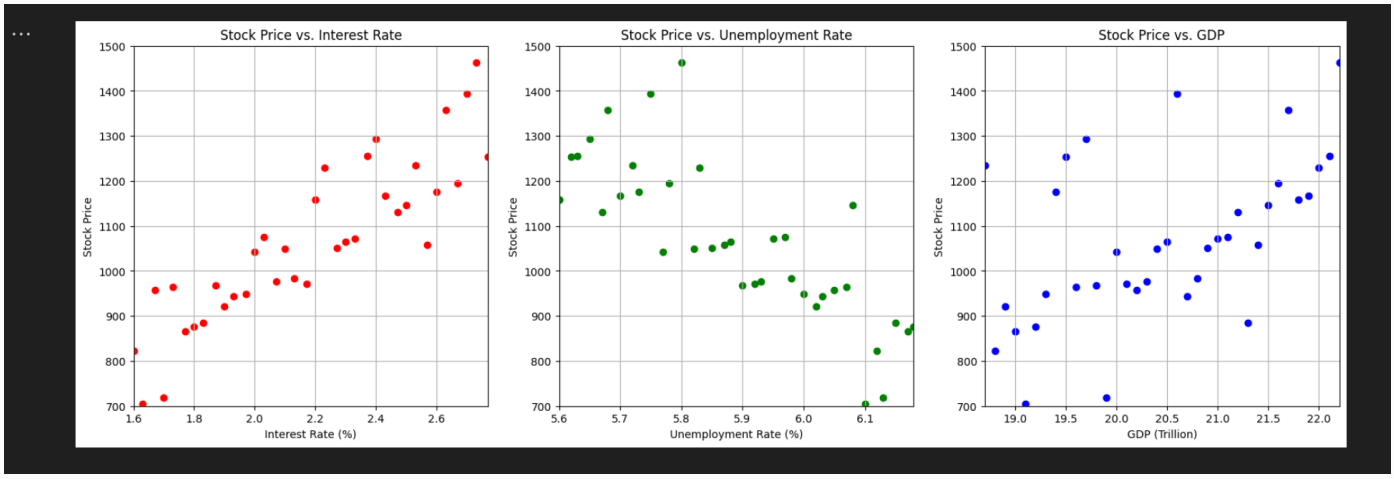
I preserved the results in rst.txt:

```python
# write into log
with open("rst.txt","w") as f:

  print('\nOutput linear model intercept, coefficients, score')
  Y = df['Stock_Index_Price'] #dependent var

  #Case 1:  Interest Rate only
  print('Case 1:  Interest Rate only:', file=f)
  X1 = df['Interest_Rate'] #extract interest rate from the dataframe as the indep vari
  X1out = MLR_sklearn(X1,Y)  #call your MLR function
  print(X1out[0],' ', X1out[1], ' ', X1out[2], '\n', file=f)
```

rst.txt

```
Case 1:  Interest Rate only:
[89.49634225]   [447.22813841]   0.7761962416936845

Case 2:  Unemployment Rate only:
[5872.5447392]   [-815.83031877]   0.6446566846067017

Case 3:  GDP only:
[-752.81591592]   [88.93693694]   0.27607375430501235

Case 4:  Interest Rate + Unemployment Rate:
[2197.75287066]   [ 327.09060652 -313.31647815]   0.815266916826149

Case 5:  Interest Rate + GDP:
[-288.47474076]   [415.47404803  21.87290205]   0.7889815680865808

Case 6:  Unemployment Rate + GDP:
[4425.83575581]   [-716.66435407   42.17389607]   0.6972114266765547

Case 7:  Interest Rate + Unemployment Rate + GDP:
[1802.39783781]   [ 299.610547   -307.27216408   20.52528504]   0.8265107950832802


```

```
[25]
...                Year     Month  Interest_Rate  Unemployment_Rate        GDP  \
count    36.000000  36.00000      36.000000          36.000000  36.000000
mean   2017.000000   6.50000       2.183333           5.891667  20.450000
std       0.828079   3.50102       0.351308           0.175507   1.053565
min    2016.000000   1.00000       1.600000           5.600000  18.700000
25%    2016.000000   3.75000       1.892500           5.745000  19.575000
50%    2017.000000   6.50000       2.185000           5.890000  20.450000
75%    2018.000000   9.25000       2.477500           6.035000  21.325000
max    2018.000000  12.00000       2.770000           6.180000  22.200000

       Stock_Index_Price
count          36.000000
mean         1065.944444
std           178.332746
min           704.000000
25%           955.750000
50%          1054.500000
75%          1180.000000
max          1464.000000

Output linear model intercept, coefficients, score
```



My Reflections:

From the graphs, we can see that the stock price is positively correlated with the interest rate and GDP, and the stock price is negatively correlated with the unemployment rate.

From the results, we can see that when we extract some features in X to predict y, we can more clearly see the relationships between the features. For example, when we only use the interest rate and the unemployment rate as the "X" to predict the stock price, we can see that it is still highly-predictable, and they are positively related. But as we include more explanatory variables, the R-square increases,

showing that our model becomes better at predictions.