Effect of Economic Growth on Health

Contents

1 Introduction

In the field of public health research and policy creation, the interaction of economic indicators and health outcomes is seen as a critical focal area. This line of investigation yields valuable insights into community health determinants and inequities, emphasising the need of understanding social well-being and developmental pathways. The report seeks to look into the complex relationship between certain economic indicators and relevant health indicators. Detailed analysis reveals underlying patterns and correlations, contributing to conversations about health equity, socioeconomic determinants, and policy formation initiatives. To understand both linear and nonlinear links, the study takes a systematic approach that includes numerical summaries, exploratory visualisations, trend analysis, and modelling. This leads to a better knowledge of how economical factors affect different areas of public health.

Ultimately, this research endeavor aims to expand our understanding of the interplay between economic determinants and health outcomes. Through rigorous analysis and empirical investigation, it seeks to inform evidence-based policy-making and interventions aimed at promoting equitable health outcomes across diverse demographic groups.

2 Methodology

The methodology follows a systematic approach aid to find the relationship between economic indicators and few of the most significant health indicators. The economic indicators selected for this specific research are Gross Domestic Product per Capita (GDP) (NY.GDP.PCAP.CD), Consumer Price Index (CPI) for Inflation (FP.CPI.TOTL.ZG), and Health Expenditure per Capita (SH.XPD.CHEX.PC.DC). These indicators were selected based on their significance and influence in a country's success and development. For the health indicators, the indicators under investigation are Infant Mortality rate (SP.DYN.IMRT.IN), Immunisation Coverage (SH.IMM.IDPT), Maternal Mortality ratio (SH.STA.MMRT) and Prevalence of undernourishment (SN.ITK.DEFC.ZS). These indicators are amongst the significant indicators which helps in evaluating the health condition of a nation.

Followed by indicator selection, various exploratory data analysis techniques are applied to gain insight into the data's characteristics and distributions. Various types of visualisations such as scatterplots are utilised to disclose the patterns and identify outliers.

After analysing the trend and visualisations, modelling was done to capture the linear and non-linear relationships to strengthen the research and find the effect and influence of economic indicators on health variables. Linear regression is used to capture the linear associations, while Generalised Additive Models (GAMs) are utilised to detect the potential non-linear associations. Separate models are fitted for each health indicator to allow for the identification of specific effects of economic indicators. Though this methodology, the study aims to provide comprehensive insights into the complex interplay between the economical and health indicators.

3 Data Collection

The data used for the research is retrieved from reputable source, the World Bank's World Development Indicators (WDI) database. Using the R programming language, a powerful tool for statistical analysis, the data was extracted using the WDI library package. The WDI package and R's functionality automated the data retrieval process, streamlining the collection of the economic and health indicators. The data collected for this research spans from the years 2015 to 2018, thereby adding depth and making more impactful through the analysis of multi-year data trends.

4 Data Understanding

Gaining knowledge about the data is a crucial step for a successful analysis. The data is meticulously explored using numerical and graphical summaries to gain insight on the basic trends and distribution of the data, followed by finding the missing values in the dataset. This is also essential as it sets the foundation of the analysis by providing evidence related to the relationships amongst the indicators and their individual trends.

4.1 Numerical Summaries

Table 1: Numerical Summary

Variable	Min	Q1	Median	Mean	Q3	Max	NA's
Infant Mortality rate	1.000	6.675	15.900	23.463	36.193	95.500	88
Immunisation coverage	19.00	84.21	92.00	87.72	96.00	99.00	100
Maternal mortality ratio	1.0	15.0	64.0	164.5	237.2	1288.0	132
Prevalence of undernourishment	2.50	2.50	6.25	10.25	14.10	58.20	192
GDP Per Capita	232.1	2125.3	6367.0	16335.9	19393.0	193968.1	32
inflation Rate	-3.749	0.901	2.272	4.613	4.272	380.000	141
Health expenditure per capita	17.71	81.58	337.16	1049.00	991.00	10284.56	125

4.1.1 Economic Indicators

From Table 1, GPD per capita in the dataset reveals disparities, ranging from \$232 to the maximum of \$193,968, highlighting the economic disparities between the nations. The median at \$6,367 indicates that half of these nations have an annual per capita economic output less than \$6,367, indicating the potential challenges in funding robust health systems. The median and average rates for the inflation are 2.272% and 4.613%, respectively, showing considerable variation. It also exhibits a dramatic range from a decrease of -3.749% to a spike of 380%. High inflation in a nation can erode purchasing power and make healthcare services more expensive and less accessible, specially to the lower section of the society, complicating efforts to improve public health.

The per capita health expenditure ranges from \$17.71 to \$10,284.56, with a median spending of \$337.16. This insight reveals how much country invests in health per person and the wide disparities suggests that health services and access can vary significantly depending on a country's economic wealth which further affects the health budget.

4.1.2 Health Indicators

The infant mortality is still a significant concern globally, particularly in economically weaker regions, as highlighted by the range of infant mortality rates in Table 1, which ranges from 1 to 95.5 deaths per 1,000 live births, with an average of 23.463 per 1,000 live births and the median at 15.9. These figures highlight the need for improved paediatric care and preventive health measures. Along with this the maternal mortality ratio's data paints a grim picture of maternal health measures. With a range from 1 to 1,288 maternal deaths per 100,000 live births and the average of 164.5, high ratios are indicative of inadequate maternal health services and a lack of access to necessary care, especially in the nations with the highest ratios.

The median immunisation coverage is very high at 92%, along with the immunisation coverage range of 19% to 99%. But the median of 87.72% points to significant under-immunised populations in some nations. This gap underscores the importance of sustained public health efforts and immunisation programs. Talking about the undernourishment prevalence, it demonstrates the challenges of ensuring food security, ranging from 2.5% to 58.2%, with a median of 6.25%. This shows that many countries still face many challenges in providing their populations with sufficient nutritional intake, directly impacting the overall health and economic productivity.

4.1.3 Missing Values

After carefully analysing the numerical summaries, from Table 1, it was found that there are numerous missing values present in the dataset. All the health and economic indicators have missing values but, the highest number of missing values is 192 for prevalence of undernourishment and lowest being in GDP per capita. These missing values were usually missing for the countries who are underdeveloped or are in the initial phase of development.

4.2 Graphical Summaries

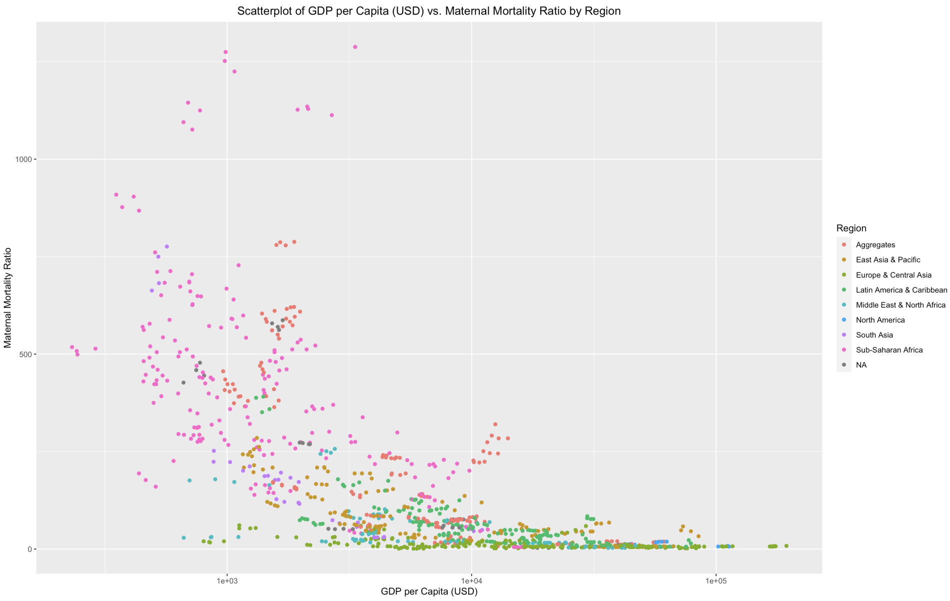

Figure 1: GDP per capita vs Maternal Mortality ratio Scatterplot

From Figure 1 it can be observed that there is a negative trend between GDP per capita and maternal mortality ratio which means as the economic output increase, the maternal mortality ratio decreases. The trend suggests that with every $1,000 increase in GDP per capita, maternal mortality ratio decreases by roughly 10 deaths per 100,000 live births. The scatter plot also shows the presence of outliers, especially in low-income regions where increase in GDP has not affected maternal health outcomes.

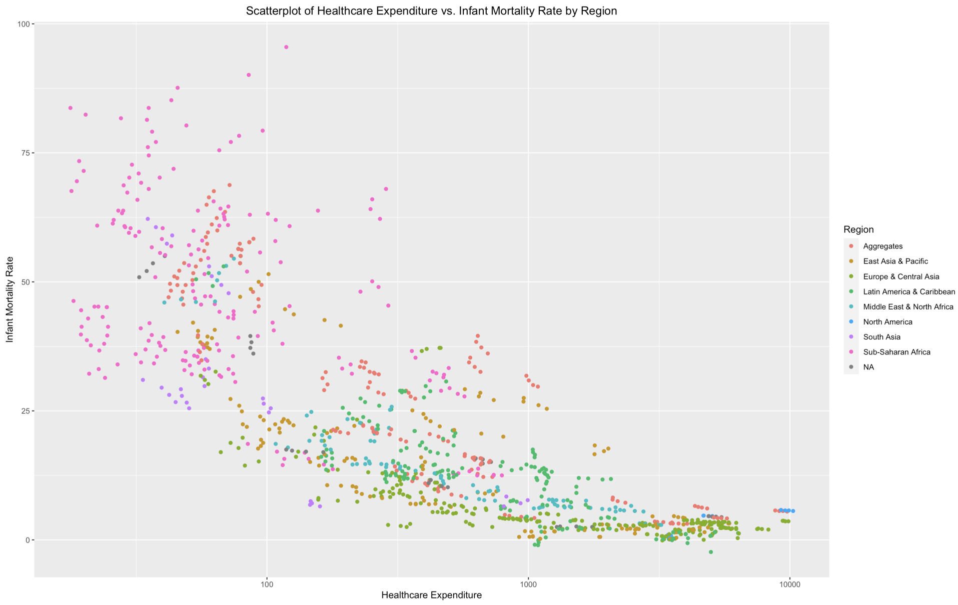

Figure 2: Healthcare Expenditure vs Infant Mortality rate Scatterplot

Just like the above relationship Figure 2 reveals that health expenditure and infant mortality rate follows the same downward trend. The inverse relationship reveals that for reducing the infant mortality, increased investments in health care is crucial, indirectly pointing towards having high GDP.

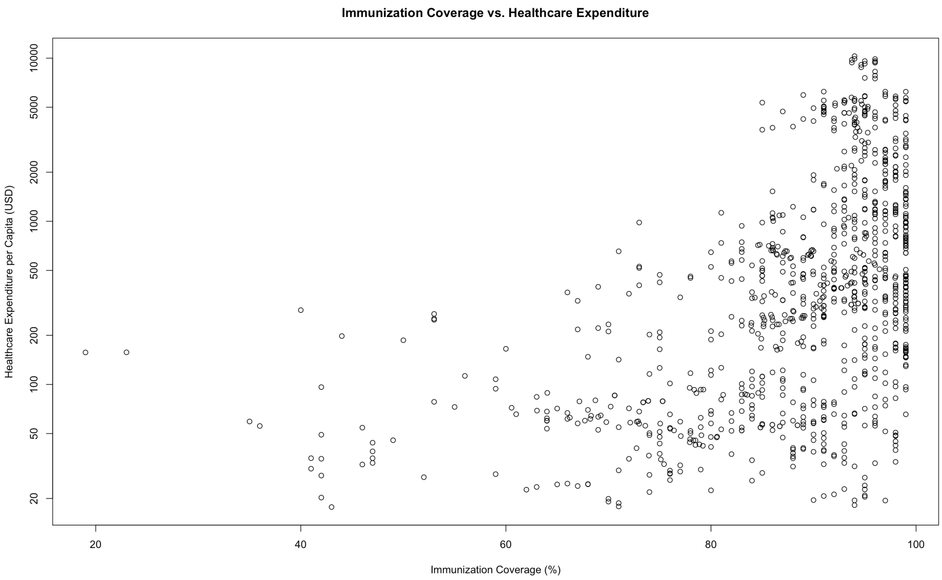

Figure 3: Healthcare Expenditure vs Immunisation Coverage Scatterplot

Figure 3 shows that healthcare spending per capita is directly proportional to immunisation rates, which means that higher investment in health leads to better immunisation outcomes, highlighting the value of healthcare expenditure in enhancing the public health infrastructure and preventing diseases.

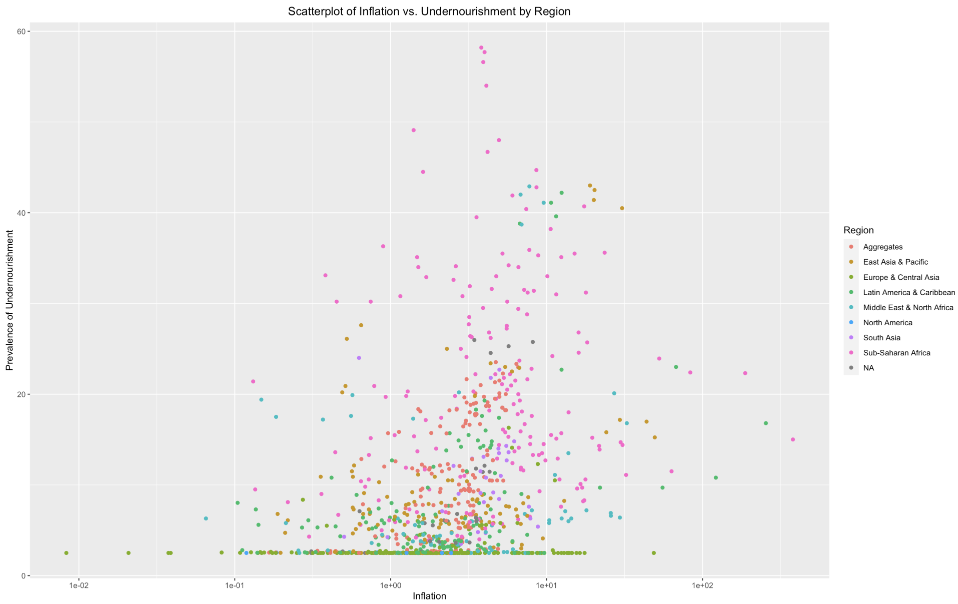

Figure 4: Inflation rate vs Prevalence of Undernourishment Scatterplot

Analysing data by region in Figure 4, reveals that increased inflation does not uniformly lead to lower undernourishment across all regions. While regions like North America and Europe show high inflation with low prevalence of undernourishment, areas like Sub-Saharan Africa exhibit less correlation. This might be due to other factors involved like governance, infrastructure, and socio-economic conditions.

With the help of numerical and graphical summaries, the foundation to find the connections between economic indicators and health indicators was established. The exploratory data analysis revealed that higher economic indicators generally correlate with the improved health metrics, disparities persist, requiring tailored economic and health policies.

5 Data Preparation

After the analysis of numerical summaries of the data, it was found that there are numerous missing values in the dataset, which made the data cleaning an essential step before modelling for robust and reliable analyses. In this phase, the data quality issues were addressed. The first step was to convert the 'country' variable into a factor and select the economical and health variables while excluding the non-essential ones such as geographical coordinates and iso codes.

For replacing the missing value to some meaningful value, data imputation was conducted using the Random Forest using 'missForest' library of R programming language. The choice of Random Forest totally depends on the way it imputes the missing values by predicting the missing values based on other variables in the dataset, capturing complex relationships between variables. On the other hand, replacing the values with mean, median, or mode may underestimate the variance and covariance in the data and introduce bias. KNN imputation was also another option instead of Random Forest, but it was dropped because it assumes that similar observations are likely to have similar missing values, which may not always true in real life. The imputation was done on a copy of dataset and then the imputed data was replaced by the missing values in the dataset where necessary, to keep the original data safe from any accidental changes.

By systematically addressing missing values, data integrity and reliability was uphold, a solid groundwork for the insightful analyses and informed decision-making in the field of public health and economic development was laid.

6 Modelling and Evaluation

To go to the root of the analysis about the effect of economic growth on health, it was essential to find how economic indicators affect the health metrics using modelling. In any real-world scenario, it is important to evaluate the non-linear relationships along with the linear relationships as it is not necessary that all the relationships are linear in nature. For this, two different types of models, Linear Regression and Generalised Additive Model (GAM) are utilised to explore the associations between the indicators.

6.1 Linear Regression Models

The linear regression model is employed as it is suitable for capturing linear relationships between the predictor variables (economic indicators) and health outcomes (health indicators). The various linear regression models provide insights into how changes in economic factors affect the health outcomes in a straightforward manner.

Table 2: Coefficients of Linear Regression Model for Infant Mortality Rate

Variable	**Estimate**	**Std. Error**	**t-value**	$Pr(>\|t\|)$	**Signf**
(Intercept)	3.064×10^{1}	1.061×10^{0}	2.888×10^{1}	$< 2 \times 10^{-16}$	***
NY.GDP.PCAP.CD	-5.463×10^{-5}	2.873×10^{-5}	-1.901	0.0575	.
FP.CPI.TOTL.ZG	2.182×10^{-2}	2.716×10^{-2}	8.030×10^{-1}	0.4219	
SH.XPD.CHEX.PC.CD	-1.939×10^{-3}	3.831×10^{-4}	-5.060×10^{0}	$< 5 \times 10^{-7}$	***
incomeHigh	-1.823×10^{1}	1.471×10^{0}	-1.239×10^{1}	$< 2 \times 10^{-16}$	***
incomeLow	2.179×10^{1}	1.659×10^{0}	1.313×10^{1}	$< 2 \times 10^{-16}$	***
incomeLower	1.042×10^{0}	1.384×10^{0}	7.530×10^{-1}	0.4518	
incomeNot classified	-1.227×10^{1}	7.317×10^{0}	-1.676	0.0940	.
incomeUpper middle	-1.351×10^{1}	1.362×10^{0}	-9.917×10^{0}	$< 2 \times 10^{-16}$	***

Table 2 reveals that health expenditure per capita is a significant predictor, with a coefficient of -0.0019, which means that for each unit increase in health expenditure per capita, IMR decreases by approximately 0.0019 units. Along with this income levels also play significant role in decreasing the IMR. The model's adjusted R-squared value of 0.5801 further tells that 58.01% of the variability in IMR is explained by the model suggesting a good fit to the data.

Table 3: Coefficients of Linear Regression Model for Immune Development

Variable	**Estimate**	**Std. Error**	**t-value**	$Pr(>\|t\|)$	**Signf**
(Intercept)	-70.37	65.87	-1.068	0.2856	
NY.GDP.PCAP.CD	0.01617	0.001784	9.065	$< 2 \times 10^{-16}$	***
FP.CPI.TOTL.ZG	-0.2026	1.686	-0.120	0.9044	
SH.XPD.CHEX.PC.CD	-0.006947	0.02379	-0.292	0.7703	
incomeHigh	336.30	91.35	3.681	0.000244	***
incomeLow	134.02	103.03	1.301	0.1936	
incomeLower middle	116.98	85.91	1.362	0.1736	
incomeNot classified	98.74	454.34	0.217	0.8280	
incomeUpper middle	68.49	84.59	0.810	0.4183	

As pointed by Table 3, the immunisation coverage positively correlates with GDP per capita with a coefficient of 0.016 highlighting that one-unit increase in GDP per capita leads to an increase of 0.016 units in immunisation coverage. Other economic indicators did not show significant effects, but this might be due to their non-linear relationship.

Table 4: Coefficients of Linear Regression Model for Maternal Mortality Ratio

Variable	**Estimate**	**Std. Error**	**t-value**	$Pr(>\|t\|)$	**Signf**
(Intercept)	2.194×10^{2}	1.244×10^{1}	1.763×10^{1}	$< 2 \times 10^{-16}$	***
NY.GDP.PCAP.CD	-2.112×10^{-4}	3.370×10^{-4}	-6.27×10^{-1}	0.531	
FP.CPI.TOTL.ZG	2.010	3.186×10^{-1}	6.310	4.15×10^{-10}	***
SH.XPD.CHEX.PC.CD	-1.585×10^{-2}	4.494×10^{-3}	-3.527	4.39×10^{-4}	***
incomeHigh	-1.499×10^{2}	1.726×10^{1}	-8.689	$< 2 \times 10^{-16}$	***
incomeLow	2.352×10^{2}	1.946×10^{1}	1.208×10^{2}	$< 2 \times 10^{-16}$	***
incomeLower middle	-2.067×10^{1}	1.623×10^{1}	-1.274	0.2031	
incomeNot classified	-2.850×10^{2}	8.583×10^{1}	-3.321	9.29×10^{-4}	***
incomeUpper middle	-1.583×10^{2}	1.598×10^{1}	-9.903	$< 2 \times 10^{-16}$	***

Inflation rate, healthcare expenditure and income levels are significantly associated with maternal mortality ratio (MMR) as seen in Table 4. The predictor variables successfully explains 47.7% of the variability in maternal mortality ratio.

Table 5: Coefficients of Prevalence of Undernourishment

Variable	**Estimate**	**Std. Error**	**t-value**	$Pr(>\|t\|)$	**Signf**
(Intercept)	1.088×10^{1}	5.379×10^{-1}	2.023×10^{1}	$< 2 \times 10^{-16}$	***
NY.GDP.PCAP.CD	-9.225×10^{-6}	1.457×10^{-5}	-6.33×10^{-1}	0.527	
FP.CPI.TOTL.ZG	-2.659×10^{-2}	1.377×10^{-2}	-1.931	0.0538	.
SH.XPD.CHEX.PC.CD	-7.108×10^{-4}	1.943×10^{-4}	-3.659	2.66×10^{-4}	***
incomeHigh income	-5.033×10^{0}	7.460×10^{-1}	-6.746	2.54×10^{-11}	***
incomeLow income	1.590×10^{1}	8.414×10^{-1}	1.890×10^{1}	$< 2 \times 10^{-16}$	***
incomeLower middle income	1.711×10^{0}	7.016×10^{-1}	2.439	0.0149	*
incomeNot classified	1.077×10^{1}	3.710×10^{0}	2.902	0.00378	**
incomeUpper middle income	-3.806×10^{0}	6.908×10^{-1}	-5.509	4.57×10^{-8}	***

The analysis from Table 5 reveals a highly significant negative correlation between health expenditure per capita and prevalence of undernourishment, revealing that for every unit increase in the health expenditure per capita leads to 0.0007108 decrease in the prevalence of undernourishment.

These relationships between economical and health indicators presents straightforward interpretations suggesting that as the economic indicators change linearly, they have a proportional impact on the health indicators. In summary, the linear regression models highlight that although all the economic indicators had associations with the health outcomes, the health expenditure per capita emerges as a consistent predictor across various health indicators, reflecting its importance in improving public health outcomes.

6.2 Generalised Additive Models (GAM)

After capturing the linear relationships, it was also important to focus on the complex non-linear relationships which exists amongst the indicators. For this, generalised additive models are used as they are known for their ability to capture potential non-linearities effectively which are overlooked by linear regression.

Table 6: Approximate Significance of Smooth Terms for Infant Mortality Rate

Term	**edf**	**Ref.df**	**F-value**	**p-value**	**Signf**
s(NY.GDP.PCAP.CD)	3.874	3.987	10.527	$< 2 \times 10^{-16}$	***
s(FP.CPI.TOTL.ZG)	1.000	1.000	4.019	0.0452	*
s(SH.XPD.CHEX.PC.CD)	3.893	3.990	12.293	$< 2 \times 10^{-16}$	***

The GAM analysis of infant mortality rate (IMR) reveals that there are significant complex non-linearities with GPD per capita and healthcare expenditure per capita. From Table 6, the estimated degree of freedom (edf) for GPD is 3.874 and 3.893 for health expenditure along with their F-value being 10.527 and 12.293 respectively, supporting the non-linear relationship claim. The adjusted R-squared value of 0.579 adds even more significance to the results, by explaining 57.9% of the variability in the IMR.

Table 7: Approximate Significance of Smooth Terms for Immune Index

Term	**edf**	**Ref.df**	**F-value**	**p-value**	**Signf**
s(NY.GDP.PCAP.CD)	3.690	3.927	46.927	$< 2 \times 10^{-16}$	***
s(FP.CPI.TOTL.ZG)	1.000	1.000	0.017	0.895	
s(SH.XPD.CHEX.PC.CD)	3.689	3.930	33.358	$< 2 \times 10^{-16}$	***

From Table 7, it can be easily observed that there exists a strong non-linear association with GPD per capita (edf: 3.690 & F-value: 46.927) and health expenditure per capita (edf: 3.689 & F-value: 33.358) for immunisation coverage.

Table 8: Approximate Significance of Smooth Terms for Maternal Mortality Ratio

Term	**edf**	**Ref.df**	**F-value**	**p-value**	**Signf**
s(NY.GDP.PCAP.CD)	3.868	3.985	8.239	1.57×10^{-6}	***
s(FP.CPI.TOTL.ZG)	3.900	3.993	16.080	$< 2 \times 10^{-16}$	***
s(SH.XPD.CHEX.PC.CD)	3.858	3.984	7.763	3.02×10^{-6}	***

For maternal mortality ratio (MMR), the GAM analysis shows significant non-linear relationships with GPD per capita and inflation rate. Their edf being 3.868 and 3.900 respectively, and their F-values being 8.239 and 3.900 respectively (From Table 8), indicating strong non-linearities between the economic indicators and MMR.

Table 9: Approximate Significance of Smooth Terms for Undernourishment

Term	**edf**	**Ref.df**	**F-value**	**p-value**	**Signf**
s(NY.GDP.PCAP.CD)	3.844	3.977	9.710	$< 2 \times 10^{-16}$	***
s(FP.CPI.TOTL.ZG)	2.590	3.086	2.787	0.03639	*
s(SH.XPD.CHEX.PC.CD)	3.736	3.953	4.438	0.00477	**

GAM analysis of prevalence of undernourishment reveals some interesting insights. The GDP per capita and healthcare show strong non-linear relationship with prevalence of undernourishment. But the most interesting insight which provides the justification to the visualisation of prevalence of undernourishment vs inflation rate is, that along with GPD per capita and healthcare expenditure, inflation rate also exhibits non-linear association with it, as it can be observed from Table 9.

GAM analysis unveiled significant non-linear associations between economic indicators and health outcomes, indicating that the relationships between these variables are not always straightforward and may exhibit complexities that require more nuanced modelling approaches. For instance, GDP per capita and health expenditure per capita demonstrated non-linear relationships with various health indicators, suggesting that the effects of economic factors on health outcomes may vary non-linearly across different income levels and healthcare spending patterns.

7 Conclusion

The thorough analysis has shed light on the complex interplay between economic growth and health outcomes. It was found that economic factors such as GDP per capita, inflation, and health expenditure per capita are closely linked to key health metrics. These relationships exhibit both linear and non-linear patterns, demonstrating the nuanced nature of the economic-health connection.

One of the main insights from the study is that higher economic growth, especially increases in GDP per capita and health expenditures, generally correlates with better health outcomes. This correlation is visible in the decrease of maternal and infant mortality rates and the increase in immunization coverage in areas where the economy is improving. However, the non-linear dynamics explored through

Generalised Additive Models (GAM) reveal that these benefits are not uniformly distributed across all socio-economic environments. This variability suggests that economic advancements don't automatically equate to health improvements across the board.

The role of inflation and different income groups also emerges as critical in shaping health outcomes. The findings indicate that high inflation can erode the affordability of healthcare, thus negating some of the potential health benefits of economic growth. This is a crucial consideration for policymakers, who need to implement strategies that buffer health systems from the adverse effects of inflation.

From a policy perspective, the evidence underscores the importance of targeted health spending. Policymakers must ensure that funds allocated to health are not only sufficient but are also directed towards the most impactful areas, such as maternal and child health services. Additionally, economic policies need to be crafted with an eye towards health equity, striving for growth that inherently supports health improvements.

Furthermore, the data suggest that health systems must be flexible and responsive to the economic conditions in which they operate. For example, during periods of high inflation, proactive measures are necessary to keep healthcare affordable and prevent a deterioration in public health standards. The sustained effort in public health programs, especially vaccination initiatives, is vital to address the gaps in immunization coverage observed in some regions.

In conclusion, while economic growth tends to support better health outcomes, the relationship is complex and influenced by a range of factors including inflation and income disparities. Effective policy-making requires a deep understanding of these intricacies to leverage economic growth for significant and equitable health benefits. This report lays the groundwork for informed policy-making, advocating for a strategic approach that synchronizes economic and health objectives to maximize benefits for all segments of society.