

**BEMM457** Topics in Business Analytics (A, TERM1 2022/3)**232089**

1082237



720073720

**Coursework:** Coursework**Submission Deadline:** Thu 22nd Dec 2022 12:00**Personal tutor:** Professor Stephen Disney**Marker name:** N/A**Word count:** 3001

By submitting coursework you declare that you understand and consent to the University policies regarding plagiarism and mitigation (these can be seen online at [www.exeter.ac.uk/plagiarism](http://www.exeter.ac.uk/plagiarism), and [www.exeter.ac.uk/mitigation](http://www.exeter.ac.uk/mitigation) respectively), and that you have read your school's rules for submission of written coursework, for example rules on maximum and minimum number of words. Indicative/first marks are provisional only.



# Economic Freedom Index Analysis

*Student Number- 720073720*

# Economic Freedom Index Analysis

## Using Correlation and Regression Analysis

1. Introduction.....	3
1.1 Preview of dataset.....	3
2. Missing Data.....	5
3. Data Cleaning.....	6
4. Descriptive Analysis.....	8
5. Exploratory Analysis	
5.1 How developed countries can help underdeveloped countries.....	10
5.2 General comparison of data for all countries.....	11
5.3 Top 3 variable developed nations should focus on .....	13
5.4 Top 3 variable developing nations should focus on.....	16
5.5 Top 3 variable underdeveloped nations should focus on.....	20
6. Relative Representation	
6.1 Using Correlation Heatmap .....	24
6.2 Using Regression Analysis .....	25
7. Data Ethics Framework and Principles.....	26
8. Conclusion.....	27
9. References.....	27

# 1. Introduction

I'm working on a modest data analytics project for the Economic Freedom Index, and I have taken this data from Kaggle. Economic freedom index is a composite measure of the quality of political-economic institution across different jurisdiction. Scores and ranks in an index are based on criteria that the creators of the index judge as being relevant, which vary from one index to another. The Economic Freedom Index is set to assist readers in tracking the economic freedom progress, profitability, and opportunity during the last two decades and promoting these standards in homes, schools and communities. These indexes are measured by the observations that economies that are more free market based tend to experience greater level of investments and higher average incomes. The Index examines 12 freedoms in 186 nations, ranging from property rights to financial independence. Economic freedom is measured using indicators that are classified into four major categories, or pillars, of economic freedom: 1. Legal Rule (property rights, government integrity, judicial effectiveness) 2. The Size of Government (government spending, tax burden, fiscal health) 3. Regulatory Efficiency (freedom of commerce, labour, and money) 4. Open Markets (trade, investment, and financial freedom).

## 1.1 Dataset Preview

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 186 entries, 0 to 185
Data columns (total 34 columns):
 #   Column              Non-Null Count  Dtype
---  -
 0   CountryID           186 non-null    int64
 1   Country Name        186 non-null    object
 2   WEBNAME             186 non-null    object
 3   Region              186 non-null    object
 4   World Rank          180 non-null    float64
 5   Region Rank         180 non-null    float64
 6   2019 Score          180 non-null    float64
 7   Property Rights     185 non-null    float64
 8   Judicial Effectiveness 185 non-null    float64
 9   Government Integrity 185 non-null    float64
10   Tax Burden          180 non-null    float64
11   Gov't Spending      183 non-null    float64
12   Fiscal Health       183 non-null    float64
13   Business Freedom    185 non-null    float64
```

14	Labor Freedom	184 non-null	float64
15	Monetary Freedom	184 non-null	float64
16	Trade Freedom	182 non-null	float64
17	Investment Freedom	184 non-null	float64
18	Financial Freedom	181 non-null	float64
19	Tariff Rate (%)	182 non-null	float64
20	Income Tax Rate (%)	183 non-null	float64
21	Corporate Tax Rate (%)	183 non-null	float64
22	Tax Burden % of GDP	179 non-null	float64
23	Gov't Expenditure % of GDP	182 non-null	float64
24	Country	186 non-null	object
25	Population (Millions)	186 non-null	object
26	GDP (Billions, PPP)	185 non-null	object
27	GDP Growth Rate (%)	184 non-null	float64
28	5 Year GDP Growth Rate (%)	183 non-null	float64
29	GDP per Capita (PPP)	184 non-null	object
30	Unemployment (%)	181 non-null	object
31	Inflation (%)	182 non-null	float64
32	FDI Inflow (Millions)	181 non-null	object
33	Public Debt (% of GDP)	182 non-null	float64

dtypes: float64(24), int64(1), object(9)  
memory usage: 49.5+ KB

Now, we should consider carefully which elements developed, emerging, and poor nations should prioritize in order to improve economic growth. The main aspects that we must examine are the questions, which we shall discuss in detail in.

## Questions

1. How developed countries can help underdeveloped countries?
2. Which of the top three variables should a developed country focus on?
3. Which of the top three variables should a developing country focus on?
4. Which of the top three variables should a developing country focus on?

## 2. Missing Data

We will first search for any missing values in the data set. If there will be any missing values then we have to clean our data first before starting analysis.

```
(df.isnull().sum())
```

CountryID	0
Country Name	0
WEBNAME	0
Region	0
World Rank	6
Region Rank	6
2019 Score	6
Property Rights	1
Judical Effectiveness	1
Government Integrity	1
Tax Burden	6
Gov't Spending	3
Fiscal Health	3
Business Freedom	1
Labor Freedom	2
Monetary Freedom	2
Trade Freedom	4
Investment Freedom	2
Financial Freedom	5
Tariff Rate (%)	4
Income Tax Rate (%)	3
Corporate Tax Rate (%)	3
Tax Burden % of GDP	7
Gov't Expenditure % of GDP	4
Country	0
Population (Millions)	0
GDP (Billions, PPP)	1
GDP Growth Rate (%)	2
5 Year GDP Growth Rate (%)	3
GDP per Capita (PPP)	2
Unemployment (%)	5
Inflation (%)	4
FDI Inflow (Millions)	5
Public Debt (% of GDP)	4
dtype: int64	

### 3. Data Cleaning

```
#step1= detecting N/A and na values  
missing_values=["N/a", "na", np.nan,]
```

```
df = pd.read_csv("economic_freedom_index2019_data.csv",na_values=missing_values)
```

```
df.isnull().sum()
```

```
df.isnull().any()
```

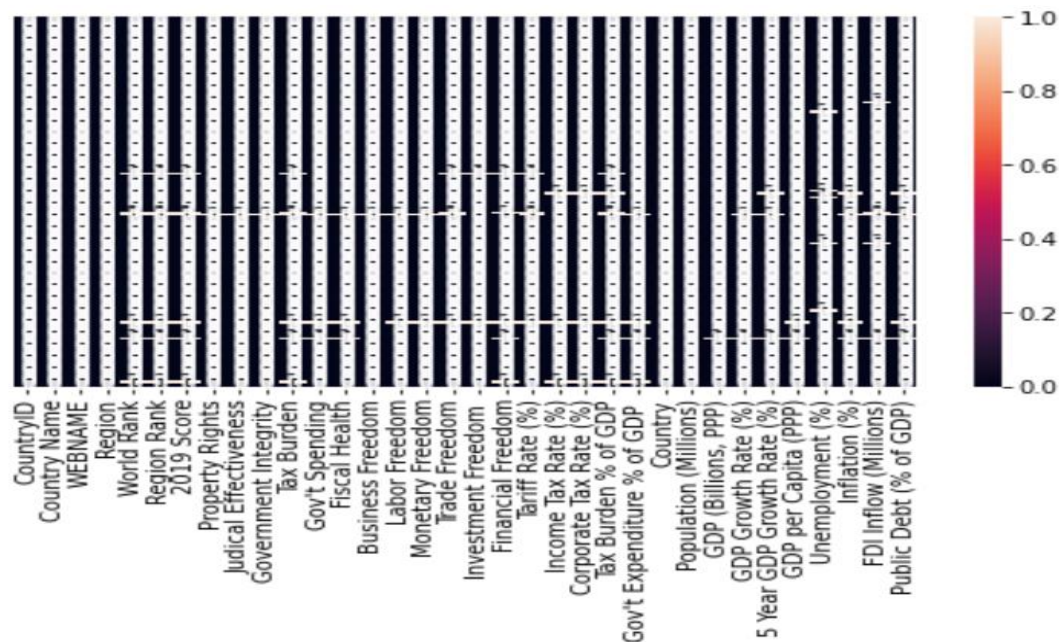
CountryID	False
Country Name	False
WEBNAME	False
Region	False
World Rank	True
Region Rank	True
2019 Score	True
Property Rights	True
Judical Effectiveness	True
Government Integrity	True
Tax Burden	True
Gov't Spending	True
Fiscal Health	True
Business Freedom	True
Labor Freedom	True
Monetary Freedom	True
Trade Freedom	True
Investment Freedom	True
Financial Freedom	True
Tariff Rate (%)	True
Income Tax Rate (%)	True
Corporate Tax Rate (%)	True
Tax Burden % of GDP	True

Gov't Expenditure % of GDP	True
Country	False
Population (Millions)	False
GDP (Billions, PPP)	True
GDP Growth Rate (%)	True
5 Year GDP Growth Rate (%)	True
GDP per Capita (PPP)	True
Unemployment (%)	True
Inflation (%)	True
FDI Inflow (Millions)	True
Public Debt (% of GDP)	True

dtype: bool



```
sns.heatmap(df.isnull(), yticklabels=False, annot=True)
```



```
df_dropped = df.dropna()
```

```
df_dropped
```

```
df_dropped1=df_dropped.interpolate()
```

```
df_dropped1
```

```
(df_dropped.isnull().sum())
```

```
CountryID      0
Country Name    0
WEBNAME         0
Region          0
World Rank      0
Region Rank     0
2019 Score      0
Property Rights 0
Judical Effectiveness 0
Government Integrity 0
Tax Burden      0
Gov't Spending  0
Fiscal Health   0
Business Freedom 0
Labor Freedom   0
Monetary Freedom 0
Trade Freedom   0
Investment Freedom 0
Financial Freedom 0
Tariff Rate (%) 0
Income Tax Rate (%) 0
Corporate Tax Rate (%) 0
Tax Burden % of GDP 0
```

```
Gov't Expenditure % of GDP  0
Country                     0
Population (Millions)       0
GDP (Billions, PPP)         0
GDP Growth Rate (%)         0
5 Year GDP Growth Rate (%)  0
GDP per Capita (PPP)        0
Unemployment (%)            0
Inflation (%)               0
FDI Inflow (Millions)       0
Public Debt (% of GDP)      0
dtype: int64
```

My dataset had several issues that I needed to address first. I changed certain string values to float so that I could perform the analysis. So I changed GDP per capita(ppp), Unemployment, FDI Inflow and Population to float.

```
df['GDP per Capita (PPP)'] = df['GDP per Capita (PPP)'].str.replace('$', '')
df['GDP per Capita (PPP)'] = df['GDP per Capita (PPP)'].str.replace(',', '')
df['GDP per Capita (PPP)'] = df['GDP per Capita (PPP)'].str.split(' ').str[0]
df['GDP per Capita (PPP)'] = df['GDP per Capita (PPP)'].astype(float)
```

```
df['Unemployment (%)'] = df['Unemployment (%)'].str.replace(',', '')
df['Unemployment (%)'] = df['Unemployment (%)'].str.split(' ').str[0]
df['Unemployment (%)'] = df['Unemployment (%)'].astype(float)
```

```
df["FDI Inflow (Millions)"] = df["FDI Inflow (Millions)"].str.replace(',', '')
df["FDI Inflow (Millions)"] = df["FDI Inflow (Millions)"].str.split(' ').str[0]
df["FDI Inflow (Millions)"] = df["FDI Inflow (Millions)"].astype(float)
```

```
df["Population (Millions)"] = df["Population (Millions)"].str.replace(',', '')
df["Population (Millions)"] = df["Population (Millions)"].str.split(' ').str[0]
df["Population (Millions)"] = df["Population (Millions)"].astype(float)
```

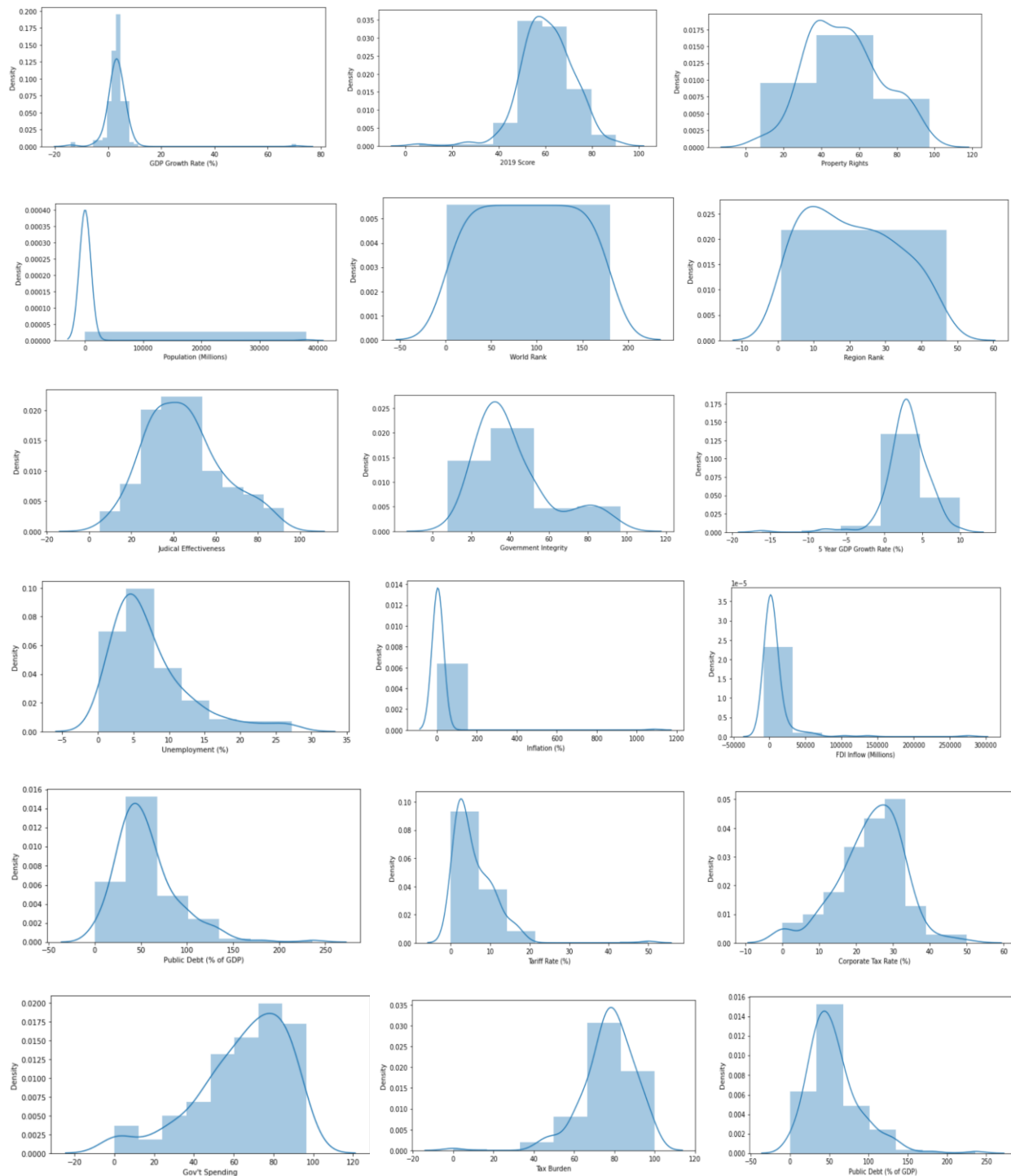
## 4. Descriptive analysis

In this part, we look at the data set's composition, the number of columns and rows, and the data structure. To begin, I utilize Python's pandas-profiling tool to produce a descriptive analysis report. The following indicators to be visualized and discussed are as following:

	CountryID	World Rank	Region Rank	2019 Score	Property Rights	Judical Effectiveness	Government Integrity	Tax Burden	Gov't Spending
count	186.000000	180.000000	180.000000	180.000000	185.000000	185.000000	185.000000	180.000000	183.000000
mean	93.500000	90.500000	20.538889	60.768333	52.327568	44.899459	41.470270	77.212778	64.203825
std	53.837719	52.105662	12.738611	11.255725	19.608526	18.104745	19.793193	13.208314	23.150984
min	1.000000	1.000000	1.000000	5.900000	7.600000	5.000000	7.900000	0.000000	0.000000
25%	47.250000	45.750000	9.750000	53.950000	37.000000	31.000000	27.200000	70.975000	51.700000
50%	93.500000	90.500000	19.500000	60.750000	50.100000	42.900000	35.500000	78.050000	68.800000
75%	139.750000	135.250000	31.000000	67.800000	65.900000	54.700000	50.300000	85.425000	82.600000
max	186.000000	180.000000	47.000000	90.200000	97.400000	92.400000	96.700000	99.800000	96.600000

8 rows × 25 columns

The above result was generated using Python and contains the count of each variable, the mean of all its values, the standard deviation between each variable's values, the minimum value in each variable, the 25%, 50% (mode), 75%, and maximum value in each variable.



**My dependant variable is Gdp per capita .**

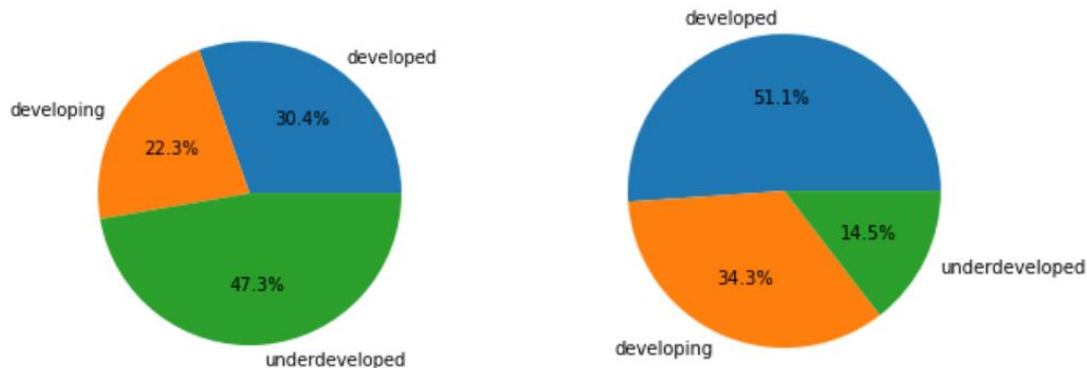
## **5. Exploratory Analysis**

In the study, I plan to give my insights on how many elements impact a country's economic growth. It will comprise elements that are critical to an economy's progress.

### **5.1. How developed countries can help underdeveloped countries?**

According to the pie chart in figure 1, Out of all , 22.3% of nations are developing, 30.4% are developed, and 47.3 are undeveloped. Figure 2 demonstrates that:

1. developed nations' GDP is 51.1%,
2. emerging countries' GDP is 34.3%,
3. The GDP of developing nations is 14.5%.



**Figure1.**

**Figure 2.**

Now, how can a developed country assist a developing country?

There are numerous methods that rich countries may assist developing and underdeveloped countries in improving their GDP. The industrialized nations might provide funding to the establishment of new schools and polytechnic institutes. These will not only improve literacy, but will also give vocational education. The primary components through which a developing nation can be developed by a developed country are as follows:

- Free trade may help a country expand its market and become a standing elephant in the global commerce competition. There are many countries who produce their own commodities, and many countries that require them in order to function.
- When a country has enough weaponry, aircraft, and ships, it is safe. When a country is created with all of the capabilities of its adjacent country, it might be less afraid of the other.

- Highly skilled and experienced economists are responsible for the development of a developed country. Their government's cabinet and advisors are carefully chosen. Those economists can transform a country in a few of years.
- Offering low-interest loans or extending credit to underdeveloped countries will be helpful.

## 5.2. General comparison of data for all countries

```
d = df.corr()
d['GDP per Capita (PPP)'].sort_values().head(6)
```

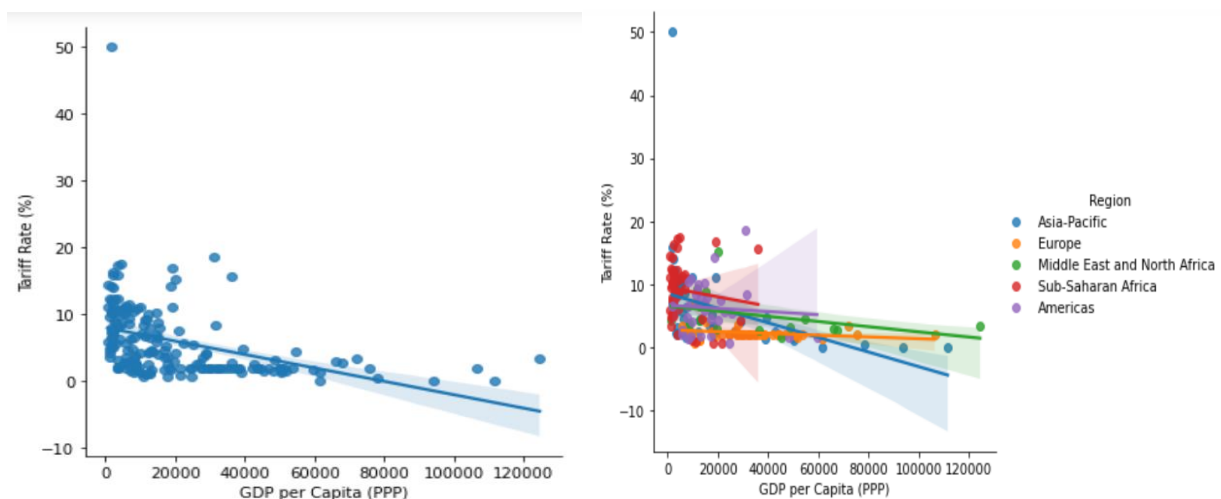
```
World Rank      -0.640095
Region Rank     -0.454021
Tariff Rate (%) -0.410434
Corporate Tax Rate (%) -0.347868
Gov't Spending  -0.255897
Unemployment (%) -0.183008
Name: GDP per Capita (PPP), dtype: float64
```

We discovered that Tariff Rate (%), Corporate Tax Rate and Government Spending have the least association with GDP per capita.

### 1. Tariff Rate and GDP per capita

```
#analysing the variable who has Least correlation with gdp per capita
sns.lmplot(data=df, x='GDP per Capita (PPP)', y='Tariff Rate (%)')
```

```
sns.lmplot(data=df, x='GDP per Capita (PPP)', y='Tariff Rate (%)', hue='Region')
```



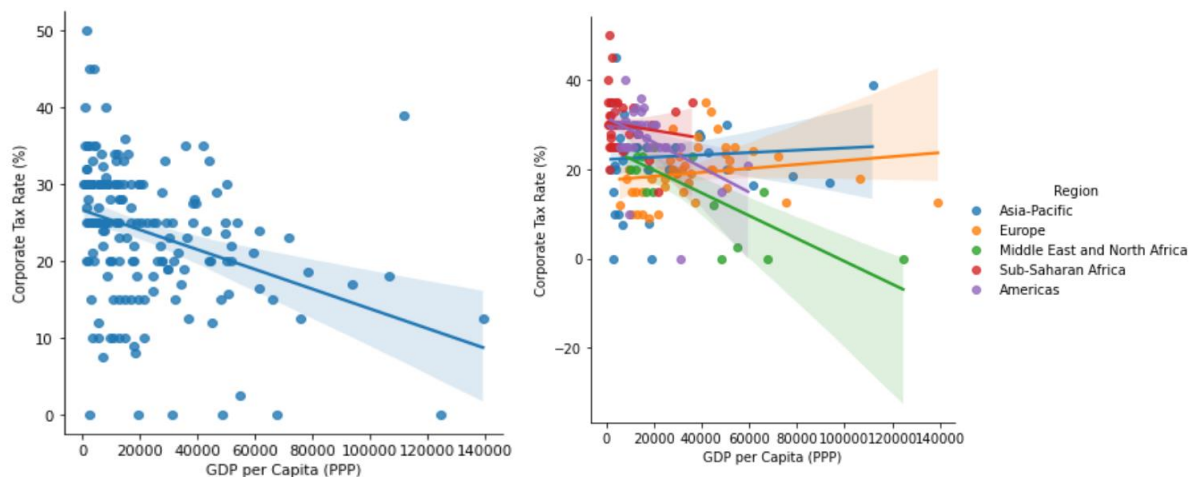
The graph clearly shows a negative relationship between GDP per capita and tariff rate for all regions. This is because when a country's GDP is high, the government lowers import taxes. One

of the key causes of the fall has been the establishment of international organizations aimed at improving free trade, such as the World Trade Organization (WTO). Tariff reduction reduces the "loss of efficiency" expenses caused by the price system distortions caused by tariffs, which leads to economic growth.

## 2 Corporate tax and GDP per capita

```
sns.lmplot(data=df, x='GDP per Capita (PPP)', y='Corporate Tax Rate (%)')
```

```
sns.lmplot(data=df, x='GDP per Capita (PPP)', y='Corporate Tax Rate (%)', hue='Region')
```

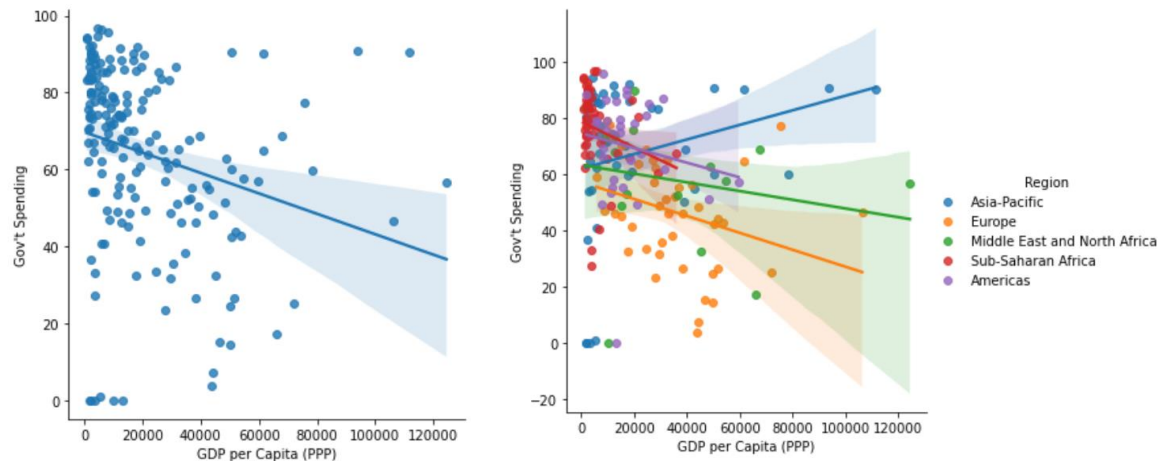


According to the findings, there is a negative association between GDP and corporate tax rates in the Americas, the Middle East and North Africa, and the Sub-Saharan region. A reduction in corporate tax will improve businesses' post-tax profitability. This will boost the amount of money available to fund capital investments such as new plants and factories, which might contribute to an increase in the country's productive capacity. However, alternative scenarios are possible. GDP and corporation taxation can also have a favourable relationship like in Asia-Pacific and Europe region. If the government raises corporate profit tax rates, the government may also contribute to economic growth through investment.

## 3 Government spending and GDP

```
sns.lmplot(data=df, x='GDP per Capita (PPP)', y='Gov't Spending')
```

```
sns.lmplot(data=df, x='GDP per Capita (PPP)', y='Gov't Spending', hue='Region')
```



The graph clearly shows that GDP is strongly tied to government spending in the Asia-Pacific region. This is because increased government spending is expected to boost aggregate demand (AD). If infrastructure spending is prioritized, it may result in greater productivity and long-run aggregate supply growth. However, there is a negative relationship between GDP and government spending in other regions. In the government sector, there may be a lack of knowledge and incentives, resulting in resource misallocation.

### 5.3. Top 3 variable a developed nation should focus on

In this analysis we will find out the 3 variable on which a developed nation should focus on and 3 variable on which developed countries need to improve for the economic growth.

*#Analysis for developed countries*

```
developed = df[df['GDP per Capita (PPP)']>=25000]
developed
```

```
d= developed.corr()
d['GDP per Capita (PPP)'].sort_values().head(6)
```

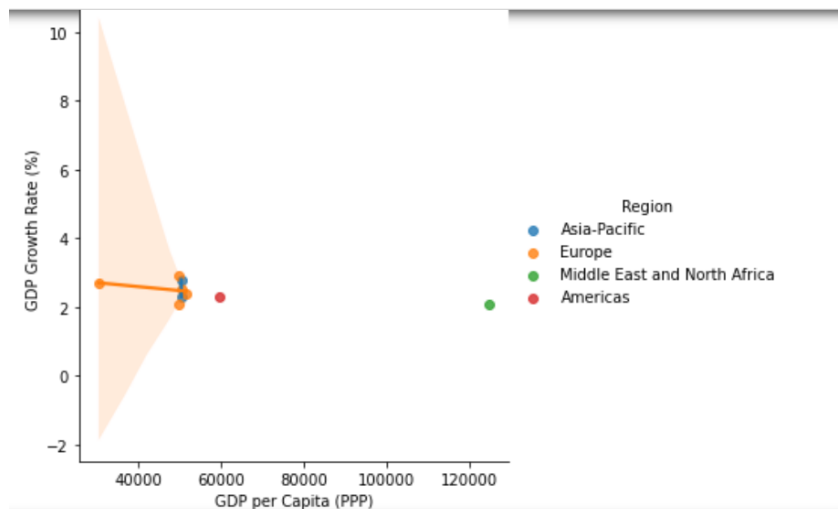
```
Region Rank          -0.435511
Unemployment (%)     -0.420893
Inflation (%)        -0.349710
World Rank           -0.348981
Tariff Rate (%)      -0.248460
Corporate Tax Rate (%) -0.244632
Name: GDP per Capita (PPP), dtype: float64
```

We discovered that the variables are least connected with GDP using this code, indicating that there is room for improvement in these areas.

```
d1=developed[(developed['GDP Growth Rate (%)']>2) &(developed['GDP Growth Rate (%)']<3)]
```

```
sns.lmplot(data=d1, x='GDP per Capita (PPP)',y= 'GDP Growth Rate (%)',hue='Region')
```





The growth rate of developed countries should be between 2% to 3%. In industrialised countries, there are just four regions whose growth rate is between 2% to 3%. Capital, labour, and technology all have a role in economic progress. Increased investment rate, the job of the financial market (such as banks) is to steer savers' assets towards suitable enterprises, Property rights are well protected by the legal system in industrialised countries. Political stability reduces investment risk while increasing international investment and economic growth.

**Now, we will find the 3 variable on which developed countries need to focus for the economic growth**

```
d['GDP per Capita (PPP)'].sort_values().tail()
```

```
2019 Score          0.382376
Judicial Effectiveness 0.405432
Government Integrity  0.411290
Population (Millions) 0.487994
GDP per Capita (PPP)  1.000000
Name: GDP per Capita (PPP), dtype: float64
```

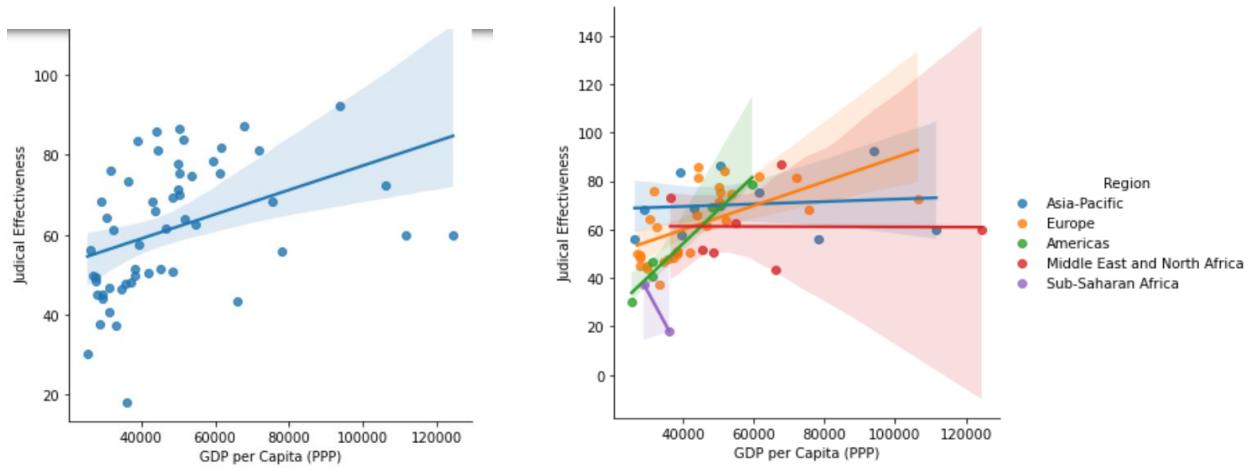
We discovered that the variables (Judicial Effectiveness, Government Integrity, Population are the most closely associated to GDP, hence there is a need to focus on these factors for further progress in economic growth.

## 1. GDP and Judicial Effectiveness

```
sns.lmplot(data=developed, x='GDP per Capita (PPP)', y="Judicial Effectiveness")
```



```
sns.lmplot(data=developed, x='GDP per Capita (PPP)', y="Judicial Effectiveness", hue ='Region')
```

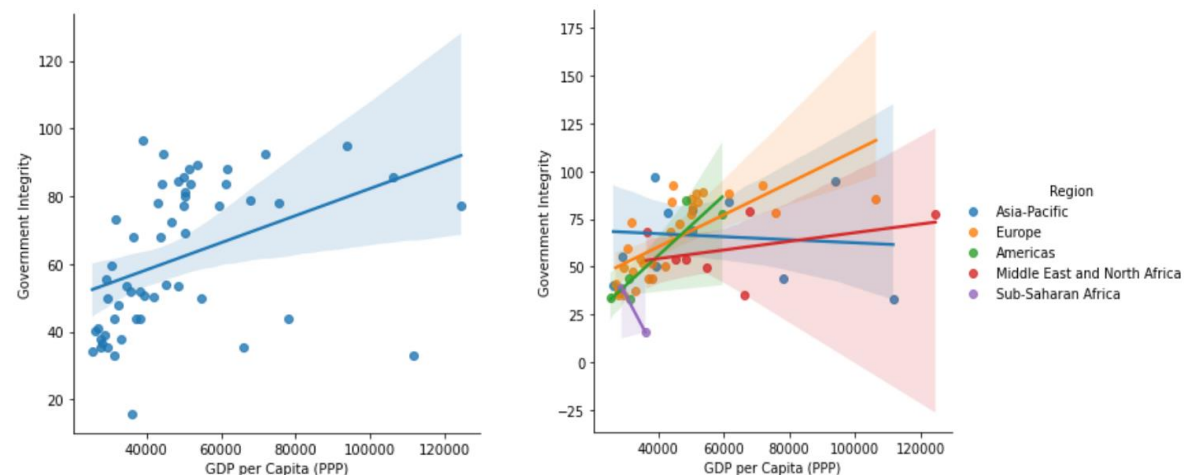


Except for the Middle East and North Africa, there is a positive correlation between GDP and Judicial effectiveness. My findings corroborate that inefficiencies in the operation of judicial systems impede economic growth, and hence beneficial advancements in judicial efficiency can be growth promoting. According to data, America, Asia, and Europe have a good relationship with GDP and judicial effectiveness, but Sub-Saharan Africa has a lower growth rate since it is not focusing on judicial effectiveness.

## 2 GDP per capita and Government integrity

```
sns.lmplot(data=developed, x='GDP per Capita (PPP)', y="Government Integrity")
```

```
sns.lmplot(data=developed, x='GDP per Capita (PPP)', y="Government Integrity", hue ='Region')
```

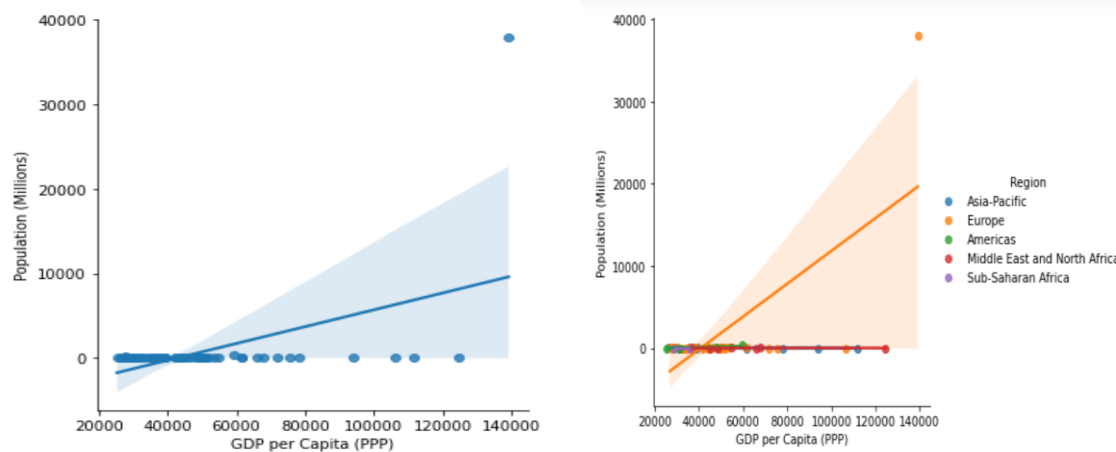


With the exception of the Asia Pacific and Sub-Saharan regions, there is a positive association between GDP per capita and government integrity. According to study, corporate leaders' decisions and preferences regarding corporate governance are positively correlated with governmental integrity. In particular, we find that corporate bribery and corruption is low in transparent environments and high in contexts with weak government integrity.

### 3. GDP per capita and population

```
sns.lmplot(data=developed, x='GDP per Capita (PPP)', y="Population (Millions)")

sns.lmplot(data=developed, x='GDP per Capita (PPP)', y="Population (Millions)", hue='Region')
```



GDP per capita and population expectations are positively correlated for the Asia-Pacific, Middle East, and North Africa regions. Population growth raises GDP because more people are working, which happens when there is population expansion. Despite the fact that the population is declining, the GDP is still rising due to additional factors like education level, standard of living, healthcare costs, and other socioeconomic and environmental aspects that can also influence the expansion of the economy.

## 5.4. Top 3 variable a developing nation should focus on

In this analysis we will find out the 3 variable on which a developing nation should focus on and 3 variable on which developing countries need to improve for the economic growth.

```
#Analysis for developing countries
developing = df[(df["GDP per Capita (PPP)"]>12000) & (df["GDP per Capita (PPP)"] <25000)]
```

```
d= developing.corr()
d['GDP per Capita (PPP)'].sort_values().head(6)
```

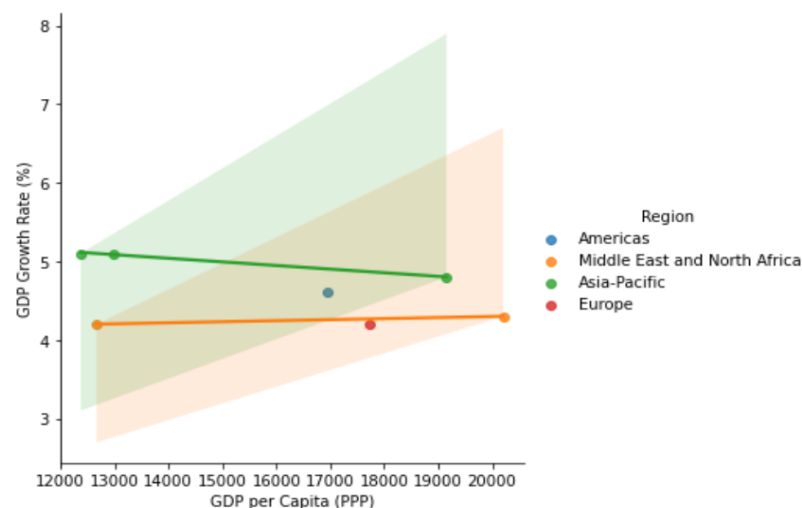
```
Unemployment (%)      -0.249920
Corporate Tax Rate (%) -0.245369
World Rank             -0.241097
Inflation (%)          -0.216727
CountryID              -0.091114
Region Rank            -0.073774
Name: GDP per Capita (PPP), dtype: float64
```

We discovered that the variables are least connected with GDP using this code, indicating that there is room for improvement in these areas.

```
d2 = developing[(developing['GDP Growth Rate (%)']>4) &(developing['GDP Growth Rate (%)']<6)]
sns.lmplot(data=d2, x='GDP per Capita (PPP)',y= 'GDP Growth Rate (%)',hue='Region')
```

```
e=developing[developing["Region"]=="Europe"]
e
e[['GDP per Capita (PPP)', "GDP Growth Rate (%)"]].corr()
```

	GDP per Capita (PPP)	GDP Growth Rate (%)
GDP per Capita (PPP)	1.000000	0.496249
GDP Growth Rate (%)	0.496249	1.000000



There are only four regions with the required growth rate for developing nations, which is between 4% and 6%. Here, we discover an association between GDP and growth rate in developing nations. Only seven developing nations, according to the graph, are actually performing well in terms of economic growth. The Middle East, North Africa, and Asia-Pacific region's GDP and growth rates are found to be negatively correlated through the aforementioned code.

**Now, we will find the 3 variable on which developing countries need to focus for the economic growth**

```
d= developing.corr()
d['GDP per Capita (PPP)'].sort_values().tail(6)
```

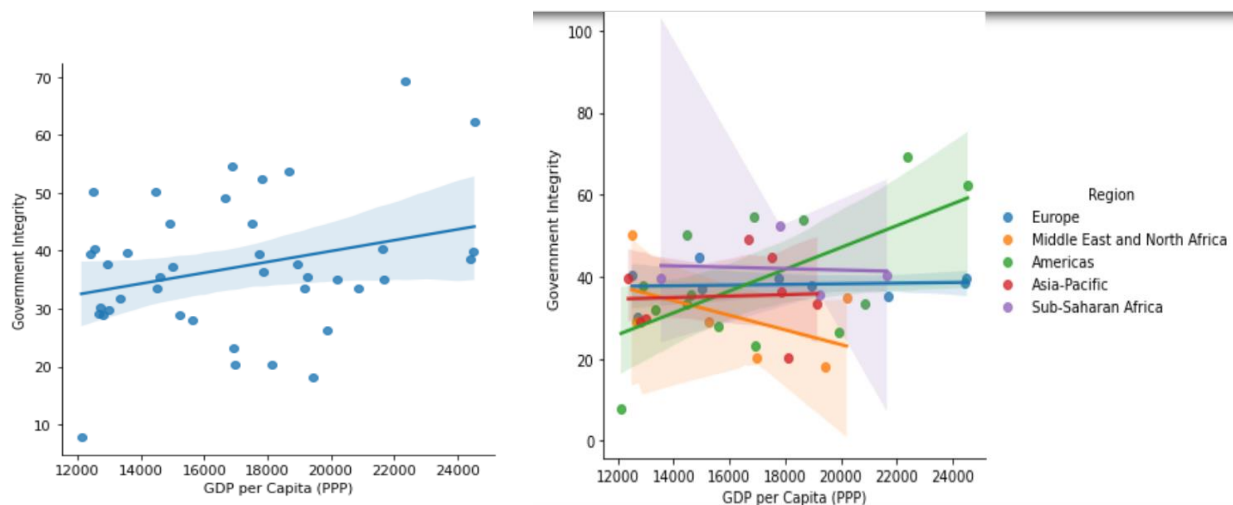
```
Government Integrity    0.285403
Investment Freedom      0.297429
Fiscal Health           0.298742
2019 Score              0.324557
Property Rights         0.379709
GDP per Capita (PPP)    1.000000
Name: GDP per Capita (PPP), dtype: float64
```

We discovered that the variables Government Integrity, Property Rights and Fiscal Health are the most closely associated to GDP, hence there is a need to focus on these factors for further progress in economic growth.

## 1.GDP and Government Integrity

```
sns.lmplot(data=developing, x='GDP per Capita (PPP)',y="Government Integrity")
```

```
sns.lmplot(data=developing, x='GDP per Capita (PPP)',y="Government Integrity", hue ='Region')
```



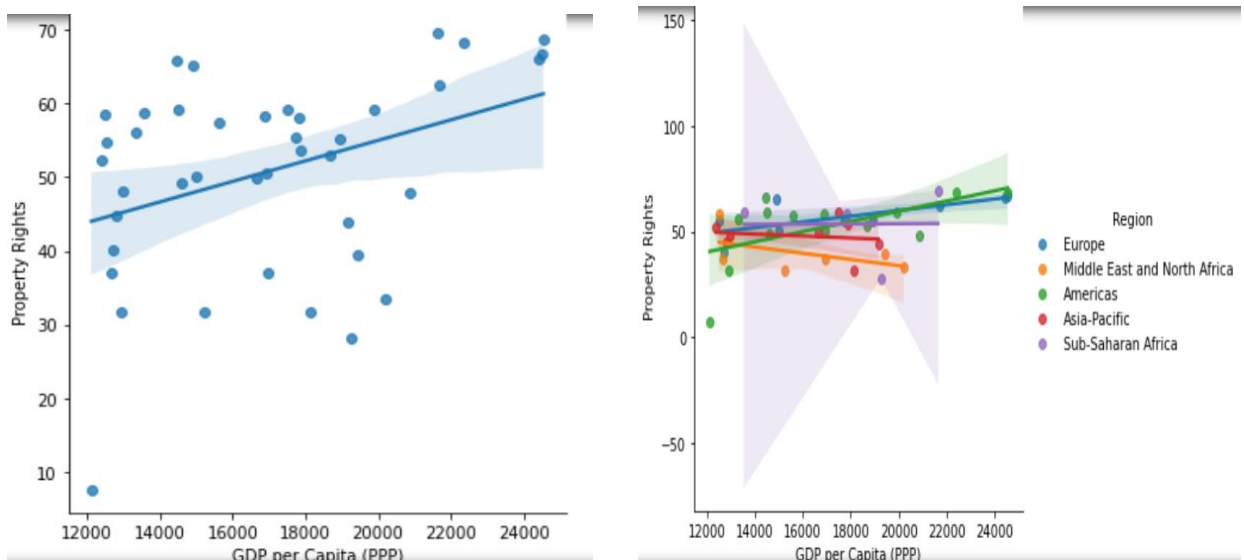
```
e=developing[developing["Region"]=="Sub-Saharan Africa"]
e
e[['GDP per Capita (PPP)', "Government Integrity"]].corr()
```

From the analysis it is found that there is a positive correlation between GDP and governance integrity In all regions except Sub-Saharan Africa and Middle East and North Africa. Higher Government Integrity reduces underinvestment which in turn leads to economic growth. Middle East and North Africa and Sub-Saharan Africa's economic growth is not good because they are not good in government integrity.

## 2.GDP per capita and Property Rights

```
sns.lmplot(data=developing, x='GDP per Capita (PPP)',y="Property Rights")
```

```
sns.lmplot(data=developing,x='GDP per Capita (PPP)',y='Property Rights',hue ='Region')
```



```
e=developing[developing["Region"]=="Sub-Saharan Africa"]
e
e[['GDP per Capita (PPP)', "Property Rights"]].corr()
```

	GDP per Capita (PPP)	Property Rights
GDP per Capita (PPP)	1.000000	0.010754
Property Rights	0.010754	1.000000

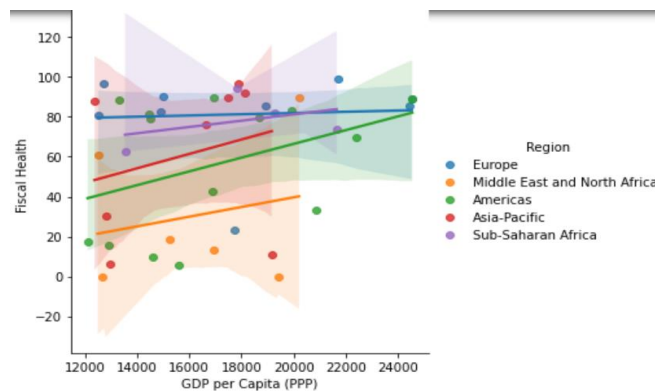
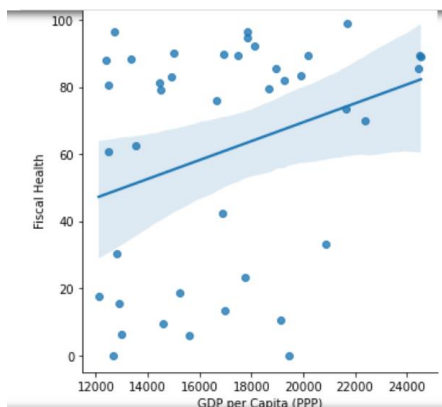
From the analysis, it is discovered that there is a positive correlation between GDP and Property rights in all regions except Middle East and North Africa and Asia-Pacific. If Property rights is increasing to a good level then GDP is also increasing.

### 3. GDP and Fiscal health

```
sns.lmplot(data=developing, x='GDP per Capita (PPP)',y="Fiscal Health")
```

```
sns.lmplot(data=developing, x='GDP per Capita (PPP)',y="Fiscal Health", hue ='Region')
```

```
e=developing[developing["Region"]=="Europe"]  
e  
e[['GDP per Capita (PPP)', "Fiscal Health"]].corr()
```



From the analysis it is discovered that there is a positive relation between GDP and Fiscal health in all the regions. It means that as fiscal health is improving GDP of developing country is also improving. Financial health of a country is improving, and this is why the economic growth of developing countries is also improving.

### 5.5. Top 3 Variable Underdeveloped nations should focus on

In this analysis we will find out the 3 variable on which a underdeveloped nation should focus on for the economic growth.

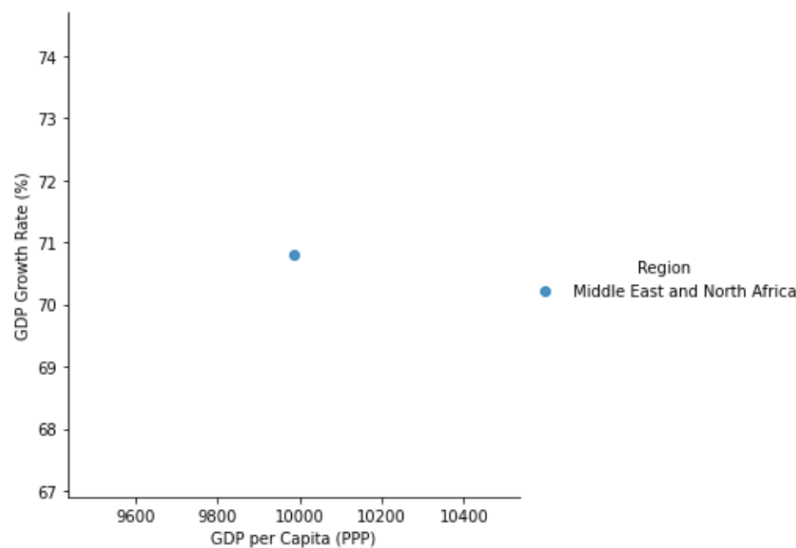
```
#analysis for underdeveloped countries  
underdeveloped = df[df['GDP per Capita (PPP)']<=12000]  
underdeveloped
```

```
d= underdeveloped.corr()
d['GDP per Capita (PPP)'].sort_values().head(6)
```

```
World Rank          -0.501351
Region Rank         -0.355148
Tariff Rate (%)     -0.320556
Corporate Tax Rate (%) -0.227763
Income Tax Rate (%) -0.186768
Inflation (%)       -0.127011
Name: GDP per Capita (PPP), dtype: float64
```

We discovered that the variables are least connected with GDP using this code, indicating that there is room for improvement in these areas.

```
d=underdeveloped[(underdeveloped['GDP Growth Rate (%)']>=15)]
sns.lmplot(data=d, x='GDP per Capita (PPP)',y= 'GDP Growth Rate (%)',hue='Region')
```



According to the above Graph, there is only 1 country from Middle East and North Africa region who has a growth rate of 71. The required growth rate for underdeveloped country is 15%.

**Now, we will find the 3 variable on which Underdeveloped countries need to focus for the economic growth**

```
e= underdeveloped.corr()
e['GDP per Capita (PPP)'].sort_values().tail(6)
```

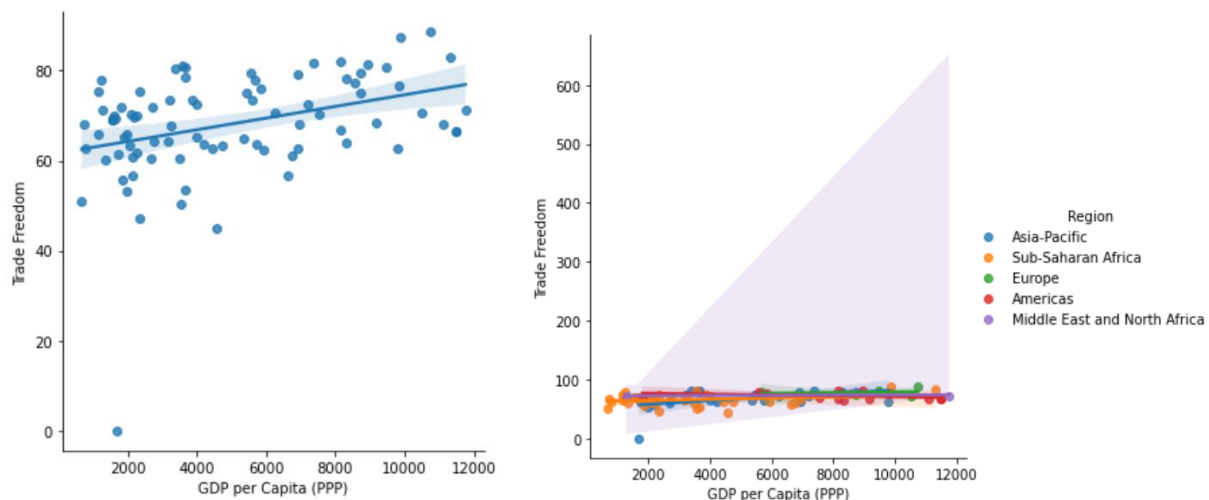
```
Trade Freedom      0.360177
Property Rights    0.384079
Tax Burden % of GDP 0.395292
2019 Score         0.416948
Business Freedom    0.552840
GDP per Capita (PPP) 1.000000
Name: GDP per Capita (PPP), dtype: float64
```

We discovered that the variables Trade Freedom, Property Rights and Tax Burden are the most closely associated to GDP, hence there is a need to focus on these factors for further progress in economic growth.

## 1. GDP per capita and Trade Freedom

```
sns.lmplot(data=underdeveloped, x='GDP per Capita (PPP)',y="Trade Freedom")
```

```
sns.lmplot(data=underdeveloped, x='GDP per Capita (PPP)',y= 'Trade Freedom',hue='Region')
```



```
e=underdeveloped[underdeveloped["Region"]=="Middle East and North Africa"]
e
e[['GDP per Capita (PPP)', "Trade Freedom"]].corr()
```

From the analysis it is discovered that there is a positive correlation between growth of gdp per capita and trade freedom in all region except America. Trade freedom enables people to concentrate on what they do best and then trade the result of their own areas of excellence, Freedom rate is deteriorating, which is why its economy isn't expanding at a good rate.

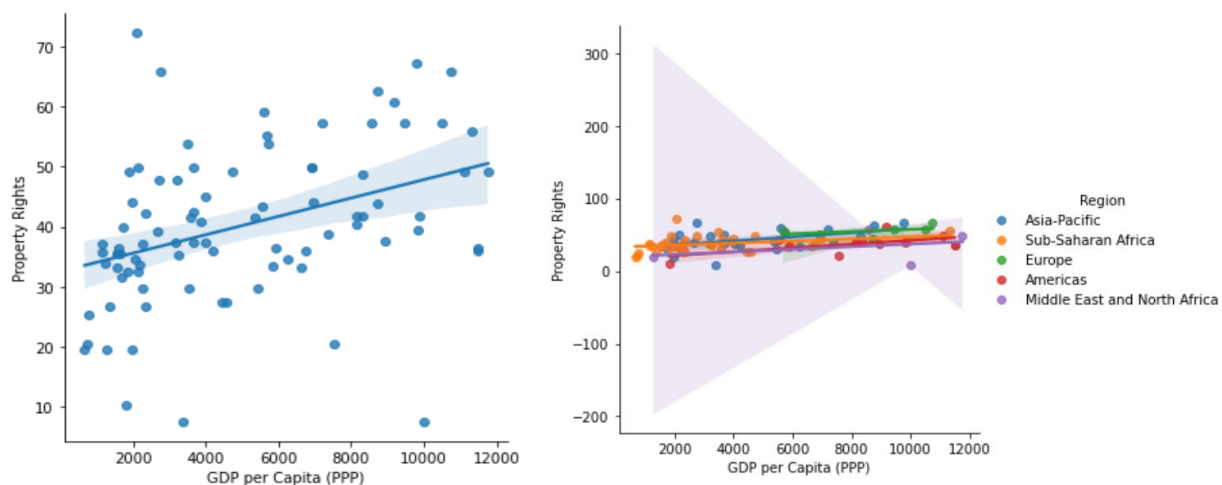


Overcoming issues such as cultural differences, language and health and safety regulation can help America to improve its Trade freedom.

## 2.GDP per capita and Property rights

```
sns.lmplot(data=underdeveloped, x='GDP per Capita (PPP)',y="Property Rights")
```

```
sns.lmplot(data=underdeveloped, x='GDP per Capita (PPP)',y= 'Property Rights',hue='Region')
```



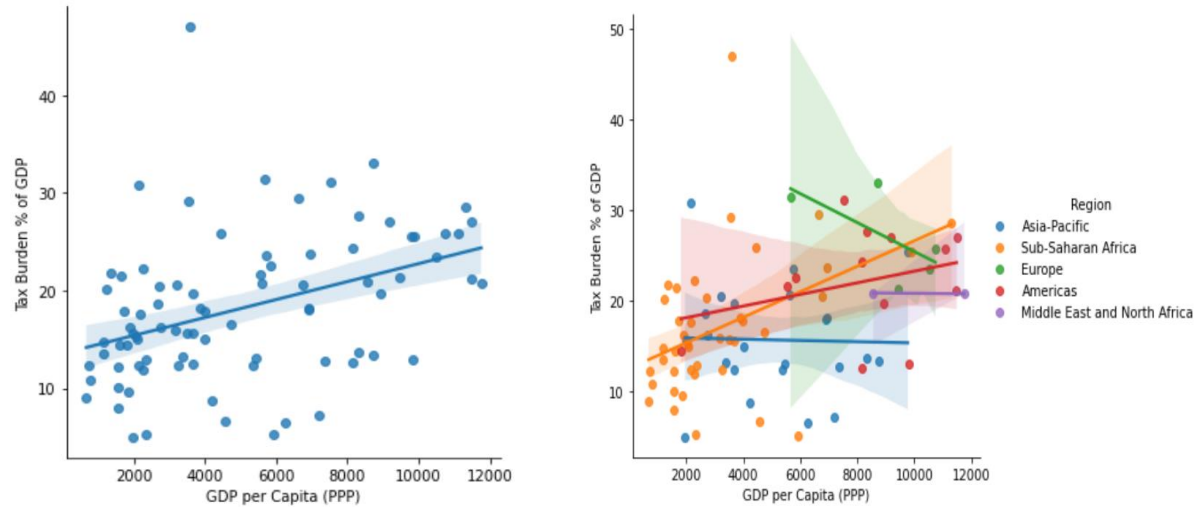
```
e=underdeveloped[underdeveloped["Region"]=="Asia-Pacific"]  
e  
e[['GDP per Capita (PPP)',"Property Rights"]].corr()
```

From my analysis it is discovered that the countries in all the region they are doing good in protecting the property rights as there is a positive correlation between growth of GDP per capita and Property right in all region. Protection of property right is a major barrier to corruption which helps in economy growth.

## 3.GDP per capita and Tax burden

```
sns.lmplot(data=underdeveloped, x='GDP per Capita (PPP)',y="Tax Burden % of GDP")
```

```
sns.lmplot(data=underdeveloped, x='GDP per Capita (PPP)',y= 'Tax Burden % of GDP',hue='Region')
```



```
e=underdeveloped[underdeveloped["Region"]=="Middle East and North Africa"]
e
e[['GDP per Capita (PPP)', "Tax Burden % of GDP"]].corr()
```

Well, as per the graph there is a negative correlation between the growth of GDP and tax burden for Europe and Asia-Pacific. From the analysis it is discovered that tax burden plays a very important role in finding the economy health of a country. As tax burden is increasing, the economy is not expanding to a satisfactory level as government is not able to pay taxes. However, Europe and Asia-Pacific region excelling at reducing tax burden.

## 6.Relative Representation

### Using correlation

I will create a heatmap using correlation analysis to establish a correlation for each feature with each other based on their coefficients, show a clear relationship between them, and solidify the correlation in this dataset. This will help us create a solution with all the quantitative components.

```
heat=df.drop(['CountryID', 'Country Name', 'WEBNAME', 'Region', 'World Rank',
              'Region Rank', '2019 Score'],axis=1).corr()
```

```
sns.heatmap(heat,vmax=1,vmin=-1,linewidths=1,linecolor='Black')
```



When the correlation coefficient is close to 0, the dependent variable is said to be weakly correlated, and when it is close to 1, the correlation coefficient is said to be strongly correlated. The degree of movement between two variables is indicated by correlation. We can see from the visualisation that property rights have weak correlations with tariff rates but very strong correlations with judicial effectiveness and government integrity.

Furthermore, some features significantly correlate with one another, such as Business Freedom and Judicial Effectiveness (correlation coefficient: 0.75), while others have a negative correlation, such as the Tariff rate and Business Freedom (correlation coefficient: -0.75). I was able to conduct a thorough investigation and identify key elements that have a significant impact on the house's price thanks to this heat map correlation.

## Using Regression

Regression analysis is a statistical technique that lets us use a dataset to determine the relationship between dependent and independent variables. This is done by testing the null hypothesis and determining the relationship between the variables using the regression coefficient  $r$  and  $p$ -value. The  $p$ -value and regression coefficient both show how strongly a relationship is related.

I've provided a table below that shows the regression table that I found using Microsoft Excel.

SUMMARY OUTPUT									
<b>Regression Statistics</b>									
Multiple	0.771507								
R Square	0.595224								
Adjusted	0.556902								
Standard	15875.4								
Observat	186								
<b>ANOVA</b>									
	<b>df</b>	<b>SS</b>	<b>MS</b>	<b>F</b>	<b>Significance F</b>				
Regression	16	62632757110	3914547319	15.53217	1.52222E-25				
Residual	169	42592805217	252028433						
Total	185	1.05226E+11							
	<b>Coefficients</b>	<b>Standard Error</b>	<b>t Stat</b>	<b>P-value</b>	<b>Lower 95%</b>	<b>Upper 95%</b>	<b>Lower 95.0%</b>	<b>Upper 95.0%</b>	
Intercept	61312.33	11870.38013	5.16515288	6.7E-07	37879.00537	84745.65	37879.01	84745.65	
X Variabl	142.8231	148.0530798	0.9646747	0.336086	-149.4485905	435.0947	-149.449	435.0947	
X Variabl	46.64913	150.3518989	0.31026631	0.756741	-250.1606221	343.4589	-250.161	343.4589	
X Variabl	573.7807	157.3946887	3.64548983	0.000355	263.067811	884.4937	263.0678	884.4937	
X Variabl	-100.565	110.0666141	-0.9136741	0.36219	-317.8475669	116.7175	-317.848	116.7175	
X Variabl	-106.396	76.52492001	-1.3903459	0.166252	-257.46399	44.67177	-257.464	44.67177	
X Variabl	28.41561	41.61260655	0.68286051	0.49563	-53.73185958	110.5631	-53.7319	110.5631	
X Variabl	-218.809	134.4660022	-1.6272413	0.105549	-484.2580335	46.64076	-484.258	46.64076	
X Variabl	-115.009	102.7639682	-1.1191598	0.26466	-317.875699	87.85709	-317.876	87.85709	
X Variabl	-181.405	127.4091364	-1.4238028	0.156347	-432.9239186	70.11295	-432.924	70.11295	
X Variabl	-351.688	195.8252163	-1.7959297	0.074292	-738.266961	34.89033	-738.267	34.89033	
X Variabl	31.70188	99.60436923	0.31827801	0.750667	-164.9271495	228.3309	-164.927	228.3309	
X Variabl	425.294	113.8677546	3.73498152	0.000256	200.5075743	650.0803	200.5076	650.0803	
X Variabl	-1371.38	396.4139935	-3.4594519	0.000685	-2153.936205	-588.814	-2153.94	-588.814	
X Variabl	-335.529	148.4043255	-2.260908	0.025041	-628.4935738	-42.5635	-628.494	-42.5635	
X Variabl	23.82721	185.2785944	0.12860206	0.897825	-341.9313471	389.5858	-341.931	389.5858	
X Variabl	-288.945	184.3299368	-1.5675417	0.118859	-652.8306823	74.94094	-652.831	74.94094	

## 7.Data Ethics Framework and Principles

Any investigation or project where we use data that has been outsourced for research must include a Data Ethics Framework. In order for the data to be used correctly and ethically, it is primarily researched by someone who has scraped or obtained it for a purpose. When developing, implementing, and evaluating new programmes or services or conducting analysis, public entities must have a set of guidelines on how to use data sensibly and carefully. Therefore, the UK Government Data Ethics Framework, which is the best fit for it, is what I used for this analysis. The analysis is evaluated by assigning a 5 out of 10 rating to each factor.

Two principles— accountability, and fairness—are tested in the inspection under the UK Government Data Ethics Framework.

**1.Accountability**- Accountability means that the task I completed and the process I followed in conducting the analysis do not negatively impact wider public or society. I shall be chargeable if any loss occurs as a result of my analysis.

**2.Fairness** - Since this is a dataset on Economic Freedom Index, I assurance that there is no content or intention that would be objectable to any group or class in society. Fairness is the elimination of any racism present in any way like sex, region, age, disability etc. during the analysis.

## **8. Conclusion**

In conclusion to analysis of Economic Freedom Index, we can make the following inference and insights:

- Developed countries can help underdeveloped countries in many ways for the growth of their economy. Free trade, experienced economist , low interest rate and human development can help underdeveloped nations in boosting economy of their country.
- Judicial effectiveness, Government Integrity and Population are the 3 variables a developed nation should focus on for the growth of the economy.
- Government integrity, Investment Freedom aand Fiscal health are the 3 variable a developing nation should focus on for the growth of the economy.
- Trade freedom, property rights and tax burden are the 3 variables a underdeveloped nation should focus on.

## **References**

1. *How developed countries can help underdeveloped countries?* | News Brio. (n.d.).
2. <https://newsbrio.com/how-developed-countries-can-help-underdeveloped-countries/>
3. CFA, FRM, and Actuarial Exams Study Notes. (2021). *Economic Growth in the Developed and Developing Economies*. [online] Available at:

<https://analystprep.com/study-notes/cfa-level-2/economic-growth-in-the-developed-and-developed-economies/>.

4. Economic Policy Institute. (2013). *Corporate tax rates and economic growth since 1947*. [online] Available at: <https://www.epi.org/publication/ib364-corporate-tax-rates-and-economic-growth/>.
5. Farley, R. (2019). *Economists: Tariffs Not Boosting GDP*. [online] FactCheck.org. Available at: <https://www.factcheck.org/2019/05/economists-tariffs-not-boosting-gdp/>.
6. scholar.google.co.uk. (n.d.). *Google Scholar*. [online] Available at: [https://scholar.google.co.uk/scholar?q=-+\(PDF\)+Judicial+Efficiency+and+Economic+Growth:+Evidence+based+on+EU+data+\(researchgate.net\)&hl=en&as\\_sdt=0&as\\_vis=1&oi=scholart](https://scholar.google.co.uk/scholar?q=-+(PDF)+Judicial+Efficiency+and+Economic+Growth:+Evidence+based+on+EU+data+(researchgate.net)&hl=en&as_sdt=0&as_vis=1&oi=scholart).
7. Tejvan Pettinger (2017). *Impact of Increasing Government Spending / Economics Help*. [online] Economicshelp.org. Available at: <https://www.economicshelp.org/blog/2731/economics/impact-of-increasing-government-spending/>.