**DETERMINANTS OF ECONOMIC GROWTH OF A NATION**

**ASHISH U. SOLLAPUR**

**RYERSON UNIVERSITY**

**TORONTO**

**STUDENT ID 501140340**

**DATA ANALYTICS, BIG DATA AND PREDICTIVE ANALYTICS**

**DR. DERYA KICI**

**04 APRIL 2022**

**Abstract**

Development of any economy is based on a broad range of indices such as agricultural income, gross domestic product (GDP) growth, employment rate, income equality index etc. An economy’s growth is indicated by such factors which can be used as guidance tools for the respective economy to decide on future course of action. The data points can help understand where the resources need to be optimally allocated for better growth. The data may also present facts where there has been a misuse of available resources thus inviting urgent corrective action.

One of the several reliable data metrics is published by The World Bank featuring a set of various monthly indicators in several broad groups such as climate change, education, health etc. The published data includes a set of data points since the year 1960 for all the countries in the world. However, since the data from 1960 will be very exhaustive for my study, I plan to analyse the data starting from 2001 up to 2020. Some of the metrics have missing data as several aspects were started to be measured only until a few decades ago. In this project, I aim to analyse the datasets published by The World Bank under the following broad groups: (1) Economy & Growth (2) Education (3) Financial Sector (4) Health

By analysing a set of attributes under each of the above broad groups, I plan to carry out exploratory data analysis to understand correlation and causality between the variables. I will also carry out time-series analysis to see how various attributes have changed over years and their interlinkage. This will help me understand if one parameter has a positive or negative impact with some other parameter; or it may be completely independent of other attributes in the dataset.

The dataset contains missing values and probable outliers in certain attributes. After cleaning the dataset, I will first attempt to draw patterns from the data. These patterns would be valid for most of the economies.

In the second part of this project, I plan to short-list 3 economies and carry out regression analysis and K-means classification on the data points of these economies. This will provide a deeper understanding about the data valid specifically for such economies. I will later interpret the findings and recommend areas where more focus is required to achieve faster and balanced overall growth.

My research questions are as follows:

1. What are the top 3 attributes that positively or negatively affect GDP growth the most?
2. Which specific group of attributes pertaining to each of the three datasets – heath, education and financial sector - affect GDP growth the most?
3. Which measures should be implemented to increase GDP for the chosen three economies

As I draw conclusions from the data points, I plan to use several data points from different datasets which are correlated to each other. The data is published in the form of various datasets interlinked with each other. Thus, assessing data metrics from these groups will enable us draw insights about the overall health of an economy, its direction, areas of improvement, strengths, and weaknesses etc.

I will use various tools such as R-programming, Python, Structured Query Language (SQL), and Tableau to analyse data and present my findings.

**Literature Review**

Economic growth means overall development of a region in terms of quality of living, sustainability, environmental friendliness, average income of individuals, inflation, employment opportunities, healthy living, higher education opportunities among other aspects. Each of these stated parameters directly impacts the quality of living of any individual. For example, higher inflation leads to depreciation of the local currency eventually reducing the purchasing power. People of a region have higher productivity if they are healthy and eventually contribute more towards the overall growth of the region. Ample employment opportunities mean a robust workforce ready to deliver higher output in times of requirements. All these positive factors have a multiplier effect on the dependant industries which leads to total economic growth and development. While each of these parameters are important for the growth, one should also take into consideration the environmental impact. Every development should intend to have minimal or no environmental impact, albeit it can leave a positive development at times.

This literature review aims to understand the research already done in the field of economic growth and the parameters directly affecting it. There are papers published for individual aspects such as the impact of education, health, inflation etc. on the growth of economies within the region. Such studies assess the impact of only certain parameters on the GDP growth of localised economies. Hongyi and Huang (2009) studied how health and education affect the economic growth in China. Whereas Aziz and Azmi (2017) studied the factors affecting the GDP growth in Malaysia. There are empirical studies on the economic growth of African countries and its factors such as one carried out by Eggoh, Houeninvo, and Sossou (2015).

Aziz and Azmi (2017) studied the various parameters affecting the GDP growth in Malaysia. The authors tried to find the impact that inflation, foreign direct investment (FDI) and women’s participation in the workforce on the GDP growth in Malaysia. The authors studied more than 30 years of data starting from 1982. Some of the methods they used to analyse were Ordinary Least Squares (OLS) method, Regression analysis to carry their in-depth analysis. The authors wanted to try and find out the reasons and factors affecting the growth of the GDP. They concluded that in their study, inflation does not contribute significantly to the GDP growth. However, the FDI has a huge contribution in deciding the GDP growth of an economy. Furthermore, the paper states that there is a positive impact of the female labour participation on the GDP growth. The paper’s findings are consistent with other relevant studies carried out by the authors earlier. This study is limited to the geographical region of Malaysia.

Eggoh, Houeninvo, and Sossou (2015) studied the impact of human capital and the impact of health on the growth of an economy, especially in the African context. The authors mention that until their study was carried out, there were opposing results regarding the impact of health on the growth of the GDP. The research used two main methods of assessing the impact of health on GDP: the OLS method and the Generalized Method of Moment (GMM) and assessed the data from 1996 to 2010 of the majority of African countries. Indicators such as inflation rate, GDP per capita and share of exports and imports as a percentage of the economy’s GDP were used to understand the changes. The study concluded by stating that more the spending on health, lower was its impact on the GDP growth. The authors cited inefficiencies such as leakages, red tape and corruption as the main reasons for their conclusion. They mentioned that the cossuption levels should be brought down and health spending increased to have overall positive impact on the GDP growth.

Karavaeva (2021) studied the relationship between sustainable development the economic growth in the European Region. The author states that while levels of education were significant deciding factors to the overall economic growth, there was not enough evidence in stating that healthcare support contributed to overall economic growth. The author studied 4 of the 17 interdependent factors contributing to a sustainable future for everyone. These 4 factors were Poverty, Health, Education, and Inequality levels. The author analysed the interdependence of these parameters by employing statistical tests such as correlation, carrying out the Durbin Watson autocorrelation test among other significant analyses. The findings were that higher income inequality promotes higher income growth in the region. Poverty alleviation and health of an individual do not contribute significantly to the economic growth of European Union (EU) member countries.

Deme and Mahmoud (2020) studied the effect of quantity and quality of education on per capita real GDP growth in the African region, specifically for the low- and middle-income countries. They carried out regression analysis for 34 African nations where the data consisted of various education parameters. The Graduate Record Examination (GRE) quantitative and verbal test scores were used to assess the outcome of the quantity and quality of education levels. The studies found that there is statistical correlation between growth in the primary school enrolments and per-capita economic growth. The findings were similar for secondary level education as well. The quality of the education was assessed by the quantum of government spending as a percentage of GDP growth on the education levels. The results supported the theory that higher economic growth is achieved by growth in the quality of education. However, the studies found that due to factors such as corruption and pilferage, there is loss in the quality of education being offered at the ground level. The authors concluded that there is strong association quantity of education and economic growth and that the policy makers should target school enrollment as a tool to achieve economic growth.

Another paper published by Chowdhury (2003) analysed the world bank data to understand the income distribution of the world. This paper focused on the income distribution analysis in terms of inequalities and disparities. The author discussed the growth-inequality parameter and studied the relationship between GDP growth rate and the income inequality by way of regression analysis and Theil’s entropy index. The paper concluded two trends concerning the income equality: (1) The inter-regional component of the inequality is stronger than the global inequality. This inequality increases as time passes. (2) The inequality between two regions increases whereas it decreases within a region over time.

Cooray (2009) published a paper on The Financial Sector and Economic Growth where the author evaluated and assessed the financial sector size, activity, efficiency, and their interaction with the economic growth of 35 low and medium sized economies. The paper states that for the economic growth, the financial sector size, activity and efficiency are important parameters. If the resources are directed for productive uses of the financial sector of the 35 economies under study, it can induce higher and faster growth. To increase the efficiency, the overhead costs should be decreased, and bank concentration should be increased. The skill levels of the population of each of these economies can be increased to promote economic growth. This research paper’s findings are consistent with the results of Beck and Levine (1999) which states that there is a positive correlation between economic growth and the development of the financial sector.

Marquez-Ramos (2019) studied the impact of education and literacy on the growth of an economy. The author studied the data from Spain’s perspective. The study mentioned that the most common assumption when studying this relationship is linearity between the education level and growth of an economy. However, the study assumed that an economy may respond differently depending on the literacy level of its population and thus went on to assess the relationship dynamically. For education levels, the admission numbers for secondary and college levels were considered. The author applied the Ng-Perron unit root test and regression analysis to carry out the research. The paper’s concluding remarks mention that there are inconsistencies in the patterns observed. The author states that education plays a vital role in the economic growth of Spain. However, there were irregularities observed in the patterns and thus, education plays a dual role.

Grant (2019) ventured to study the effect of primary, secondary and tertiary levels of education on the economic activity of a country. The study did not focus on any specific geographical region and hence can be considered to be a global research. The data from a sample set of countries was studied and this sample data was divided among countries with varying levels of income. The sample countries were divided into three groups – low income, lower middle income, and upper middle income. The data for the study was sourced from the World Bank’s publicly available datasets. Instead of applying any statistical methods to assess the impact, the study was carried out empirically. Thus, only observational analyses were made. The study mentions that if the primary education attendance levels are doubled, it will have a positive impact of 4% on food availability. More than 170 million people can come out of poverty if all the children have at least the primary education. The paper mentions that for secondary education, the countries should strive to achieve gender equality as women who complete secondary education have higher chances of earning more than their male counterparts. Tertiary and voluntary education leads to higher aptitude for the students and it enables them to think independently. The paper mentions that the low income countries should focus more on enrolling students and spend more on the education infrastructure while the middle income countries need to stop pilferage of funds so that they are put to better use for infrastructure development. Overall, the higher levels of education lead to higher economic activity eventually leading to higher GDP growth.

While there have been studies concentrating on the factors affecting economic growth in certain specific regions, there are very few research available for assessing the reasons for economic growth and finding its determinants on a global level. I plan to study three determinants of economic growth – Education, financial sector and health and their impact on the growth of the economy. These determinants were selected based on empirical analyses of existing research available. In this research, I aim to find out the if these parameters affect the economic growth of a nation and to what extent.

**Data Description**

World Bank conducts ongoing research of various parameters to depict the economic well being of a nation. I have selected the data available up to the year 2020 which is published by the World Bank. As my research focuses on understanding the impact of the financial sector, health conditions and literacy levels in a country and their impact on the GDP growth of the country, I have used only the relevant four of the several data sets published by the World Bank.

I plan to use four different data sets for the detailed analyses. Each of the datasets has several parameters as tabulated below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dataset** | **Attributes** | **Regions** | **Start Year** | **End Year** |
| Financial Sector | 77 | 267 | 2001 | 2020 |
| Health | 256 | 267 | 2001 | 2020 |
| Education | 163 | 267 | 2001 | 2020 |
| Economy and Growth | 267 | 255 | 2001 | 2020 |

As the available data is multi-tiered – For each attribute, data is available for several years for multiple countries, the description of the important attributes relevant to my study from each of the four data sets is provided below:

Financial Sector:

|  |  |  |  |
| --- | --- | --- | --- |
| Feature Name | Description | Type of Variable | Basic Statistics |
| Country Name | Name of the country | Nominal | Examples:  Canada, Australia, Italy |
| Indicator Name | Types of various indicators monitored by World Bank | Nominal | Examples:  Bank capital to assets ratio (%),  Automated teller machines (ATMs) (per 100,000 adults) |
| 2001-2020 | Year for which the data is available | Categorical | 2001:  Min: -3.312x1011  1st Quartile: 4  Median: 24  Mean: 1.068 x 1012  3rd Quartile: 1.09 x 107  Max: 1.19 x 1015 |

Bank capital to assets ratio (%) (Continuous variable) – This ratio signifies the strength of the banking sector in the country with respect to its quantum of lending.

Automated teller machines (ATMs) (per 100,000 adults) (Discrete variable) - Automated teller machines are computerized telecommunications devices that provide clients of a financial institution with access to financial transactions in a public place.

Inflation, consumer prices (annual %) (Continuous variable) - Inflation as measured by the consumer price index reflects the annual percentage change in the cost to the average consumer of acquiring a basket of goods and services that may be fixed or changed at specified intervals.

Health:

Life expectancy at birth, total (years) (Continuous variable) - Life expectancy at birth indicates the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.

People using at least basic sanitation services (% of population) (Continuous variable) - The percentage of people using at least basic sanitation services, that is, improved sanitation facilities that are not shared with other households.

Education:

|  |  |  |  |
| --- | --- | --- | --- |
| Feature Name | Description | Type of Variable | Basic Statistics |
| Country Name | Name of the country | Nominal | Examples:  Canada, Australia, Italy |
| Indicator Name | Types of various indicators monitored by World Bank | Nominal | Examples:  Current education expenditure, total (% of total expenditure in public institutions),  Literacy rate, adult total (% of people ages 15 and above) |
| 2001-2020 | Year for which the data is available | Categorical | 2020:  Min: 0  1st Quartile: 8  Median: 46  Mean: 4.698 x 106  3rd Quartile: 92  Max: 3.391 x 109 |

Current education expenditure, total (% of total expenditure in public institutions) - Current expenditure is expressed as a percentage of direct expenditure in public educational institutions (instructional and non-instructional) of the specified level of education.

Educational attainment, at least Bachelor's or equivalent, population 25+, total (%) (cumulative) - The percentage of population ages 25 and over that attained or completed Bachelor's or equivalent.

Educational attainment, at least completed primary, population 25+ years, total (%) (cumulative) - The percentage of population ages 25 and over that attained or completed primary education.

Literacy rate, adult total (% of people ages 15 and above) - Adult literacy rate is the percentage of people ages 15 and above who can both read and write with understanding a short simple statement about their everyday life.

Economy & Growth:

|  |  |  |  |
| --- | --- | --- | --- |
| Feature Name | Description | Type of Variable | Basic Statistics |
| Country Name | Name of the country | Nominal | Examples:  Canada, Australia, Italy |
| Indicator Name | Types of various indicators monitored by World Bank | Nominal | Examples:  GDP (constant 2015 US$),  Exports of goods and services (constant 2015 US$) |
| 2001-2020 | Year for which the data is available | Categorical | 2015:  Min: -8.531 x 1014  1st Quartile: 10  Median: 1.224 x 107  Mean: 1.14 x 1013  3rd Quartile: 2.32 x 1010  Max: 1.17 x 1016 |

GDP (constant 2015 US$) (Continuous Variable) - GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. Data are in constant 2015 prices, expressed in U.S. dollars.

Exports of goods and services (constant 2015 US$) (Continuous Variable) - Exports of goods and services represent the value of all goods and other market services provided to the rest of the world.

Imports of goods and services (constant 2015 US$) (Continuous Variable) - Imports of goods and services represent the value of all goods and other market services received from the rest of the world.

As the number of variables available in each of the datasets is more than 200, I will restrict my study to only a few variables which are relevant to my study. These variables are tabulated below:

|  |  |
| --- | --- |
| Health | 1. Life expectancy at birth, total (years) 2. People using at least basic sanitation services (% of population) 3. Physicians (per 1,000 people) |
| Education | 1. Literacy rate, adult total (% of people ages 15 and above) 2. Current education expenditure, total (% of total expenditure in public institutions) |
| Financial sector | 1. Net foreign assets (current LCU) 2. Account ownership at a financial institution or with a mobile-money-service provider (% of population ages 15+) |
| Economic Growth | 1. GDP (constant 2015 US$) (Continuous Variable) |

I have chosen only one variable of GDP in the Economic Growth dataset as I am assessing the impact of health, education and financial sector improvements on the GDP growth of countries. Thus, the GDP will be the dependant variable and others will be independent variables.

**Methodology and Tools to be used:**

**Regression**

**Exploratory Data Analysis**

**Findings Presentation**

**Classification Algorithms**

Naïve Bayes

K-Means classification

I plan to start with exploratory data analysis on the 4 datasets. This will help understand the basic statistical measures of central tendencies.

After exploring the dataset, I will carry out regression analysis and implement classification algorithms such as Naïve Bayes and K-means classification on the data points of these economies.

These analyses will yield details regarding the correlation of the data points and their impact on the GDP growth attribute. Thus, the dependent variable will be the GDP growth rate which I will analyse using the above-mentioned methods.

Finally, I will present the findings of the research including recommendations to increase the GDP growth rate of the countries. In this project, I will use R-programming along with Python wherever necessary. To present the findings, I will use R-programming and Tableau to a certain extent.

**Descriptive Statistics:**

Health dataset

## ï..Country.Name Country.Code Indicator.Name Indicator.Code   
## Length:67830 Length:67830 Length:67830 Length:67830   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## X2001 X2002 X2003   
## Min. : -4 Min. : -18705036 Min. : -1   
## 1st Qu.: 5 1st Qu.: 5 1st Qu.: 5   
## Median : 19 Median : 19 Median : 20   
## Mean : 7175267 Mean : 7120828 Mean : 7248174   
## 3rd Qu.: 78 3rd Qu.: 78 3rd Qu.: 78   
## Max. :6193663732 Max. :6272724236 Max. :6351855732   
## NA's :31129 NA's :30364 NA's :30516   
## X2004 X2005 X2006   
## Min. : -1 Min. : -2 Min. : -2   
## 1st Qu.: 5 1st Qu.: 5 1st Qu.: 5   
## Median : 20 Median : 19 Median : 20   
## Mean : 7287595 Mean : 7157695 Mean : 7302404   
## 3rd Qu.: 78 3rd Qu.: 79 3rd Qu.: 79   
## Max. :6431527221 Max. :6511724848 Max. :6592711655   
## NA's :30221 NA's :28323 NA's :29293   
## X2007 X2008 X2009   
## Min. : -23126156 Min. : -2 Min. : -2   
## 1st Qu.: 5 1st Qu.: 5 1st Qu.: 5   
## Median : 20 Median : 20 Median : 20   
## Mean : 7283690 Mean : 7584909 Mean : 7637512   
## 3rd Qu.: 79 3rd Qu.: 80 3rd Qu.: 81   
## Max. :6674181848 Max. :6757000414 Max. :6839553692   
## NA's :28699 NA's :29743 NA's :29511   
## X2010 X2011 X2012   
## Min. : -2 Min. : -3 Min. : -18168999   
## 1st Qu.: 5 1st Qu.: 5 1st Qu.: 5   
## Median : 19 Median : 20 Median : 19   
## Mean : 7156771 Mean : 7699480 Mean : 7649071   
## 3rd Qu.: 78 3rd Qu.: 80 3rd Qu.: 80   
## Max. :6921854591 Max. :7003760440 Max. :7089254548   
## NA's :25626 NA's :28831 NA's :28070   
## X2013 X2014 X2015   
## Min. : -4 Min. : -5 Min. : -4   
## 1st Qu.: 5 1st Qu.: 5 1st Qu.: 5   
## Median : 20 Median : 20 Median : 19   
## Mean : 8011374 Mean : 7863758 Mean : 7890978   
## 3rd Qu.: 81 3rd Qu.: 80 3rd Qu.: 81   
## Max. :7175500378 Max. :7261846543 Max. :7347679005   
## NA's :29349 NA's :28105 NA's :26963   
## X2016 X2017 X2018   
## Min. : -3 Min. : -15664439 Min. : -4   
## 1st Qu.: 5 1st Qu.: 5 1st Qu.: 5   
## Median : 20 Median : 20 Median : 18   
## Mean : 7994944 Mean : 8365534 Mean : 8507001   
## 3rd Qu.: 81 3rd Qu.: 83 3rd Qu.: 78   
## Max. :7433569330 Max. :7519183459 Max. :7602454161   
## NA's :27739 NA's :29031 NA's :29200   
## X2019 X2020   
## Min. : -2 Min. : -2   
## 1st Qu.: 5 1st Qu.: 5   
## Median : 17 Median : 11   
## Mean : 10037114 Mean : 15647120   
## 3rd Qu.: 77 3rd Qu.: 81   
## Max. :7683372259 Max. :7761620146   
## NA's :34702 NA's :46346

Education dataset

## ï..Country.Name Country.Code Indicator.Name Indicator.Code   
## Length:43092 Length:43092 Length:43092 Length:43092   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## X2001 X2002 X2003   
## Min. :0.000e+00 Min. :0.000e+00 Min. :0.000e+00   
## 1st Qu.:1.000e+01 1st Qu.:1.000e+01 1st Qu.:1.000e+01   
## Median :4.800e+01 Median :4.800e+01 Median :4.800e+01   
## Mean :2.462e+06 Mean :2.471e+06 Mean :2.578e+06   
## 3rd Qu.:9.500e+01 3rd Qu.:9.600e+01 3rd Qu.:9.600e+01   
## Max. :2.798e+09 Max. :2.840e+09 Max. :2.885e+09   
## NA's :23379 NA's :23217 NA's :23751   
## X2004 X2005 X2006   
## Min. :0.000e+00 Min. :0.000e+00 Min. :0.000e+00   
## 1st Qu.:1.000e+01 1st Qu.:1.000e+01 1st Qu.:1.000e+01   
## Median :4.900e+01 Median :4.900e+01 Median :4.900e+01   
## Mean :2.470e+06 Mean :2.479e+06 Mean :2.533e+06   
## 3rd Qu.:9.600e+01 3rd Qu.:9.600e+01 3rd Qu.:9.600e+01   
## Max. :2.933e+09 Max. :2.978e+09 Max. :3.015e+09   
## NA's :22676 NA's :22496 NA's :22626   
## X2007 X2008 X2009   
## Min. :0.000e+00 Min. :0.000e+00 Min. :0.000e+00   
## 1st Qu.:1.000e+01 1st Qu.:1.100e+01 1st Qu.:1.100e+01   
## Median :5.000e+01 Median :5.000e+01 Median :5.000e+01   
## Mean :2.480e+06 Mean :2.457e+06 Mean :2.464e+06   
## 3rd Qu.:9.600e+01 3rd Qu.:9.600e+01 3rd Qu.:9.600e+01   
## Max. :3.055e+09 Max. :3.091e+09 Max. :3.124e+09   
## NA's :21861 NA's :21414 NA's :21289   
## X2010 X2011 X2012   
## Min. :0.000e+00 Min. :0.000e+00 Min. :0.000e+00   
## 1st Qu.:1.100e+01 1st Qu.:1.000e+01 1st Qu.:1.000e+01   
## Median :5.000e+01 Median :5.000e+01 Median :5.100e+01   
## Mean :2.441e+06 Mean :2.406e+06 Mean :2.458e+06   
## 3rd Qu.:9.600e+01 3rd Qu.:9.600e+01 3rd Qu.:9.600e+01   
## Max. :3.153e+09 Max. :3.185e+09 Max. :3.221e+09   
## NA's :20859 NA's :20256 NA's :20483   
## X2013 X2014 X2015   
## Min. :0.000e+00 Min. :0.000e+00 Min. :0.000e+00   
## 1st Qu.:1.000e+01 1st Qu.:1.000e+01 1st Qu.:1.000e+01   
## Median :5.000e+01 Median :5.100e+01 Median :5.100e+01   
## Mean :2.516e+06 Mean :2.475e+06 Mean :2.493e+06   
## 3rd Qu.:9.600e+01 3rd Qu.:9.600e+01 3rd Qu.:9.700e+01   
## Max. :3.255e+09 Max. :3.288e+09 Max. :3.325e+09   
## NA's :20689 NA's :20050 NA's :19975   
## X2016 X2017 X2018   
## Min. :0.000e+00 Min. :0.000e+00 Min. :0.000e+00   
## 1st Qu.:1.000e+01 1st Qu.:1.000e+01 1st Qu.:9.000e+00   
## Median :5.300e+01 Median :5.500e+01 Median :5.500e+01   
## Mean :2.538e+06 Mean :2.689e+06 Mean :2.976e+06   
## 3rd Qu.:9.700e+01 3rd Qu.:9.700e+01 3rd Qu.:9.700e+01   
## Max. :3.360e+09 Max. :3.398e+09 Max. :3.435e+09   
## NA's :20096 NA's :21182 NA's :23274   
## X2019 X2020   
## Min. :0.000e+00 Min. :0.000e+00   
## 1st Qu.:7.000e+00 1st Qu.:8.000e+00   
## Median :5.400e+01 Median :4.600e+01   
## Mean :3.602e+06 Mean :4.698e+06   
## 3rd Qu.:9.900e+01 3rd Qu.:9.200e+01   
## Max. :3.474e+09 Max. :3.391e+09   
## NA's :28175 NA's :35254

Financial Sector dataset

## ï..Country.Name Country.Code Indicator.Name Indicator.Code   
## Length:20216 Length:20216 Length:20216 Length:20216   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## X2001 X2002 X2003   
## Min. :-3.312e+11 Min. :-4.246e+11 Min. :-4.171e+11   
## 1st Qu.: 4.000e+00 1st Qu.: 3.000e+00 1st Qu.: 4.000e+00   
## Median : 2.400e+01 Median : 2.000e+01 Median : 2.400e+01   
## Mean : 1.068e+12 Mean : 1.080e+12 Mean : 1.202e+12   
## 3rd Qu.: 1.090e+07 3rd Qu.: 3.620e+06 3rd Qu.: 1.853e+07   
## Max. : 1.190e+15 Max. : 1.160e+15 Max. : 1.170e+15   
## NA's :12524 NA's :12054 NA's :12227   
## X2004 X2005 X2006   
## Min. :-6.754e+11 Min. :-9.325e+11 Min. :-3.503e+12   
## 1st Qu.: 3.000e+00 1st Qu.: 3.000e+00 1st Qu.: 4.000e+00   
## Median : 2.100e+01 Median : 2.100e+01 Median : 2.200e+01   
## Mean : 1.206e+12 Mean : 1.256e+12 Mean : 1.533e+12   
## 3rd Qu.: 3.233e+05 3rd Qu.: 1.819e+05 3rd Qu.: 1.048e+04   
## Max. : 1.160e+15 Max. : 1.280e+15 Max. : 1.390e+15   
## NA's :11245 NA's :10387 NA's :10762   
## X2007 X2008 X2009   
## Min. :-1.252e+13 Min. :-2.572e+13 Min. :-1.256e+13   
## 1st Qu.: 4.000e+00 1st Qu.: 4.000e+00 1st Qu.: 4.000e+00   
## Median : 2.300e+01 Median : 2.200e+01 Median : 2.300e+01   
## Mean : 1.816e+12 Mean : 2.020e+12 Mean : 2.301e+12   
## 3rd Qu.: 2.874e+04 3rd Qu.: 5.840e+02 3rd Qu.: 1.320e+03   
## Max. : 1.650e+15 Max. : 1.900e+15 Max. : 2.190e+15   
## NA's :10456 NA's :10545 NA's :10459   
## X2010 X2011 X2012   
## Min. :-2.020e+12 Min. :-1.126e+12 Min. :-4.739e+12   
## 1st Qu.: 4.000e+00 1st Qu.: 5.000e+00 1st Qu.: 4.000e+00   
## Median : 2.300e+01 Median : 2.600e+01 Median : 2.100e+01   
## Mean : 2.519e+12 Mean : 2.582e+12 Mean : 3.383e+12   
## 3rd Qu.: 4.343e+04 3rd Qu.: 1.290e+02 3rd Qu.: 1.165e+03   
## Max. : 2.730e+15 Max. : 3.520e+15 Max. : 4.440e+15   
## NA's :9792 NA's :8682 NA's :9902   
## X2013 X2014 X2015   
## Min. :-1.926e+13 Min. :-2.037e+13 Min. :-1.94e+13   
## 1st Qu.: 4.000e+00 1st Qu.: 5.000e+00 1st Qu.: 4.00e+00   
## Median : 1.800e+01 Median : 2.600e+01 Median : 2.00e+01   
## Mean : 3.858e+12 Mean : 4.025e+12 Mean : 5.03e+12   
## 3rd Qu.: 3.460e+02 3rd Qu.: 1.300e+02 3rd Qu.: 1.97e+03   
## Max. : 5.530e+15 Max. : 7.470e+15 Max. : 9.29e+15   
## NA's :9577 NA's :8276 NA's :9262   
## X2016 X2017 X2018   
## Min. :-1.812e+13 Min. :-5.424e+13 Min. :-1.818e+13   
## 1st Qu.: 4.000e+00 1st Qu.: 5.000e+00 1st Qu.: 3.000e+00   
## Median : 1.900e+01 Median : 3.000e+01 Median : 1.800e+01   
## Mean : 6.151e+12 Mean : 3.479e+12 Mean : 4.432e+12   
## 3rd Qu.: 4.350e+02 3rd Qu.: 1.420e+02 3rd Qu.: 3.640e+02   
## Max. : 1.190e+16 Max. : 7.770e+15 Max. : 8.760e+15   
## NA's :9828 NA's :8187 NA's :10086   
## X2019 X2020   
## Min. :-1.107e+12 Min. :-9.931e+11   
## 1st Qu.: 4.000e+00 1st Qu.: 4.000e+00   
## Median : 1.900e+01 Median : 2.800e+01   
## Mean : 4.988e+12 Mean : 6.971e+12   
## 3rd Qu.: 5.350e+02 3rd Qu.: 4.277e+03   
## Max. : 9.950e+15 Max. : 1.130e+16   
## NA's :10391 NA's :12701

Economy and Growth dataset

## ï..Country.Name Country.Code Indicator.Name Indicator.Code   
## Length:67564 Length:67564 Length:67564 Length:67564   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## X2001 X2002 X2003   
## Min. :-8.940e+15 Min. :-2.790e+15 Min. :-9.820e+15   
## 1st Qu.: 1.000e+01 1st Qu.: 1.000e+01 1st Qu.: 1.100e+01   
## Median : 3.658e+06 Median : 4.628e+06 Median : 6.208e+06   
## Mean : 2.891e+12 Mean : 3.550e+12 Mean : 3.318e+12   
## 3rd Qu.: 1.116e+10 3rd Qu.: 1.148e+10 3rd Qu.: 1.249e+10   
## Max. : 4.420e+15 Max. : 4.550e+15 Max. : 4.790e+15   
## NA's :23428 NA's :22562 NA's :22233   
## X2004 X2005 X2006   
## Min. :-5.360e+15 Min. :-7.160e+15 Min. :-6.200e+15   
## 1st Qu.: 1.200e+01 1st Qu.: 1.100e+01 1st Qu.: 1.200e+01   
## Median : 8.163e+06 Median : 7.224e+06 Median : 1.202e+07   
## Mean : 4.004e+12 Mean : 4.272e+12 Mean : 4.767e+12   
## 3rd Qu.: 1.431e+10 3rd Qu.: 1.599e+10 3rd Qu.: 1.779e+10   
## Max. : 5.030e+15 Max. : 5.310e+15 Max. : 5.690e+15   
## NA's :21983 NA's :21132 NA's :20580   
## X2007 X2008 X2009   
## Min. :-1.692e+14 Min. :-4.636e+14 Min. :-2.242e+14   
## 1st Qu.: 1.200e+01 1st Qu.: 1.100e+01 1st Qu.: 9.000e+00   
## Median : 1.295e+07 Median : 9.639e+06 Median : 6.983e+06   
## Mean : 5.785e+12 Mean : 6.301e+12 Mean : 6.552e+12   
## 3rd Qu.: 2.028e+10 3rd Qu.: 2.173e+10 3rd Qu.: 1.943e+10   
## Max. : 5.990e+15 Max. : 6.180e+15 Max. : 6.700e+15   
## NA's :20052 NA's :19510 NA's :19144   
## X2010 X2011 X2012   
## Min. :-1.828e+14 Min. :-2.169e+14 Min. :-3.014e+14   
## 1st Qu.: 1.100e+01 1st Qu.: 1.100e+01 1st Qu.: 1.000e+01   
## Median : 1.105e+07 Median : 1.088e+07 Median : 1.028e+07   
## Mean : 7.815e+12 Mean : 8.676e+12 Mean : 9.276e+12   
## 3rd Qu.: 2.152e+10 3rd Qu.: 2.294e+10 3rd Qu.: 2.348e+10   
## Max. : 6.860e+15 Max. : 7.830e+15 Max. : 8.680e+15   
## NA's :18652 NA's :18183 NA's :18102   
## X2013 X2014 X2015   
## Min. :-3.496e+14 Min. :-5.189e+14 Min. :-8.531e+14   
## 1st Qu.: 1.000e+01 1st Qu.: 1.000e+01 1st Qu.: 1.000e+01   
## Median : 1.082e+07 Median : 1.263e+07 Median : 1.224e+07   
## Mean : 1.029e+13 Mean : 1.112e+13 Mean : 1.140e+13   
## 3rd Qu.: 2.440e+10 3rd Qu.: 2.441e+10 3rd Qu.: 2.320e+10   
## Max. : 1.000e+16 Max. : 1.160e+16 Max. : 1.170e+16   
## NA's :17999 NA's :17733 NA's :17361   
## X2016 X2017 X2018   
## Min. :-1.370e+15 Min. :-1.350e+15 Min. :-1.300e+15   
## 1st Qu.: 1.000e+01 1st Qu.: 1.100e+01 1st Qu.: 1.000e+01   
## Median : 1.285e+07 Median : 1.882e+07 Median : 1.830e+07   
## Mean : 1.250e+13 Mean : 1.380e+13 Mean : 1.565e+13   
## 3rd Qu.: 2.418e+10 3rd Qu.: 2.670e+10 3rd Qu.: 2.956e+10   
## Max. : 1.320e+16 Max. : 1.530e+16 Max. : 1.920e+16   
## NA's :17992 NA's :18399 NA's :19215   
## X2019 X2020   
## Min. :-1.380e+15 Min. :-2.020e+15   
## 1st Qu.: 1.000e+01 1st Qu.: 1.700e+01   
## Median : 2.636e+07 Median : 4.462e+08   
## Mean : 1.788e+13 Mean : 2.627e+13   
## 3rd Qu.: 3.438e+10 3rd Qu.: 7.237e+10   
## Max. : 2.490e+16 Max. : 3.600e+16   
## NA's :20638 NA's :31549

**References:**

Aziz, R. N. A. R., & Azmi, A. (2017). Factors affecting gross domestic product (GDP) growth in Malaysia. *International Journal of Real Estate Studies*, 11(4), 61-67. https://www.utm.my/intrest/files/2017/09/07-FACTOR-AFFECTING-GROSS-DOMESTIC-PRODUCT-GDP-GROWTH-IN-MALAYSIA1.pdf

Beck, T., & Levine, R. (1999). *A new database on financial development and structure* (Vol. 2146). World Bank Publications.

Chowdhury K. (2003). Empirics for World Income Distribution: What Does the World Bank Data Reveal? *The Journal of Developing Areas*, *36*(2): 59–83. http://www.jstor.org/stable/4192920

Cooray A. (2009). The Financial Sector and Economic Growth. *Economic Record* https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-4932.2009.00584.x

Deme, M., & Mahmoud, A. M. A. (2020). Effect of Quantity and Quality of Education on Per Capita Real-GDP Growth: Evidence from Low- and Middle-Income African Countries. Applied Economics, 52:57, 6248-6264. [https://doi.org/10.1080/00036846.2020.1789058](about:blank)

Eggoh, J., Houeninvo, H., & Sossou, G. A. (2015). Education, health and economic growth in African countries. Journal of Economic Development, 40(1), 93. <https://www.researchgate.net/profile/Hilaire-Houeninvo/publication/282646931_EDUCATION_HEALTH_AND_ECONOMIC_GROWTH_IN_AFRICAN_COUNTRIES/links/5615219208aec62244119a7c/EDUCATION-HEALTH-AND-ECONOMIC-GROWTH-IN-AFRICAN-COUNTRIES.pdf>

Grant Catherine (2017). The contribution of education to economic growth. *Knowledge, evidence and learning for development* [K4D\_HDR\_The\_Contribution\_of\_Education\_to\_Economic\_Growth\_Final.pdf (publishing.service.gov.uk)](https://assets.publishing.service.gov.uk/media/5b9b87f340f0b67896977bae/K4D_HDR_The_Contribution_of_Education_to_Economic_Growth_Final.pdf)

Hongyi, L. I., & Huang, L. (2009). Health, education, and economic growth in China: Empirical findings and implications. *China Economic Review*, 20(3), 374-387. [https://doi.org/10.1016/j.chieco.2008.05.001](about:blank)

Indicators | Data. (2020). The World Bank. https://data.worldbank.org/indicator?tab=featured

Karavaeva, K. (2021). The Relationship between Sustainable Development and GDP Growth in EU Countries. *Charles University* <http://hdl.handle.net/20.500.11956/126546>

Marquez-Ramos, Laura. (2019) Education and Economic Growth: An Empirical Analysis of Nonlinearities. *Applied Economic Analysis* <https://www.emerald.com/insight/content/doi/10.1108/AEA-06-2019-0005/full/html>

|  |  |
| --- | --- |
| Health | 1. Life expectancy at birth, total (years) 2. People using at least basic sanitation services (% of population) 3. Physicians (per 1,000 people) |
| Economic Growth | 1. GDP (constant 2015 US$) (Continuous Variable) |
| Education | 1. Literacy rate, adult total (% of people ages 15 and above) 2. Current education expenditure, total (% of total expenditure in public institutions) |
| Financial sector | 1. Net foreign assets (current LCU) 2. Account ownership at a financial institution or with a mobile-money-service provider (% of population ages 15+) |