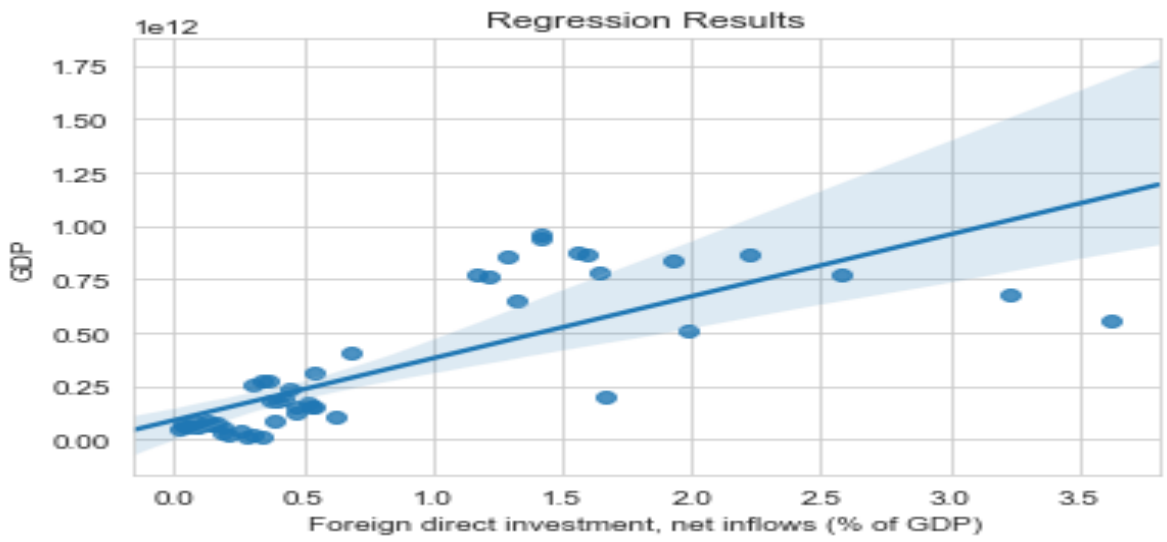
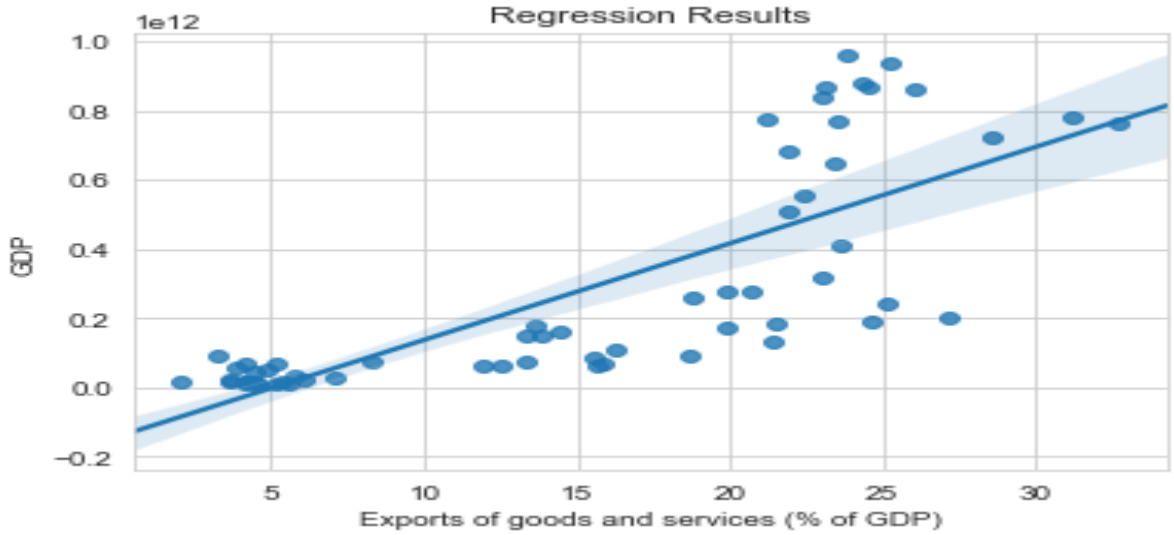
Homework 2:

Advanced Data Analysis in Python

The purpose of this assignment is to create a linear module with the help of numpy and pandas modules using the formula in hw2.pdf and compare it with a built-in linear regression module.

Plots about the relationship of 'Foreign Direct Investment' and 'Exports of Good and Services' with GDP are as follows.





Since ‘Export and Import good&services’’ correlation coefficient is 0.95, I eliminated ‘import good&services’ . The results of linear regression by hand are below:

	coefficients	standard error	lower bound	upper bound
	0	0	0	0
0	58.295464	1.451748	2.979502e+11	1.467416e+12
1	-1.840258	0.910352	-3.057217e+10	-1.039636e+10
2	0.307577	0.096352	5.333258e+10	1.888286e+11

And the results of linear regression by built-in module are below.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      0   R-squared:                       0.178
Model:                            OLS   Adj. R-squared:                  0.143
Method:                 Least Squares   F-statistic:                     5.097
Date:                Sun, 12 Dec 2021   Prob (F-statistic):            0.00992
Time:                        12:29:45   Log-Likelihood:                -142.04
No. Observations:                  50   AIC:                             290.1
Df Residuals:                      47   BIC:                             295.8
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         58.2955      1.452     40.155      0.000      55.375      61.216
0             -1.8403      0.910     -2.021      0.049      -3.672      -0.009
1              0.3076      0.096      3.192      0.003       0.114       0.501
==============================================================================
Omnibus:                        0.059   Durbin-Watson:                   0.531
Prob(Omnibus):                  0.971   Jarque-Bera (JB):                0.023
Skew:                          -0.003   Prob(JB):                        0.989
Kurtosis:                       2.896   Cond. No.                         45.9
==============================================================================
```

When we compare the results of the two modules, we see that the standard error and coefficients are same. As module by hand the formula: GDP = - FDI *1.8403 + EGS*0.3076 + 58.2955