DOES A HIGH/LOW INCOME TAX STIMULATE THE COUNTRY'S GDP PER CAPITA

FUNDAMENTALS OF DATA ANALYSIS - FINAL PROJECT

by

Aim: To analyse how the role of income taxes and other independent variables in the European Union affects the Economic Growth by evaluating the GDP growth and GDP per capita to understand the impact.

Team Members: Aswin Subramanian Maheswaran, Andrei Baraitaru, Luis Gomez-Acebo, Jaskaran Singh Ghai, Angel Lopez, Nicolas Leyva

Professor: Borja Mesa Sanchez

Date: 27th May, 2023.

TABLE OF CONTENTS

ABSTRACT

This study examines the relationship between income tax rates and economic growth, using GDP per capita and GDP growth rate as dependent variables in the European Union, as our focus group. Our aim is to critically investigate the common claims on whether low- or high-income tax rates have an influence on the economic growth of a country. Specifically, our study will be based on the null hypothesis that low-income taxes have a negative effect on economic growth, against the alternative hypothesis that low-income taxes stimulate economic growth. Our economic growth is based on the two-evaluation metrics – GDP growth rate and GDP per capita.

The paper uses a European Union dataset that encompasses a broad range of countries and tax regimes, thereby ensuring a comprehensive exploration of the effects of income tax rates. The data was subjected to an intensive exploratory data analysis and econometric statistical techniques using ANOVA which accounted for the comparisons and variation changes in the country-specific characteristics. The results suggest that the low-income tax rates can have detrimental effects on economic growth too, but we are going to explore our data and understand the relationship between the variables in our data and make our model more accurate.

The major findings indicate that the relationship between income tax rates and economic growth (Variables - GDP growth rate and GDP per capita) are related and we have an additional independent variable, "Unemployment" to account for the additional changes. Also theoretically based on the literature review, the unemployment rate can affect income tax revenue because individuals who are unemployed typically do not pay income taxes. Therefore, higher unemployment rates could result in lower income tax revenue.

GLOSSARY OF TERMS

- ANOVA (Analysis of Variance): A statistical method used to analyse the differences between group means by comparing the variances within and between groups.

- GDP Growth Rate: A measure that quantifies the change in the value of a country's Gross Domestic Product (GDP) over a specific period, typically expressed as a percentage.

- GDP per Capita: A measure of the economic output per person in a country, obtained by dividing the Gross Domestic Product (GDP) by the population.

- Income Tax: A tax levied by the government on an individual's income, which is typically a percentage of the income earned and is used to fund public expenditures.

- Correlation: A statistical measure that indicates the strength and direction of the relationship between two variables. It quantifies the extent to which changes in one variable are associated with changes in another.

- Outliers: Data points that deviate significantly from the overall pattern of a dataset, either being unusually high or low. Outliers can indicate potential errors, anomalies, or exceptional cases within the data.

- Null and Alternative Hypothesis: The null hypothesis represents the assumption of no effect or no relationship between variables, while the alternative hypothesis suggests the presence of a significant effect or relationship. In hypothesis testing, statistical analysis is conducted to determine whether the evidence supports rejecting the null hypothesis in favour of the alternative hypothesis.

- Inferential Statistics: Statistical methods and techniques used to make inferences and draw conclusions about a population based on a sample. Inferential statistics involve estimating population parameters and testing hypotheses.

- Regression: A statistical technique used to model the relationship between a dependent variable and one or more independent variables. It helps to understand how changes in the independent variables impact the dependent variable.

INTRODUCTION

Income tax policies play an important role in shaping a country's economic landscape. The question of whether high- or low-income tax rates stimulate a country's GDP per capita has been a subject of significant interest and debate among economists, policymakers, and researchers. In this study, we aim to analyse the relationship between income tax and GDP per capita, exploring the potential impacts of high- and low-income tax rates. By examining various perspectives, conducting a comprehensive analysis, and evaluating empirical evidence, our study seeks to shed light on the complex dynamics between income tax policies and economic outcomes.

The motivation behind selecting this topic *"Does a high/low-income tax stimulate the country's GDP per capita"* for our data analysis report focuses on its ability to display the economic secrets of Europe. By delving into the interesting, yet a complex relationship between income tax and GDP per capita, we unlock insights into the factors that drive economic growth and prosperity. Through the lens of econometrics, we harness the power of statistical analysis to decipher the hidden patterns and connections between these variables. This exploration not only deepens our knowledge of the European economy but also equips us with invaluable skills to unravel the complex patterns of economic dynamics. Lastly, gaining knowledge in this area equips us with valuable skills and understanding of economic principles, applicable beyond the scope of this specific analysis.

In our hypothesis testing, we formulate a null hypothesis and an alternative hypothesis to make claims about the relationship between variables. In our case, we have two dependent variables: GDP_Capita (GDP per capita) and GDP_Growth (GDP growth rate). The hypotheses related to the effect of income taxes on economic growth can be stated as follows:

- **Null Hypothesis:** Income taxes have a negative effect on economic growth. This means that the null hypothesis assumes that higher income taxes are associated with lower levels of economic growth for both GDP_Capita and GDP_Growth.
- **Alternative Hypothesis:** Income taxes have a positive effect on economic growth. The alternative hypothesis suggests that lower income taxes are linked to higher levels of economic growth for both GDP_Capita and GDP_Growth.

In hypothesis testing, we will analyse the data to determine whether the evidence supports rejecting the null hypothesis in favour of the alternative hypothesis. By conducting appropriate statistical tests and examining the significance levels, we can assess the strength of the evidence and make conclusions about the relationship between income taxes and economic growth for the specified dependent variables.

To complete successful research, we conducted an extensive literature review for the year 2021 because the datasets were latest and convenient for the particular year. We relied on the findings of the recent research performed on tax inflation rates, articles such as, ***"Google Scholar // "Financial crisis, austerity, and health in Europe" - M Karanikolos, P Mladovsky, J Cylus, S Thomson***" to confirm the relationship among the variables based on previous theory and find other important variables.

METHODOLOGY

The topic of interest is whether a low income tax rate has an impact on a country's GDP per capita within the European Union (EU). To investigate our research title, the data was collected and utilised from the OECD and EU official sources for the year 2021 (*only recent available sources)*. The population of interest is the entire EU, and the sample consists of the countries within the EU. The study considers two independent variables: "*Unemployment" and "Income Tax"*, and two dependent variables: "*GDP per capita and growth."*

To prepare the data, the values for each country were manually inputted into an Excel sheet. This likely involved collecting data on unemployment rates, income tax rates, GDP per capita, and growth figures for each country. After organising the data, statistical analysis will be conducted to examine the relationship between income tax and GDP per capita based on our hypothesis claims.

By analysing the data, the study aims to determine whether higher or lower income tax rates correspond to changes in a country's GDP per capita. This investigation can provide insights into the potential stimulatory effects of income tax policy on economic growth within the EU.

a. Assessment and Measures

The primary assessment measure includes the data cleaning process to identify null values and the outliers. To understand the type of data better, we will perform exploratory data analysis, starting with the descriptive statistics to know the interaction between our independent and dependent variables. To recognize our interaction and understand our hypothesis claims, ANOVA test will be applied to our variables and a table will be printed to analyse our results and evaluate the claims we produced. Since our research focuses on the effects of Income Tax. We would compare the income tax rate and their influence on the economic metrics by categorising sublevels of the Tax Group.

b. Conditions of the Model

For our empirical analysis of the model, we based our results on a set of assumptions on the validity and reliability of the results.
Looking at the Linearity and Relationship existence among the variables, correlation was considered when comparing the variables and relationship was explained with the evidence. For instance, we found a negative correlation of -0.04 between Income tax and GDP growth rate which highlights that both variables move in opposite directions. Although it cannot indicate causation, we are certain that there are other variables which could show the relationship in a better way. Our second external variable, "Unemployment" is to account for the changes in the Income Tax and explain the economic variation in a better way. The correlation coefficient against Unemployment with GDP Growth and GDP per capita is -0.31 and -0.21 respectively, indicating that there is a negative correlation between the variables and helps us to know that the *low* and *high* unemployment rates should be differentiated, so we can identify which type of unemployment rate affects the economic metrics.

c. ANOVA Test

To begin the model development for our research project, the inferential statistics is performed using ANOVA as it examines the sources of variance within our dataset by comparing the means of different groups and determining whether the differences between them are significant.

ANOVA is used to test the null hypothesis and the alternative hypothesis for different groups of countries, each characterised by different income tax rates. We have partitioned these countries into groups based on their tax rates (e.g., low, medium, high) and compared the means of the dependent variables (GDP per capita and GDP growth rate) for these groups. In our ANOVA table, we look at the four variables – degrees of freedom, SS (Sum of Squares), MS (Mean Squares), F-statistics and the p-value which will be compared with the significance level of the test.

It's important to remember that ANOVA has several pre-conditions, such as, the observations are independent, the data follows a normal distribution within each group, and the variances are equal across groups (homoscedasticity is satisfied). Violation of these assumptions may lead to different conclusions, so you should make sure these conditions are met before proceeding with ANOVA.

DESCRIPTIVE STATISTICS

The provided statistics are relevant to the member states of the European Union (EU). In terms of GDP growth rates, the average rate of 6.23 indicates positive economic performance across the EU, with variations observed as denoted by the standard deviation of 2.59. Regarding income tax, the mean rate of 21.92 reflects the percentage of taxation within the EU, while the standard deviation of 8.78 suggests differences in tax policies among member states. As for GDP per capita, the average value of 57,185.85 signifies the average economic output per individual, with considerable disparities evident from the large standard deviation of 25,426.99. Finally, the mean unemployment rate of 35.19 reveals the average joblessness within the EU, with variations in labour market conditions observed through the standard deviation of 11.40. These statistics shed light on the economic performance, tax levels, living standards, and employment situations across the diverse member states of the European Union. Our two dependent variables display right skewness and the distribution of the data for both variables is shifted towards higher values, resulting in a longer tail on the right side of the distribution. This indicates that there are relatively fewer data points with lower values and a greater concentration of data towards the higher end of the scale for both variables.

Table. Shows countries with their index numbers and their variables and z-score.

```
#We keep the data within 3 standard errors.
df_clean = df[(df_z<3).all(axis=1)]
df_clean
```

	INCOME_TAX	GDP_GROWTH_RATE	GDP_CAPITA	UNEMPLOYMENT
0	22.4	4.6	68775	31.523942
1	26.7	6.1	65818	42.296179
2	10.3	7.6	32233	28.248467
3	20.1	13.1	40214	20.285690
4	19.7	6.6	43318	24.156616
5	9.1	3.5	49957	29.592654
7	20.4	8.0	47721	62.914613
8	29.9	3.0	59808	31.153003
9	21.0	6.8	55698	29.943598
10	26.6	2.6	64101	57.981495
11	9.0	8.4	37212	33.993836
12	15.4	7.1	42665	20.298172
13	32.8	13.6	128343	26.580647
14	25.9	6.7	52461	43.325757
15	19.6	4.1	40685	56.623103
16	23.4	6.0	49265	41.731004
18	25.0	6.3	48726	49.447023
19	21.7	4.9	70966	36.959445
20	14.6	6.8	43606	34.165306
21	19.6	5.5	42341	25.283304
22	23.0	5.1	42372	30.331301
23	10.9	3.0	36704	36.669417
24	14.5	8.2	50713	27.079835
25	22.8	5.5	46748	41.054389
26	29.0	5.1	65209	36.581418

In our descriptive statistics, we found some outliers through the boxplots we plotted to identify them. There are two different outliers in two distinct countries. Our method to find the outliers relied on the python code and their corresponding z-score and as a rule of statistics, the ones above > 3 are ruled as outliers. From our above results, it's interesting to note that number 6, "*Denmark"* has an outlier in the income tax rate part and *"Luxembourg"* belonging to number 17, has an outlier in the GDP per capita variable section. The outliers could be caused by a specific event, policy change, or economic condition which lead to an unusually high- or low-income tax rate compared to other countries in the dataset, whereas for Luxembourg, there is a small population and a highly developed financial sector, which could lead to a significantly higher GDP per capita compared to other countries.

MODEL

To create the best possible model to predict the economic growth based on the independent variables we have selected, our first approach was to separate the Income Tax into 3 groups: the low category, "*L*" are the countries with an income tax lower than 17.5%, medium, "*M*" is higher than 17.5% and lower than 26.05% and the high, "*H*" are values higher than 26.05%. We also separated the unemployment countries by halves, L for the lower half and H for the upper half. By performing this step, we had a model which could consider the different levels of a variable and help us measure the effect more accurately.

Performing two ANOVAs to investigate the influence of different income tax levels and unemployment rates on GDP growth and GDP per capita, along with assessing the interaction effect, holds significant importance in our analysis. By conducting separate ANOVAs for income tax and unemployment, we can understand their individual impacts on

GDP growth and GDP per capita, allowing us to discern the specific contribution of each variable in isolation. Additionally, analysing the interaction effect provides insights into whether the combined influence of income tax and unemployment has a significant impact on the economic indicators. This comprehensive analysis helps us gain a deeper understanding of the complex relationship between income tax, unemployment, and the economic outcomes of GDP growth and GDP per capita. By considering these multiple variables and their interaction, we can unravel the intricate dynamics shaping the economic landscape and shed light on the factors driving economic growth and prosperity.

RESULTS

ANOVA Table 1. GDP_Growth Rate

	df	sum_sq	mean_sq	F	PR(>F)
C(Tax_Group_3levels)	2.0	1.987725	0.993862	0.140781	0.869493
C(Unemployment_Group)	1.0	17.880944	17.880944	2.532838	0.126442
C(Tax_Group_3levels):C(Unemployment_Group)	2.0	5.955008	2.977504	0.421764	0.661323
Residual	21.0	148.252619	7.059649	NaN	NaN

ANOVA Table 2 GDP_ per capita

	df	sum_sq	mean_sq	F	PR(>F)
C(Tax_Group_3levels)	2.0	8.114487e+09	4.057244e+09	13.407602	0.000177
C(Unemployment_Group)	1.0	6.398072e+08	6.398072e+08	2.114312	0.160711
C(Tax_Group_3levels):C(Unemployment_Group)	2.0	1.700783e+09	8.503917e+08	2.810212	0.082898
Residual	21.0	6.354762e+09	3.026077e+08	NaN	NaN

After conducting the first ANOVA test to examine the impact on GDP growth rate, with an alpha level of 0.05, the obtained results indicate the following:

The p-value of the F-statistic for the variable "Income Tax" is 0.87, which is considered not statistically significant . Similarly, the p-value for the variable "Unemployment" is 0.13, also considered insignificant. Additionally, the p-value for the interaction between the variables "Income Tax" and "Unemployment" is 0.66, which is also deemed insignificant. Since all of these p-values are insignificant, we cannot reject the null hypothesis, implying that there is no evidence to suggest a difference in means among the groups.

After conducting the second ANOVA test to assess the impact on GDP per capita, with an alpha level of 0.05, the following observations can be made based on the obtained results:

The p-value of the F-statistic for the variable "Income Tax" is 0.00018, which is considered statistically significant. This suggests that the variable "Income Tax" has a significant impact on GDP per capita. The p-value for the variable "Unemployment" is 0.16, which is greater than the chosen alpha value. Therefore, it is not considered statistically significant, indicating that the variable "Unemployment" may not have a significant effect on GDP per capita.

The p-value for the interaction term between "Income Tax" and "Unemployment" is 0.08, which is slightly above the chosen alpha value. While it is not statistically significant, it suggests a potential trend or tendency towards significance. It indicates that the variable

"Income Tax" has a significant effect on GDP per capita, while the impact of "Unemployment" alone may not be statistically significant.

From our inferential statistics, we have reached a conclusion which states that income tax rates of three different groups affects the GDP per capita. The interaction between both our independent variables influences affects the GDP per capita, proving there are additional factors which affects the GDP Growth Rate and hence there are external unaccounted variables, such as *Government Spending, Education Levels and Inflation Rate* which could be the causes of this result.

CONCLUSION

We have learned important things about the relationship between income tax rates, unemployment, and their effects on GDP growth rate and GDP per capita based on the findings of our ANOVA tests. The first ANOVA showed that, as shown by the negligible p-values, neither income tax nor unemployment had a significant impact on GDP growth rate. The second ANOVA, however, produced conflicting findings, demonstrating that while unemployment had no statistically significant impact on GDP per capita, income tax had a considerable impact. An inclination toward significance was seen in the income tax and unemployment interaction term.

These results imply that the observed results may be influenced by various external factors, such as government spending, educational attainment, and inflation rates. Therefore, it is essential to take these extra factors into account when analysing the intricate dynamics affecting GDP growth and GDP per capita. Our analysis underlines the complex nature of economic events and underscores the necessity for a thorough strategy that takes into account a variety of elements in order to fully comprehend the connection between income tax, unemployment, and economic consequences.

In conclusion, our research on the impact of income taxes on GDP per capita has shown the complex interrelationships between these factors. Our data indicate that there are numerous variables at work, making it difficult to draw a simple conclusion about how income tax affects GDP per capita. While we found that income taxes had a considerable impact on GDP per person, the direction of the effect changed depending on the particular situation and other influencing factors.

These findings underline the necessity for a detailed analysis of the relationship between income tax and GDP per capita that takes into consideration a variety of economic factors, political frameworks, and regional differences. Overall, this analysis offers a starting point for additional investigation and lays the groundwork for scholars and policymakers to delve further into the complex relationship between income tax laws and economic outcomes.

In retrospect, we could have added few qualitative variables and analyse the trends in a better way. This could have given us a insight onto how economic variables other than mentioned in this report could have influenced the dependent variables and affected our study. Using regression techniques, could have also been more efficient in our study since it's an econometric study.

REFERENCES

- OECD (2023), Tax on goods and services (indicator). doi: 10.1787/40b85101-en (Accessed on 24 April 2023)

- OECD (2023), General government spending (indicator). doi: 10.1787/a31cbf4d-en (Accessed on 24 April 2023)

- Google Scholar // "Financial crisis, austerity, and health in Europe" - M Karanikolos, P Mladovsky, J Cylus, S Thomson

- Stoilova, D. (2017, July 1). *Tax structure and economic growth: Evidence from the European Union*. Contaduría y Administración. https://www.elsevier.es/es-revista-contaduria-administracion-87-articulo-tax-structure-economic-growth-evidence

- GARETH D. MYLES. (n.d.). *Taxation and economic growth - JSTOR*. Taxation and Economic Growth. https://www.jstor.org/stable/pdf/24437591.pdf?ab_segments=0%2Fbasic_search_gsv2%2Fcontrol&initiator=

- Page, Benjamin R., Joseph Rosenberg, James R. Nunns, Jeffrey Rohaly, and Daniel Berger. 2017. "Macroeconomic Analysis of the Tax Cuts and Jobs Act." Washington DC: Urban-Brookings Tax Policy Center.

APPENDIX

Table 1. Data Set for European Union in 2021

	A	B	C	D	E
1	Country	GDP_growth	Income_Tax	GDP_capita	Unemployment
2	Austria	4.6	22.4	68775	31.52394204
3	Belgium	6.1	26.7	65818	42.29617899
4	Bulgaria	7.6	10.3	32233	28.24846715
5	Croatia	13.1	20.1	40214	20.28569023
6	Cyprus	6.6	19.7	43318	24.15661586
7	Czech Republi	3.5	9.1	49957	29.59265423
8	Denmark	4.9	52.2	74962	32.56779605
9	Estonia	8	20.4	47721	62.91461301
10	Finland	3	29.9	59808	31.15300293
11	France	6.8	21	55698	29.94359845
12	Germany	2.6	26.6	64101	57.98149513
13	Greece	8.4	9	37212	33.99383619
14	Hungary	7.1	15.4	42665	20.2981723
15	Ireland	13.6	32.8	128343	26.58064663
16	Italy	6.7	25.9	52461	43.32575721
17	Latvia	4.1	19.6	40685	56.62310254
18	Lithuania	6	23.4	49265	41.73100364
19	Luxembourg	5.1	26.2	143397	19.31458284
20	Malta	6.3	25	48726	49.44702305
21	Netherlands	4.9	21.7	70966	36.95944517
22	Poland	6.8	14.6	43606	34.16530614
23	Portugal	5.5	19.6	42341	25.28330423
24	Romania	5.1	23	42372	30.33130144
25	Slovakia	3	10.9	36704	36.66941714
26	Slovenia	8.2	14.5	50713	27.07983494
27	Spain	5.5	22.8	46748	41.05438857
28	Sweden	5.1	29	65209	36.58141833

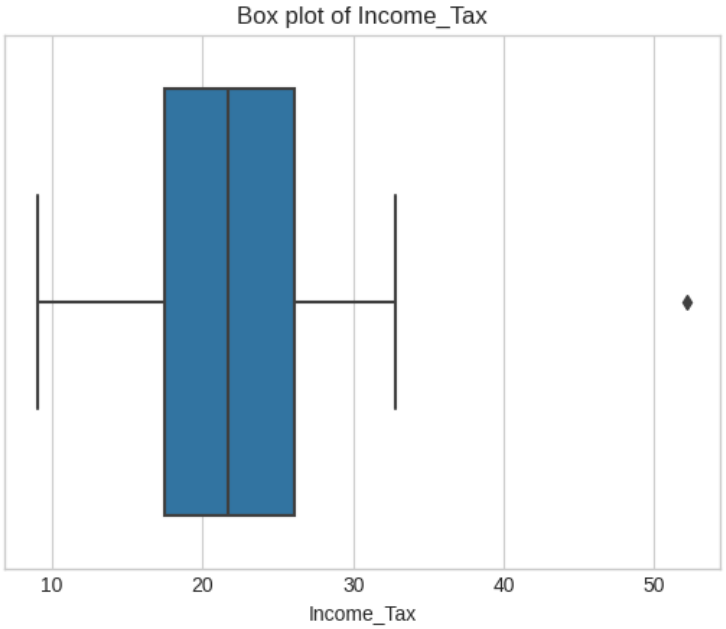
Table 2. Description and Correlation of the Data Set

	GDP_growth	Income_Tax	GDP_capita	Unemployment
count	27.000000	27.000000	27.000000	27.000000
mean	6.229630	21.918519	57185.851852	35.188985
std	2.587517	8.776982	25426.999171	11.397225
min	2.600000	9.000000	32233.000000	19.314583
25%	4.900000	17.500000	42518.500000	27.664151
50%	6.000000	21.700000	49265.000000	32.567796
75%	6.950000	26.050000	64655.000000	41.392696
max	13.600000	52.200000	143397.000000	62.914613

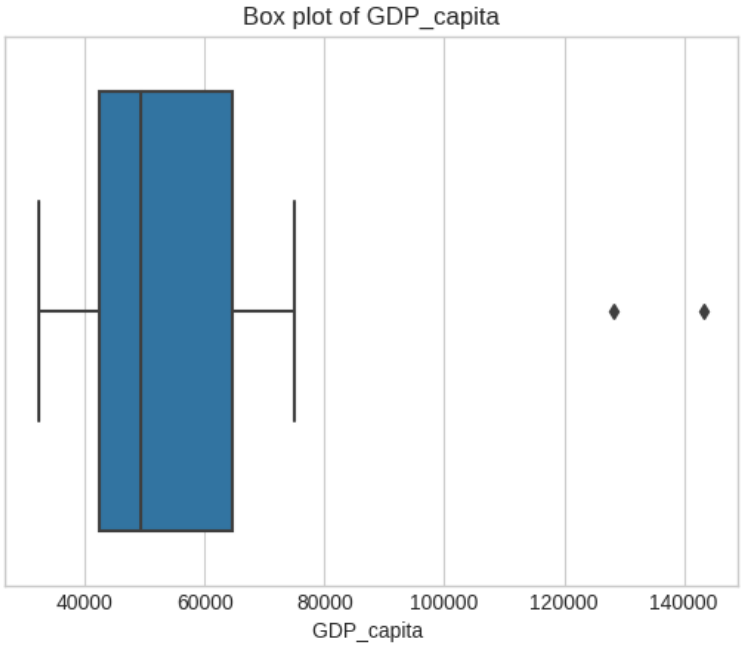
	GDP_growth	Income_Tax	GDP_capita	Unemployment
GDP_growth	1.000000	-0.041737	0.133081	-0.307790
Income_Tax	-0.041737	1.000000	0.519125	0.101849
GDP_capita	0.133081	0.519125	1.000000	-0.211271
Unemployment	-0.307790	0.101849	-0.211271	1.000000

Table 3. Boxplots of all our variables

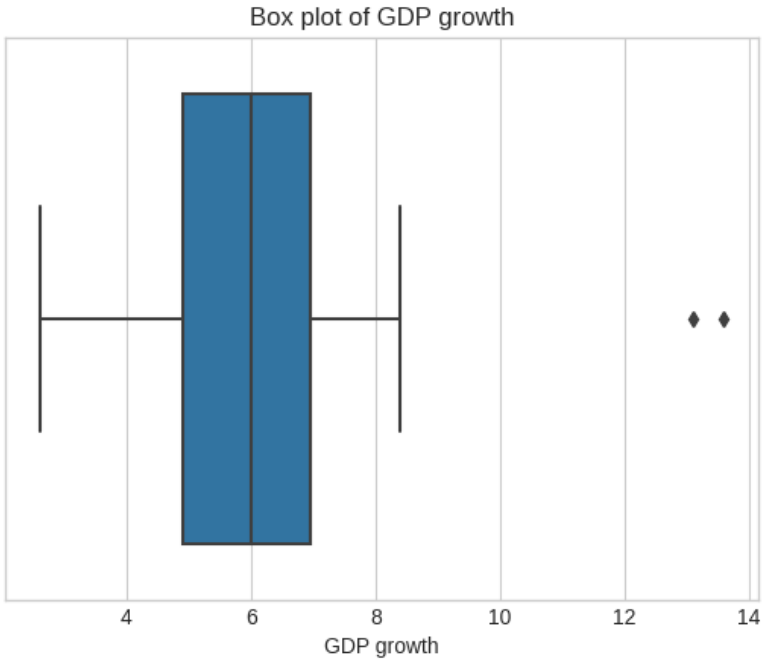
a. Graph 1. Income Tax



b. Graph 2. GDP per capita



c. Graph 3. GDP Growth Rate



d. Graph 4. Unemployment

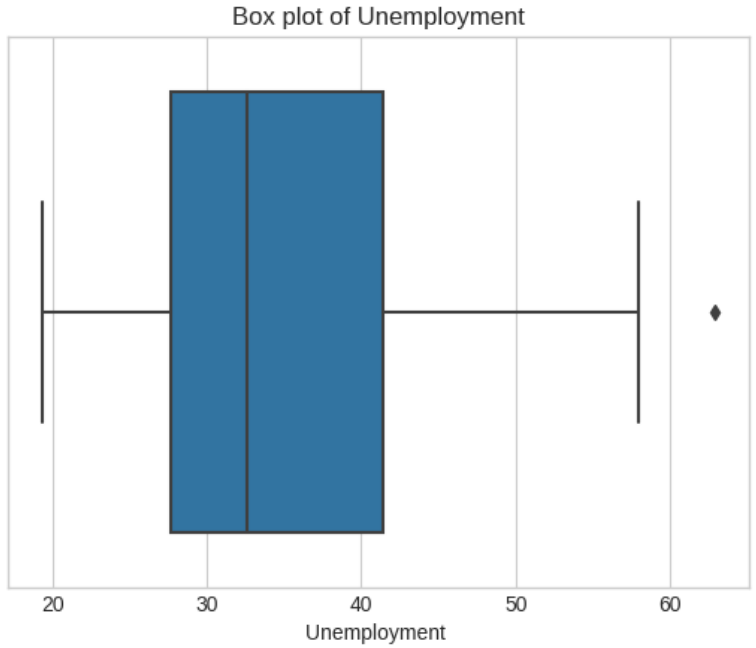
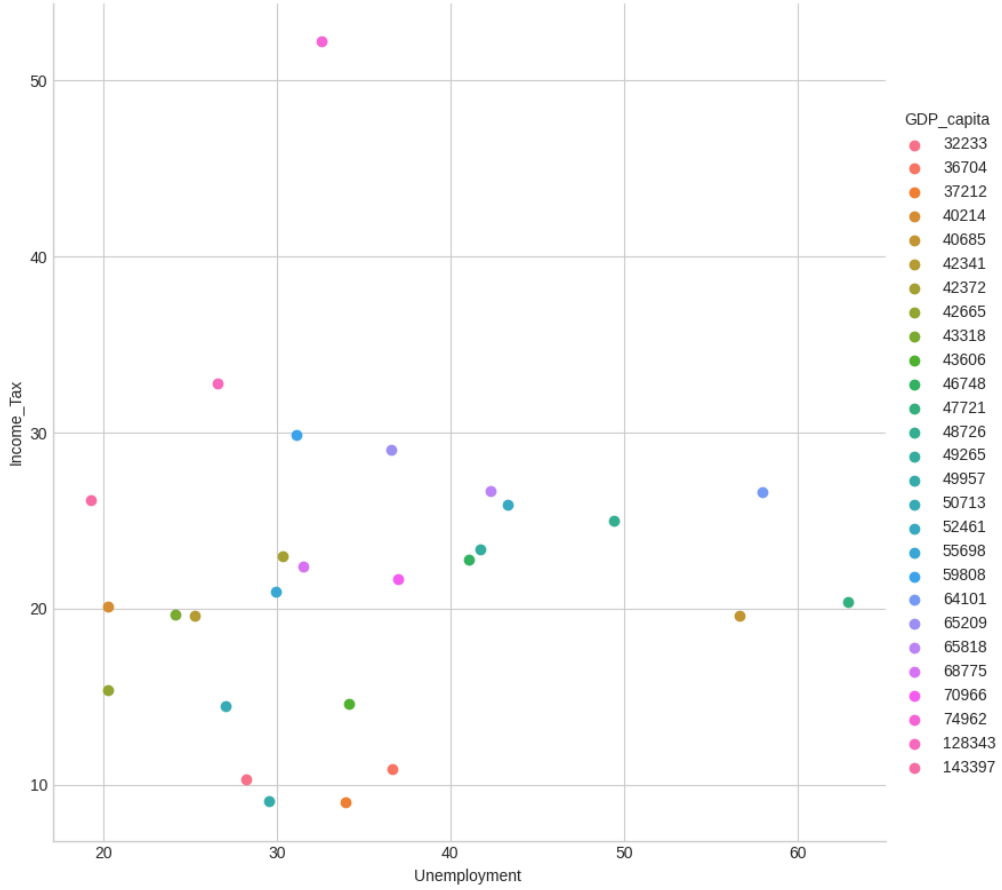


Table 4. Bi-Variate Analysis

a. GDP_capita



b. GDP Growth Rate

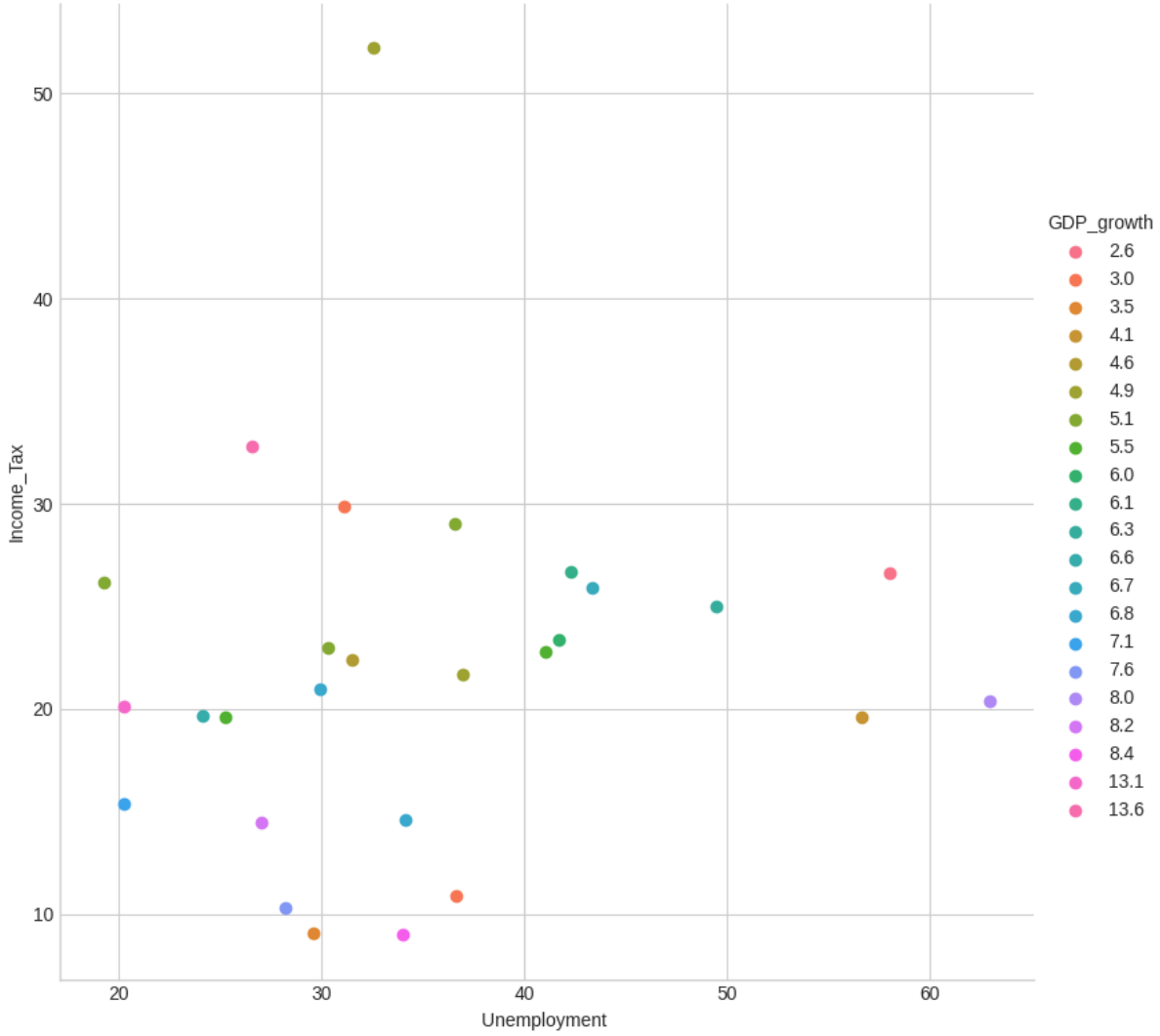
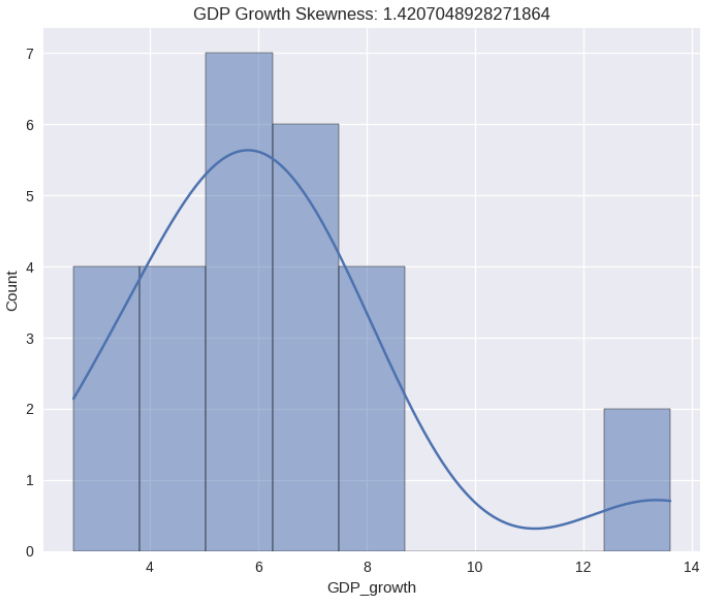
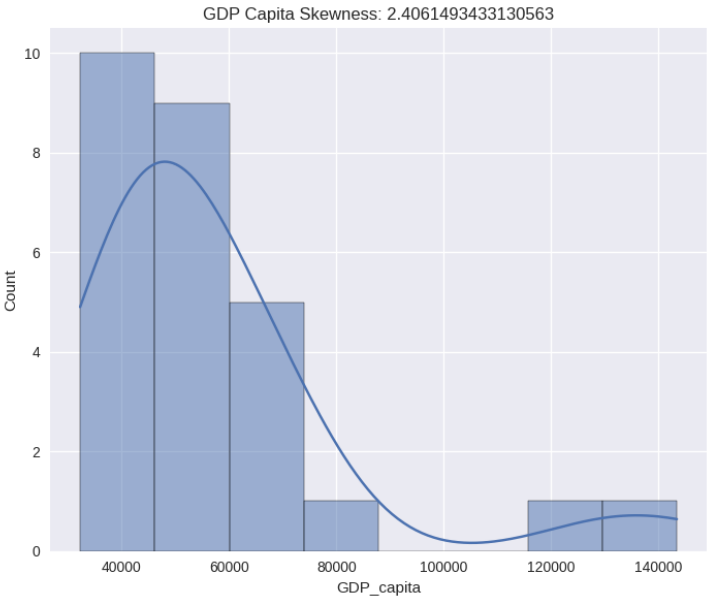


Table 5. Skewness of Dependent Variables


GDP Growth Skewness: 1.4207048928271864
Count
GDP_growth
0
1
2
3
4
5
6
7
4
6
8
10
12
14


GDP Capita Skewness: 2.4061493433130563
Count
GDP_capita
0
2
4
6
8
10
40000
60000
80000
100000
120000
140000