Predicting Gross Domestic Product

Purpose:

To predict the GDP of countries based on Trade, Real Interest Rate, and Total Revenue.

Dataset:

The data is taken from the world bank indicator (WBI) Website which consists of 3 variables and each variable has 468 observations. The data is taken from 1970 to 2021 for nine developed countries the United States, the United Kingdom, Italy, France, Portugal, Germany, Australia, New Zealand, and India. The variables are described is given below.

Variables	Description
GDP	Gross Domestic Product Per Capita
LnGDP	Log of Gross Domestic Product Per Capita
Trade	Trade
IR	Real Interest Rate
TR	Total Revenue

Dependent Variable: Log of Gross Domestic Product Per Capita (ln GDP)

Independent Variables: Trade, Real Interest Rate (IR), Total Revenue (TR).

Model Specification:

We used a multiple linear regression model because we have more than one independent variable and the dependent variable is in quantitative form. The independent variables Trade, IR, and TR are used according to significance because the P-values of these three variables are less than the critical value 0.05, which means that these three variables have a statistically significant effect on GDP per capita.

Assumptions:

The assumptions of the multiple linear regression model are given below.

1) **Linearity:** There should be a linear relationship between dependent and independent variables.
2) **Normality**: The residuals of the model should follow the normal distribution.
3) **Homoscedasticity:** The variances of the residuals should be constant.
4) **No Autocorrelation:** There should be no high correlation between the values of residuals.
5) **No Endogeneity:** No relationship between residuals and the independent variables should exist.
6) **No Multicollinearity:** There should be no high correlation between independent variables.

Model Output:

The output of model using Python is given below.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   lnGDP   R-squared:                       0.505
Model:                             OLS   Adj. R-squared:                  0.502
Method:                  Least Squares   F-statistic:                     158.0
Date:                 Wed, 30 Nov 2022   Prob (F-statistic):           1.43e-70
Time:                         15:40:14   Log-Likelihood:                -667.09
No. Observations:                  468   AIC:                             1342.
Df Residuals:                      464   BIC:                             1359.
Df Model:                            3
Covariance Type:             nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      7.7004      0.178     43.203      0.000       7.350       8.051
Trade          0.0082      0.004      2.241      0.025       0.001       0.015
IR            -0.1125      0.009    -12.119      0.000      -0.131      -0.094
TR             0.0664      0.007      9.656      0.000       0.053       0.080
==============================================================================
Omnibus:                       14.267   Durbin-Watson:                   0.217
Prob(Omnibus):                  0.001   Jarque-Bera (JB):               15.472
Skew:                          -0.368   Prob(JB):                     0.000437
Kurtosis:                       3.500   Cond. No.                         216.
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Interpret:

As one unit increases in trade, GDP per capita also increases by 0.8%, which means that there has a positive effect of trade on GDP per capita. When one unit increases in interest rate, GDP per capita decreases by 11.3%, which means that there has a negative effect of interest rate on GDP per capita. As one unit increases in total revenue, GDP per capita also increases by 6.6%, which means that there has a positive effect of total revenue on GDP per capita.

The coefficient of determination (R^2) value of the model is 0.51, which means that the variation in GDP per capita is explained 52 % by the variation in trade, real interest rate, and total revenue. The P-value of the F-test is less than the critical value, which means that the overall model is statistically significant.

Reflection:

We learned how to use Python in this project. We also learned how to estimate the multiple linear regression model for testing the effect of one or more than one independent variable on the dependent variable. We faced the issue of the insignificance effect of total revenue on GDP per capita, due to this problem we transformed the dependent variable GDP by taking the log of the variable.