



UNIVERSITY OF LEEDS

The Effects of Economic Growth on Health

MATH5747M Learning Skills through Case Studies
Teaching staff: Benjamin Thorpe, Charles Taylor, Sofya Titarenko

Shreyansh Parihar, Student ID: 201684933

Contents

1	Introduction	3
2	Data and Methodology:	3
2.1	Data	3
2.2	Methodology	3
3	Procedure	4
4	Results	5
4.1	Relationship between GDP per Capita and IMR between countries and Comparison of cross-sectional and longitudinal relationships over time(2000 –2017)	5
4.2	Variation of the relationship across different countries and possible correlation with historical events (2000 –2017)	6
4.3	Relationship between GDP per Capita and MMR between countries and Comparison of cross-sectional and longitudinal relationships over time(2000 –2017)	8
4.4	Variation of the relationship across different countries and possible correlation with historical events (2000 –2017)	9
4.5	Relationship between GDP per Capita and Life Expectancy at birth between countries and Comparison of cross-sectional and longitudinal relationships over time(2000 –2017)	11
4.6	Variation of the relationship across different countries and possible correlation with historical events (2000 –2017)	12
4.7	GDP per capita and health indicator relationships between countries over time (2000 –2017)	13
4.8	Assessing causality: Does wealth cause health or vice versa?	14
5	Conclusions	15
	References	15

1 Introduction

The relationship between economic growth and health outcomes has been a topic of great interest and study among researchers, policymakers, and public health experts. Country leaders and policymakers are always interested in question that how great is the effect of the economic indicators and health factors one each other as it allows them to efficiently allocate the resources available. It is also important to discuss on causal relationship between these factors to approach towards a reliable answers.

In this report we will discuss the relationship between economical indicator and health indicators in details and apply necessary statistical model along with other dependent factors to access the causality. We are considering GDP per capita as one of the economic indicators and study the with following 4 health indicators:

1. Life Expectancy - Life Expectancy refers to the average number of years a person is expected to live based on current mortality rates and demographic factors. It is a commonly used indicator to assess the overall health and well-being of a population.
2. Infant Mortality rate (IMR) - MR stands for Infant Mortality Rate. It is a crucial health indicator that measures the number of deaths of infants under one year of age per 1,000 live births in a given population. IMR is a significant measure of the overall health and well-being of a population, reflecting the quality of healthcare, access to prenatal and neonatal care, nutrition, and socioeconomic factors.
3. Maternal Mortality rate (MMR) - Maternal Mortality Rate (MMR) refers to the number of maternal deaths per 100,000 live births in a specific population or region. MMR reflects the quality of maternal healthcare services, access to prenatal and antenatal care, skilled birth attendance, and the overall status of women's health and rights.
4. Health Expenditure per capita - Health Expenditure per capita refers to the average amount of money spent on healthcare services per person in a given population. It is an important measure that provides insights into the level of investment in healthcare and the financial resources allocated to maintaining and improving population health.

In this report we will specifically look into below points for aforementioned indicators:

- GDP per capita and health indicator relationships between countries over time
- Relationship between countries (cross-sectional) the same as within the United Kingdom in period 2000-2017 (longitudinal)
- Variation of the relationship across different countries and possible correlation with historical events.
- Assessing causality: Does wealth cause health or vice versa?

2 Data and Methodology:

2.1 Data

For the analysis we have collected our data from World Bank site (BANK, 2023) and WHO(Wealth Health Organization) site (Organization, 2023). The dataset comprises different csv files for each health indicators and GDP per capita by country, years and region.

2.2 Methodology

1. Data Collection: The first step involves gathering relevant data on GDP per capita, health indicators, and other potential confounding factors from reputable sources such as international databases, research institutions, and official government reports. The data should cover a sufficient time period to capture meaningful trends and variations in economic growth and health outcomes.
2. Data Preprocessing: Once the data is collected, preprocessing steps are performed to ensure data quality and compatibility. This includes data cleaning, manipulating handling missing values, and standardizing variables if necessary. Additionally, normalization techniques is applied during statistical analysis.
3. Exploratory Data Analysis (EDA): EDA is conducted to gain insights into the distribution, variability, and relationships between variables. Descriptive statistics, visualizations, and correlation analyses were

performed to explore the initial patterns and identify potential associations between GDP per capita and health indicators.

4. Cross-Sectional Analysis: In the cross-sectional analysis, the relationship between GDP per capita and health indicators is examined across different countries at a specific point in time. Various statistical techniques, such as correlation analysis, regression models, and hypothesis testing, may be employed to quantify and assess the strength and significance of the relationship.
5. Longitudinal Analysis: The longitudinal analysis focuses on examining the relationship between GDP per capita and health indicators within a specific country over time. Time series analysis, panel data models, or other suitable methodologies are used to account for temporal dependencies and assess how changes in economic growth relate to changes in health outcomes.
6. Causality Assessment: To determine whether health causes wealth or vice versa, causal inference technique is applied, here we identify potential confounding factors that may influence the relationship between economic growth and health. Additional variables such as Human Development Index(HDI) and External Balance are considered as potential confounding factors.
7. Statistical Modeling and Interpretation: Statistical model such as multiple linear regression is used to estimate the relationship between economic growth and health outcomes along with confounding factors. The coefficients, significance levels, and effect sizes are interpreted to draw conclusions regarding the causal relationship between wealth and health.

3 Procedure

In this section we will discuss in detail on the procedure involved in methodology:

1. For data collection, we imported the necessary csv file and unpivot the data to structure it properly using MS Excel power-query option. Once structured, we can combine using the merge query option in Excel within power-query window to get all the health and wealth indicators in single source file.
2. Once we have the csv file ready, we used python library called Pandas and R to identify missing values, and then we decided to remove those missing values.
3. Once our data is cleaned, we used various visualization tools to plot the graphs. Following tools were used:
 1. Tableau - A data visualization software. It was used to plot graph between MMR and GDP per capita for the United Kingdom over the years and observe the data
 2. ggplot2 - R library package. Used to plot various graphs to study relationship between Life Expectation and GDP per capita and between IMR and GDP per capita.
 3. Matplotlib - A python library used to plot various graphs to study relationship between Health Expenditure per capita and GDP per capita.
4. Next we will prepare our data to perform statistical analysis. We observed that there were columns having different scales and negative values, thus we performed Normalization using min-max normalization technique to bring the input variables in range 0-1.
5. Next is to use multiple linear regression to determine causality, we generated and evaluate multiple regression model using Scikit-learn module in python and `glm()` function in R, we used following inputs and outputs (Equations represented will of form $Output = \sum_{i=1}^n \beta * Inputs$):
 1. Model 1 equation:

$$GDPpercapita = \beta_0 + \beta_1(LifeExpectancy) + \beta_2(UnemploymentRate) + \beta_3(LiteracyRate) + \beta_4(HealthExp.percapita) + \beta_5(ExternalBalance)$$

Here health factors are considered as input along with confounding factor - External balance and wealth as output determined based on inputs. Performed via Scikit-learn in python.

2. Model 1 equation without External Balance:

$$GDPpercapita = \beta_0 + \beta_1(LifeExpectancy) + \beta_2(UnemploymentRate) + \beta_3(LiteracyRate) + \beta_4(HealthExp.percapita)$$

This is the same model, but this time we will not use the confounding factor - External balance. Performed via Scikit-learn in python.

3. Model 2 equation:

$$GDPpercapita = \beta_0 + \beta_1(GDPpercapita) + \beta_2(GNI) + \beta_3(InflationPercentage) + \beta_4(HDI)$$

Here wealth factors are considered as input along with confounding factor - HDI and health as output determined based on inputs. Used `glm()` in R.

4. Model 2 equation without HDI:

$$GDPpercapita = \beta_0 + \beta_1(GDPpercapita) + \beta_2(GNI) + \beta_3(InflationPercentage)$$

This is the same model as model-2, but this time we will not use the confounding factor - HDI. Used `glm()` in R.

6. Once we have models created, we observed the R-Square and necessary coefficients to compare them and get the results.

4 Results

As per the analysis done, below are the results discussed for each health indicators.

4.1 Relationship between GDP per Capita and IMR between countries and Comparison of cross-sectional and longitudinal relationships over time(2000–2017)

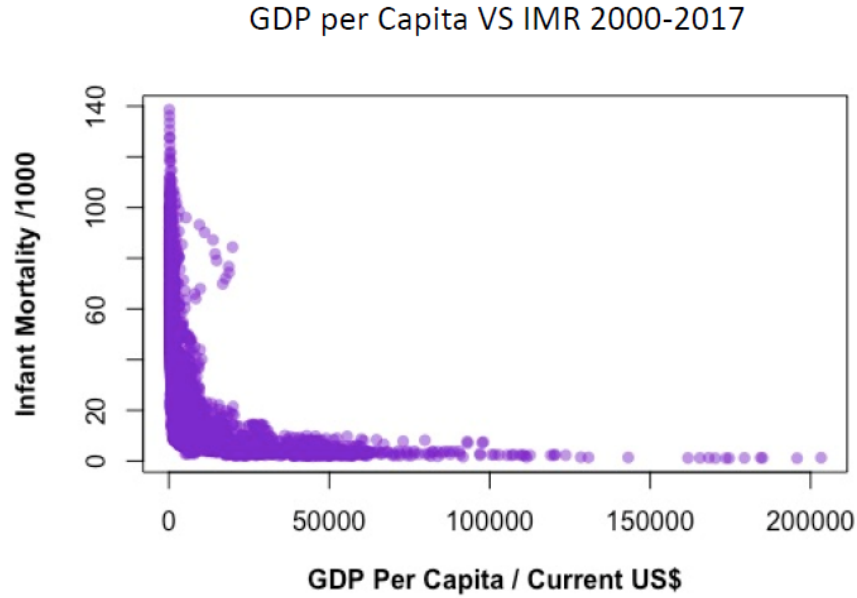


Figure 1: GDP per capita VS IMR 2000-2017

In Figure 1 we can observe as the GDP per capita is less, the IMR is at its peak. It is denser where GDP is low. But as the GDP increases the value of IMR lowers up till a certain point. Till 100000 USD the value of GDP gets really low but after that it does not have much effect on the IMR. We can see some values of IMR where the GDP is 200000 USD. This tells us that the relationship between GDP per capita and IMR

between countries is noticeably decreasing up to a certain point. After that point, the GDP per capita value does not have much effect on IMR.

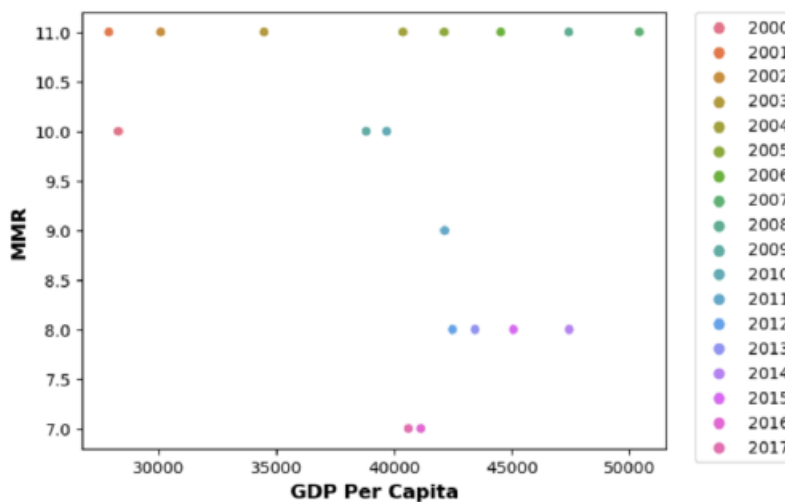


Figure 2: GDP per Capita VS IMR 2000-2017 for UK

In Figure 2, we are discussing the relationship between GDP Per Capita and IMR in the UK. We can see in the early 2000s when the GDP was around <30000 USD the IMR of the UK was really high indicated in red around 5.5. But as the years pass by the value gradually lowers. In 2017, when the GDP is around 41000 USD, the value of IMR is at low <4. If we compare the above two graphs, we can see a drastic difference in how GDP plays an important role in IMR. In the 1st graph, where GDP is close to 0 the IMR value is around 140, which is no near the UK. And when the GDP is close to 50000 USD between countries that is when the values are quite comparable to the UK.

4.2 Variation of the relationship across different countries and possible correlation with historical events (2000 –2017)

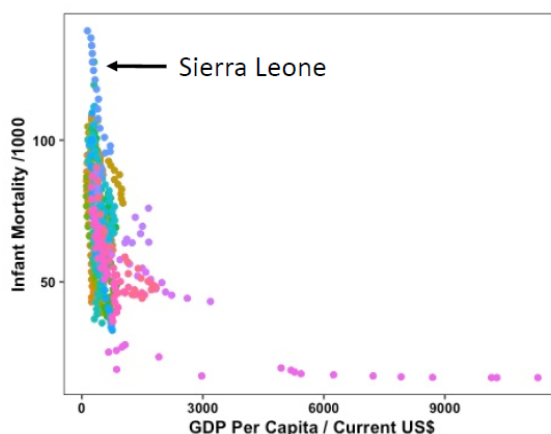


Figure 3: GDP per capita VS Infant Mortality for Low income level countries

Figure 3 shows the values of IMR in Low-income level countries. We could see the saturation is more at the upper region of the graph, which is high value of IMR. We also see that, there is a significant change in IMR

in Sierra Leone (2017), in Sierra Leone the IMR went up to 27.6 after the Ebola outbreak.

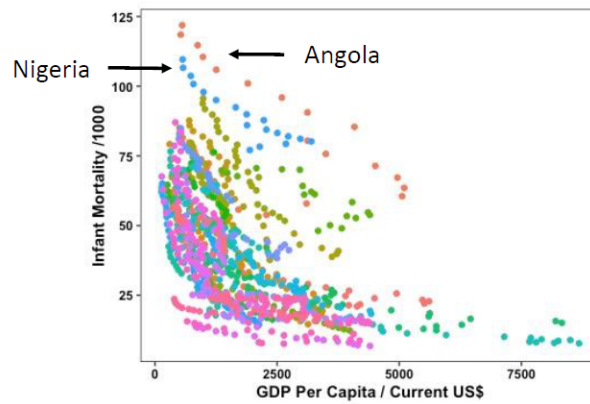


Figure 4: GDP per capita VS Infant Mortality for Lower middle income level countries

In the Angola and Nigeria region as seen in Figure 4, the value of IMR is close to 110, which is very high. This could be due to Neonatal causes, child pneumonia, malaria, diarrhoea, HIV/AIDS, measles. And unsafe water. Such factors affect the IMR values.

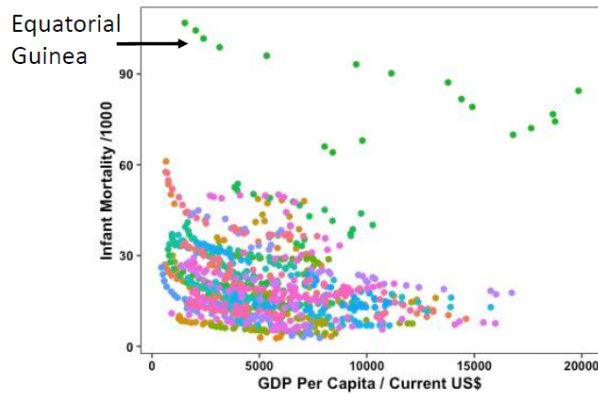


Figure 5: GDP per capita VS Infant Mortality for Upper middle income level countries

In Figure 5, we see the relationship of IMR with the upper middle-income level countries. We could see a significant change that most of the IMR values are in the lower region of the graph. We can see that countries with higher GDP has stable values of IMR.

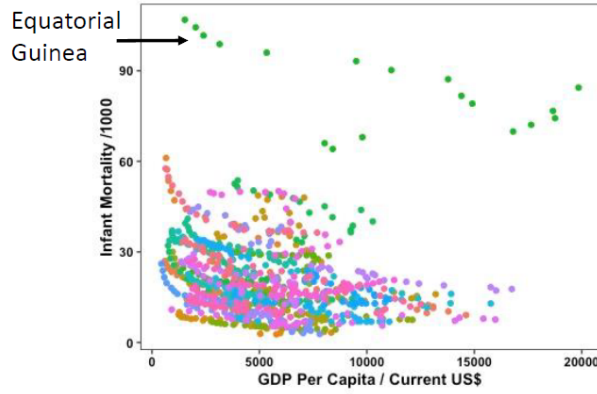


Figure 6: GDP per capita VS Infant Mortality for High income level countries

Figure 6 shows the relationship of IMR with the high income level countries. We can observe that the IMR values are more saturated at the origin. Which means countries with higher GDP has low IMR values.

4.3 Relationship between GDP per Capita and MMR between countries and Comparison of cross-sectional and longitudinal relationships over time(2000–2017)

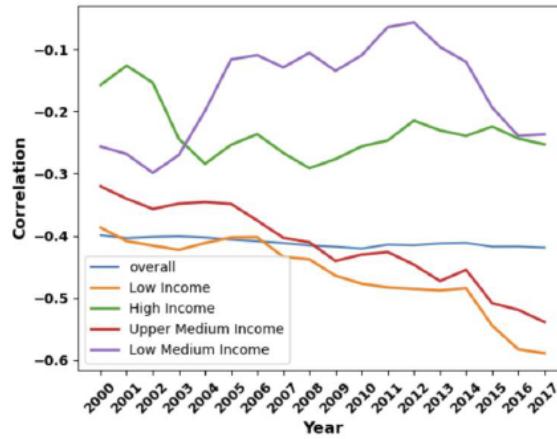


Figure 7: Correlation between GDP and MMR

In Figure 7, we see the correlation between GDP per capita has an overall increase, but the increase is not significant in 17 years as it has a dynamic behaviour, which may suggest that there can be other factors which has more affect on the relation rather than the relation among them. We can see in Figure 8, that the increase in GDP per capita for UK does not relate to MMR, the change in MMR is random.

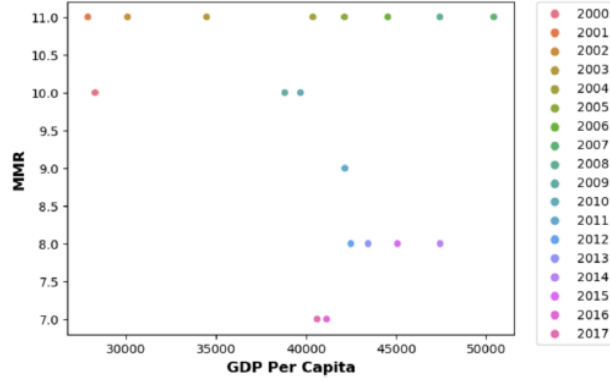


Figure 8: GDP Per Capita Vs MMR UK 2000-2017

4.4 Variation of the relationship across different countries and possible correlation with historical events (2000 –2017)

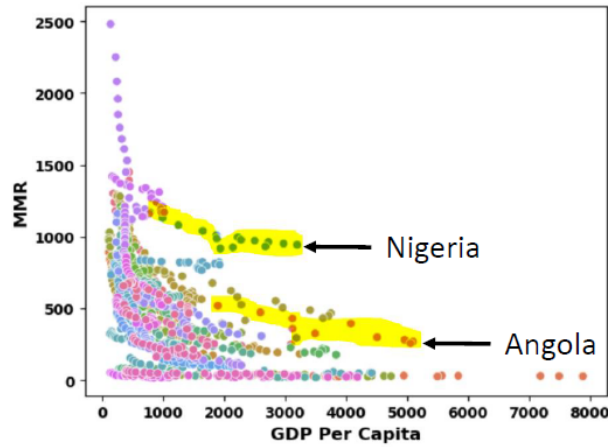


Figure 9: GDP Vs MMR for Low Income Countries

In figure 9, we can see that countries that have lower GDPs have a higher value of MMR. For e.g. in Angola in 2014-16 the Ebola outbreak and droughts 2014 caused significant damage to society. Due to this, the MMR value increased. Similarly, in Boko Haram Insurgency (2009), due to which the value of MMR increased.

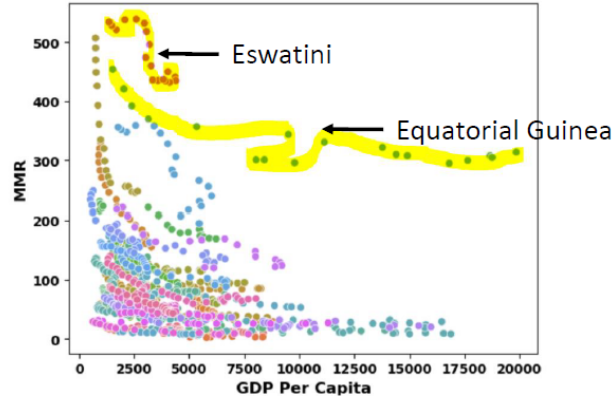


Figure 10: GDP Vs MMR for Low Medium Income Countries

From figure 10, we observe following anomalies from low medium income countries: Eswatini: Droughts (2016), HIV (since 1986) Equatorial Guinea: Decrease in oil prices (2014), economic inequality, poverty

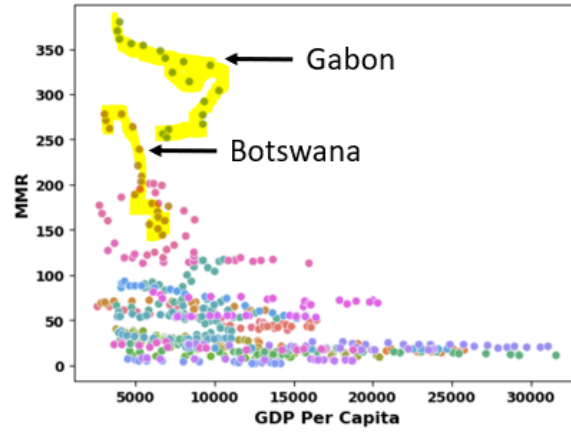


Figure 11: GDP Vs MMR for Upper Medium Income Countries

From figure 11, we observe following anomalies from upper medium income countries: Gabon: Drop in Oil prices (2014), Political instability (2009,2016) - These factors might have caused high MMR value of 300. Botswana: Financial Crisis (2008) can because for MMR reaching near 200.

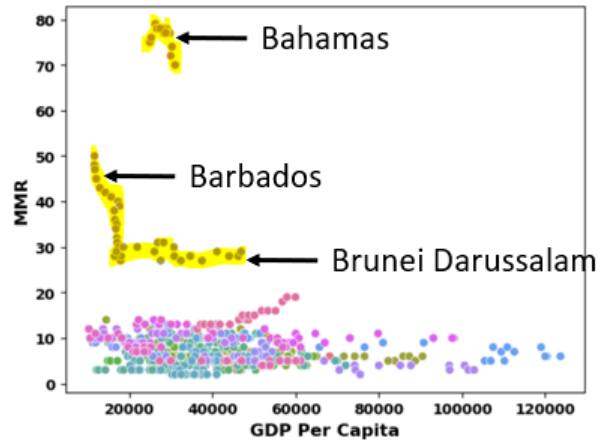


Figure 12: GDP Vs MMR for High Income Countries

In figure 12, we could see a drastic change in the MMR range. In high income countries the MMR value did not go high as in lower income countries. But in high income countries can also have some anomalies which caused the MMR value to go high. In the financial crisis of 2008 Bahamas MMR value went to 75, similarly for Barbados, around 45.

4.5 Relationship between GDP per Capita and Life Expectancy at birth between countries and Comparison of cross-sectional and longitudinal relationships over time(2000 –2017)

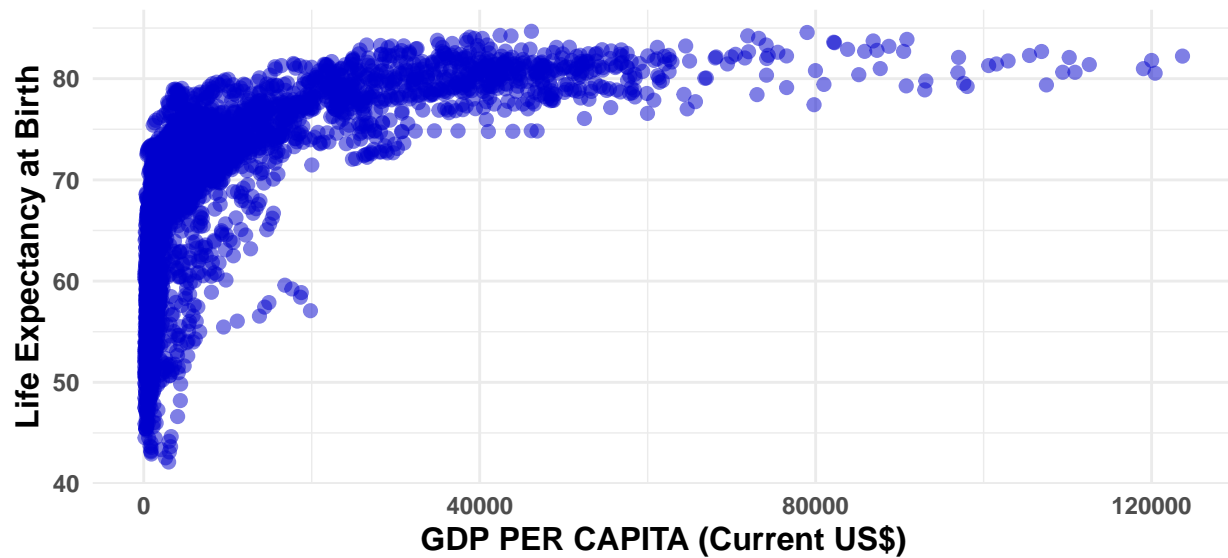


Figure 13: GDP Vs Life Expectancy 2000-2017

Figure 13 shows the relationship between Life expectancy at birth vs GDP per capita. In countries where the GDP is low life expectancy is also low. As gradually the GDP per capita is increasing the life expectancy is also increasing. We can see a major saturation at 40000 USD. But after that one point, there is not much change in the life expectancy. This may be because up to that point, all the basic necessities are getting

fulfilled but after that point, people are investing in luxuries things which maybe irrelevant to the contribution in life expectancy. We still see slight increase, that is maybe due to increase in population.

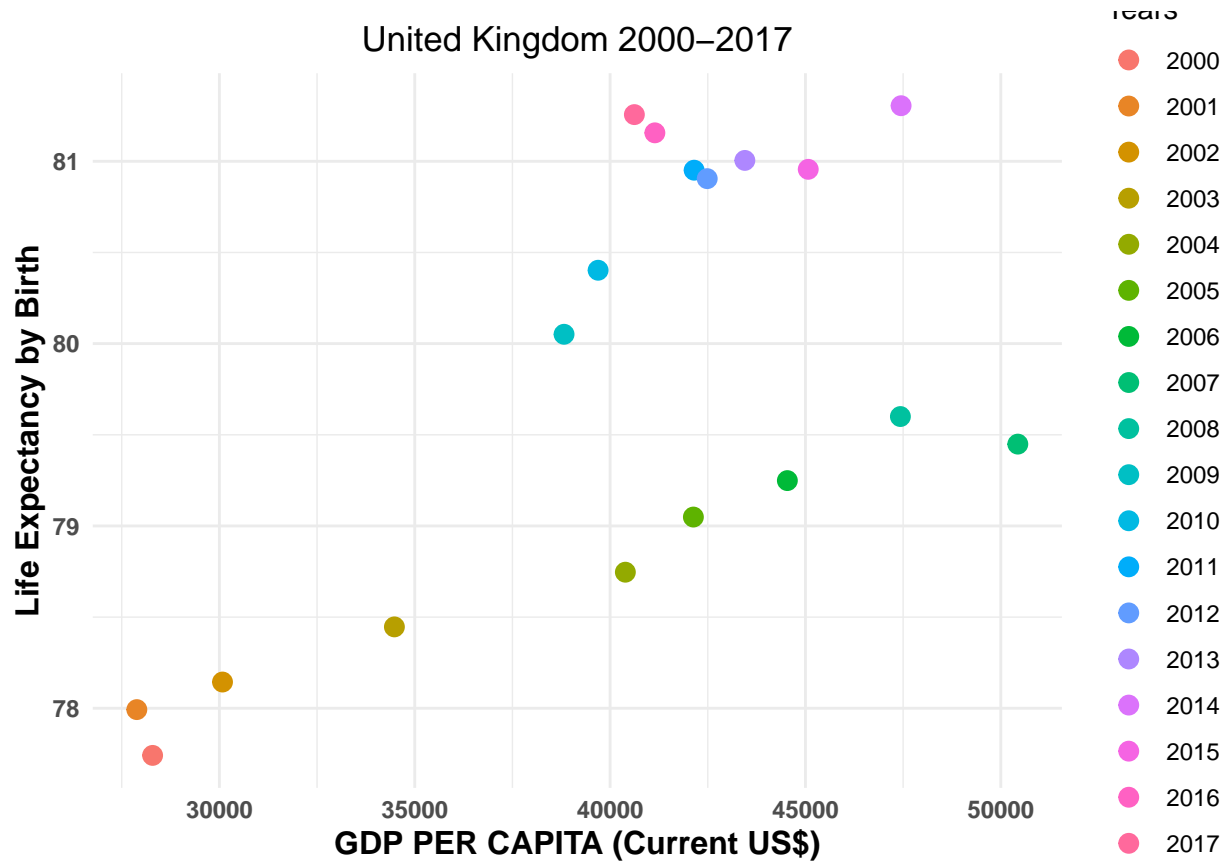


Figure 14: GDP Vs Life Expectancy UK 2000-2017

Figure 14 shows a relationship between life expectancy and GDP in the UK. We could see that as the years pass by life expectancy is increasing. In 2016 the life expectancy value is around 81. We could see a slight change in the behaviour in 2008. This may be because of the Financial crisis in 2008. We can see the same behaviour as before after a certain point there is not much change in the value similarly like we saw in Figure 11. The life expectancy value gets stable.

4.6 Variation of the relationship across different countries and possible correlation with historical events (2000 –2017)

From figure 15, we can make following observation:

- We can see a breakdown of life expectancy in the period 2011-2012 for Syria and Libya. This is likely due to the Civil war in Libya and Syria.
- The impact on life expectancy is on a larger-scale in Syria as there was a high death toll which may result in a drastic decrease in life expectancy since the war was still going on till the recent cease-fire agreement in 2021.
- And that may be the reason we can there is no data from Syria after 2012.
- South Africa: Suffering from HIV/AIDS epidemic since 1995 may cause a drop in life expectancy to the lowest in 2005.
- Same as South Africa, it can also be seen in Lesotho, potentially due to the same reason.

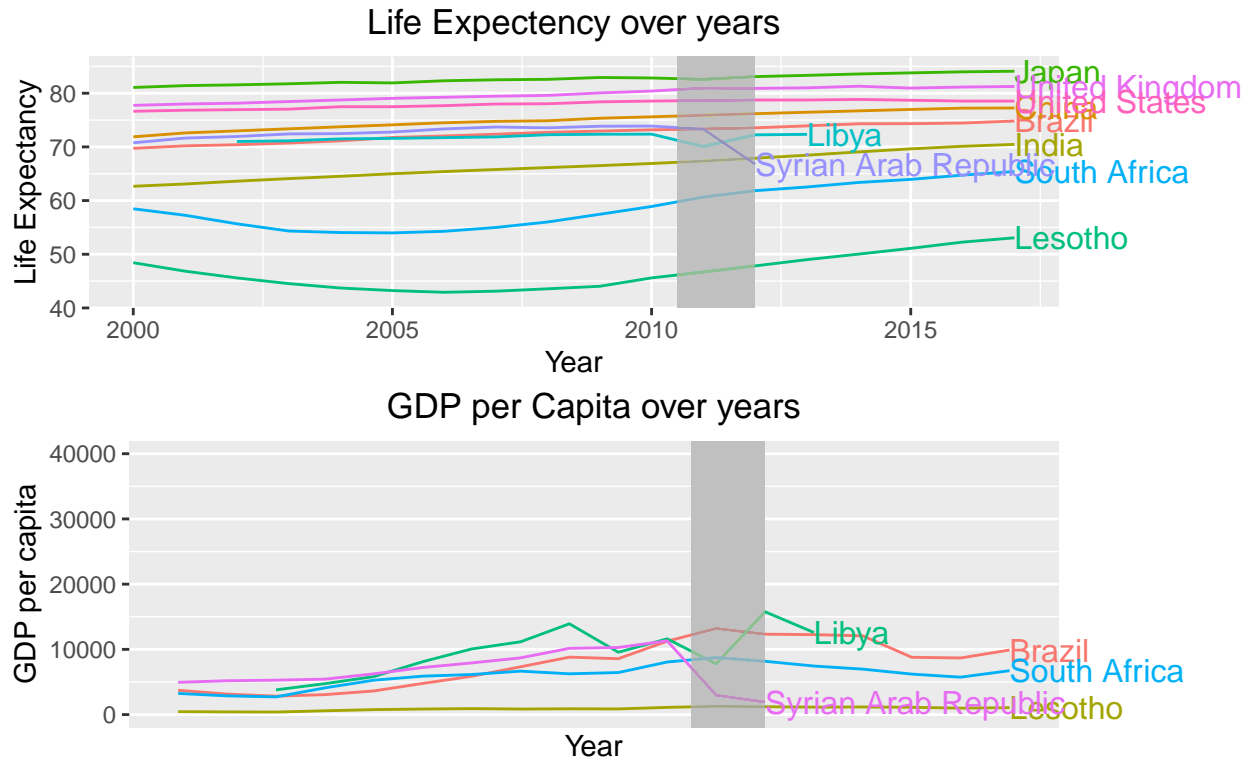


Figure 15: Historic events in certain countries 2000-2017

4.7 GDP per capita and health indicator relationships between countries over time (2000 –2017)

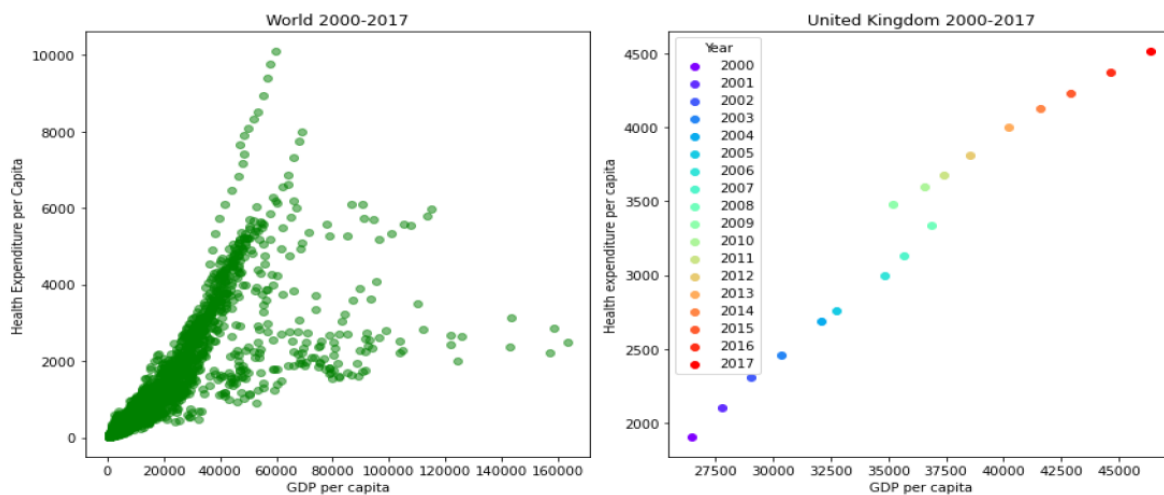


Figure 16: Health Expenditure UK and world 2000-2017

In the above graph we see the relationship between Health expenditure per capita and GDP. In the left graph in figure 16, we can see that countries with lower GDP value have much less Health expenditure per capita. We can see a linear growth in expenditure as the GDP increases. But after a point, there is not much increase

maybe because after that point the government must be investing the money somewhere else. We can see the highest to be 10000 where the GDP is 60000 USD. And the lowest is at 0. In the right graph in figure 16 we see the relationship is only of the UK. We can see a linear increase in the value of health expenditure. The maximum is at 4500 in 2017. We can see a slight change in the graph at 2008. This is caused because of the Financial crisis.

4.8 Assessing causality: Does wealth cause health or vice versa?

As seen in Procedure section we get 2 models each for wealth and health indicators as input and outputs, one with and one without confounding factors. let's see the results.

```
## #MODEL WITH HDI
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      34.8595      0.3715  93.826 < 2e-16 ***
## GDP_per_capita_norm -2.8239      0.6527  -4.327 1.56e-05 ***
## GNI_norm          0.1542      0.9835   0.157  0.875
## inflation_percent_norm -10.7069    2.2908  -4.674 3.09e-06 ***
## HDI              51.9431      0.5694  91.223 < 2e-16 ***
## Multiple R-squared:  0.8353, Adjusted R-squared:  0.774
##
## #MODEL WITHOUT HDI
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      67.3053      0.2111 318.856 < 2e-16 ***
## GDP_per_capita_norm 36.3360      0.9267  39.208 < 2e-16 ***
## GNI_norm          5.8555      1.8959   3.088 0.00203 **
## inflation_percent_norm -34.7123    4.3943  -7.899 3.85e-15 ***
##
## Multiple R-squared:  0.3837, Adjusted R-squared:  0.369
```

Figure 17: Health Expenditure UK and world 2000-2017

As we see the R-square value from the output snippet in figure 17 where wealth indicators are input, the model without HDI values performed much worse with R-square value of 0.38 which is a drastic drop in performance as compare to model with HDI in it, although the GDP per capita seems statistically significant as it has high coefficient magnitude and <0.05 p-value. While this is not true when we have health indicators as inputs, we observed that we got 0.78 with external balance as input and 0.75 without external balance as confounding factor. The result can be seen in figure 18.

R-Squared	0.774		
Adjusted R-Squared	0.769		
Input Variable	Coefficient	P-value	VIF
Life Expectancy	561	0.001	2.84
Unemployment Rate	-258	0.090	1.06
Literacy Rate	85	0.330	1.64
Health Exp. per Capita	8	0.000	2.05

R-Squared	0.78		
Adjusted R-Squared	0.77		
Input Variable	Coefficient	P-value	VIF
Life Expectancy	497	0.004	2.95
Unemployment Rate	-241	0.111	1.06
Literacy Rate	97	0.263	1.65
Health Exp. per Capita	8	0.000	2.11
External Balance	3e-83	0.049	1.04

Figure 18: Health Expenditure UK and world 2000-2017

5 Conclusions

In the results we have seen the impact of GDP per capita on multiple health factors, and we got to see that for some indicators we see a change upto certain value of GDP per capita while for some health indicators like MMR the relationship seems independent. We have also seen that presence of various historical events had major impact on health factors as we see multiple breaks in various countries for these indicators despite their GDP per capita increasing.

We also ran multiple regression model and found out that the certain wealth factors we took were not a good fit for the health indicators, although out of all the wealth indicators we took, we did found out that GDP per capita was statistically significant, but after doing causality assessment with the help of confounding factor we can conclude overall wealth indicators including GDP per capita has less impact on health. While on the other hand we see that health indicators were at least significant to GDP per capita which is considered as one of the most important economic indicator. Resulting in the conclusion that health causes wealth.

References

- BANK, T.W. 2023. World bank open data. [Accessed 1 May 2023]. Available from: [Https://data.worldbank.org/](https://data.worldbank.org/).
- Organization, W.H. 2023. THE GLOBAL HEALTH OBSERVATORY. [Accessed 1 May 2023]. Available from: <https://www.who.int/data/gho/indicator-metadata-registry/>.