The University of Hong Kong

BSDS4999 Project

Project Research Proposal

Unraveling the Drivers of Global Poverty: A Comprehensive Analysis of Geographic and Social Factors

Student name and ID: Chan Yip Wan 3036141757, Lai Chun Kit 30361409233

Supervisor: Professor LIN, Jionghao

2. Table of Content

3. Abstract

The United Nations' 2030 Agenda for Sustainable Development listed poverty eradication as the first Sustainable Development Goal (SDG 1), which aims to eliminate all forms of poverty globally. To achieve such an objective, a deep understanding of the complicated and interconnected factors that lead to poverty, especially for those rooted geographically is required. This article aims to investigate the correlation between poverty rates and both geographic and social factors across different regions in the world. By analyzing data across the globe, the article discusses the impacts of different independent variables on poverty rates including geographically and socially. Statistical analyses are conducted to find the correlation and significance of the independent variables, and distinguish the factors that are the most impactful to a region’s poverty rate, and propose solutions to solve poverty. The research aims to provide insights for policymakers and NGOs as a reference for strategic poverty alleviation policies, to solve the roots of poverty in an effective and efficient way instead of only providing financial aid, so the goal of poverty elimination can be achieved.

4. Introduction

4.1. Background

The United Nations established 17 Sustainable Development Goals in the 2030 Agenda for Sustainable Development (United Nations, 2024). The first goal is to end all forms of poverty globally, which means it aims to eliminate all the people living under the poverty line, who live on less than $1.25 per day. In order to address the issue, geospatial poverty is a topic that cannot be ignored, as it is one of the roots of poverty. According to (World Economic Forum, 2023), the economic losses caused by weather, climate, and water extremes have tremendously increased from US$183.9 billion to 1476.2 billion from 1970 to 2019, and it stated that the real figures are likely to be higher since many losses are unreported. Therefore, regions that are often affected by extreme weather, climate, and natural disasters, are encountering economic losses, which will restrict their economic development and increase the poverty rate. Moreover, there are differences in economic development between urban and rural areas, coastal and inland areas, as well as resourceful and unresourceful areas. According to (Jin et al., 2023), In 2018, the GDP of all “near regions” in the world was $515.236 billion US dollars, while “far regions” was $302.380 billion. Due to the advantage of using shipping to import and export, coastal areas are more likely to develop a robust economy. In terms of the difference between rural and urban areas, according to the statistics of the World Bank (CollegeNP, 2023), urban areas were only 2% of the total land area on Earth but occupied 80% of the global GDP in 2020. Therefore, it can be seen that the economy of urban areas is far superior to urban areas. In terms of the number of natural resources, countries that contain large amounts of natural gas and oil can export to other

countries and build a strong economy, such as the Middle Eastern countries, which rely so much on the export of oil for their economy.

Besides geographical factors, social factors are one of the crucial determinants of geospatial poverty including healthcare, education, infrastructure, tax rate, social welfare, and corruption. According to the report from (OECD, 2024), educational attainment is strongly correlated with labour market participation in terms of employment, unemployment or inactivity rates. Among all countries within the OECD, the average employment rate of people in age 25-34 with no upper secondary qualification is about 60%, while the average employment rate of those with a college degree is 87%. Besides that, the average inactivity rate of those with no upper secondary qualification is 31%, while it is 9% for those with a college degree for the same age group. And the report stated that this correlation between educational level and labour market participation is the same for both genders and still remains the same over decades. Apart from education, healthcare is being generally considered as a factor that has an impact on economic development, because a complete healthcare system secures the health of citizens and their ability to work, which ensures their productivity. The above mentioned factors are always the indicators of developed countries. It is believed that they are both the byproducts and the causes of a developed economy as rich regions can have higher government expenditures on the above social factors, but doing well on the above social factors is beneficial for the economic development either.

4.2 Connection with UN's SDG

According to the agenda (United Nations, 2024), target 1.1 of goal number one, "By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than $1.25 a day", also target 1.4, "By 2030, ensure that all men and women, in

particular the poor and the vulnerable, have equal rights to economic resources, as well as access to basic services, ownership and control over land and other forms of property, inheritance, natural resources, appropriate new technology and financial services, including microfinance", and target 1.b, "Create sound policy frameworks at the national, regional and international levels, based on pro-poor and gender-sensitive development strategies, to support accelerated investment in poverty eradication actions". To achieve the total elimination of poverty, and giving everyone the equal rights to economic resources, geospatial poverty needs to be solved. It is extremely hard for people to get rid of poverty while living in regions with geographic and social factors that restrict economic development. The findings and recommendations of the research will give insights for constructing policy frameworks to eradicate poverty.

4.3 Significance and Purpose of the Study

To eliminate poverty across the globe, it is a must to investigate what factors are restricting the economic development of certain regions and find solutions to tackle the problem instead of only giving financial assistance to them. There is not a lot of research on the comparison of different factors of geospatial poverty, so it is important to evaluate the impacts of different factors and focus on solving those that affect the most. The research aims to quantify the contribution of different factors towards geospatial poverty by finding the correlation between the two to investigate the most crucial factors and suggest solutions to tackle the issues.

4.4 Research Questions

1. Do geographic and social factors correlate with poverty rates? If yes, to what extent?

2. What policies should be released in various regions to overcome the poverty caused by geographic and social factors?

5. Literature Review

5.1 Research Hypotheses

For geographical factors, the research of (Liu et al., 2017) discovered that ecologically fragile areas are usually poor areas in China, there are more than 40 million people living under the Chinese national poverty line in rural areas across the nation, while (Zhou and Xiong, 2018) suggested that the complexity of topography is shaping the spatial distribution of poverty in China as the poor regions in China are mainly located at the Northeast, where are under difficult mountainous topographical conditions. The finding of (Gallup et al., 1998) shows similar result, it shows that almost all of the countries in the geographical tropics are poor due to the limitation of agriculture, in contrast, almost all of the rich countries are in the mid and high latitudes, it also found that coastal areas and regions linked to ocean-navigable waterways are usually richer than the landlocked areas, but it suggested there is a weak correlation between population density and economic development. (Scherr, 2000) investigated how soil erosion impacts agricultural land quality, creating a vicious cycle of ecological deterioration, pauperization, farmland loss, soil erosion, and steep slope reclamation. Also, natural disasters perpetuate poverty and make it more difficult for poor people to escape from poverty, as stated by (Rozenberg and Hallegatte, 2016) and (Akter and Mallick, 2013). For example, due to natural disasters, approximately 26 million people worldwide fall into extreme poverty every year (Hallegatte et al., 2017; Rozenberg and Hallegatte, 2016). By 2030, some 325 million extremely poor people are projected to live in the 49 most hazard-prone countries worldwide, with the majority being located in South Asia

and sub-Saharan Africa (Shepherd et al., 2013). However, it is surprised that (Skidmore & Toya, 2002) argued that natural disasters increase the growth of GDP, they suggested that natural disasters are able to stimulate new technology adoption and update capital stock, causing the economies have a better development, but (Botzen et al., 2019) suggested that Skidmore's research is potentially having variable bias.

For social factors , in the research of (Castaño et al., 2015), it is discovered that economic freedom and the rule of law are the indicators of entrepreneurial activity stimulation, and it is suggested that social factors are more correlated with entrepreneurial intentions in European countries than Latin American and the Caribbean. (Akkoyunlu & Ramella, 2018) investigated that the control of corruption surely has a positive impact on productivity, innovation and income per capita. In the findings of (Srinivasu, B., & Rao, P. S., 2013), economic and social infrastructures are essential for economic growth, but it cannot be done only on its own. In terms of education, (Hanushek & Wößmann, 2010) have found that the quality of education is strongly correlated to the performance of economies as economic activities are highly dependent on the skills of workers. As well as influential research from (Barro, R. J. ,2013), he found that economic growth is positively related to the years of schooling at secondary or dietary education, which aligns with the the report from (OECD, 2024), Across the OECD countries, people with a higher education attainment have a higher income. People who work full time with an upper secondary qualification have an income that is on average 18% higher than those who do not, and the situation is even more obvious among people with a tertiary education qualification, on average, they have an income that is almost double higher than those with a qualification below secondary. Apart from the above, the research of (GATAWA, 2022) found that healthcare and education expenditure have a strong correlation with the GDP and GNI (or GNP) of a region. The conducted linear regression models in the research

show the p-value of less than 0.0001 of healthcare expenditure in both GNI and GNI per capita, and the p-value of 0.0051 for GNI and less than 0.0001 for GNI per capita in education expenditure.

Null Hypothesis (H_0): No correlation between geographic/social factors and poverty rates.

Alternative Hypothesis (H_1): There is a correlation between geographic/social factors and poverty rates.

6. Research Methods

6.1 Research Design

This study aims to investigate the relationships between Gross National Income (GNI), Worldwide Governance Indicators (WGI), and various geographic and social factors. By utilizing robust datasets from reputable sources such as the World Bank and UNESCO, we seek to provide a comprehensive analysis of how governance quality and regional classifications influence economic performance.

Data Collection:

GDP Data:

Global GDP data from the World Bank (Our World in Data, 2025) was collected, covering the years from 1990 to 2023. This dataset represents the total value of goods and services produced within a country's borders during a specific period. By analyzing the entire dataset rather than focusing solely on 2023, we aimed to mitigate the influence of short-term economic fluctuations and provide a clearer view of long-term trends.

GNI Data:

Global GNI data, also sourced from the World Bank (2025), encompasses total income earned by residents, including income from abroad. Similar to the GDP analysis, we calculated the average GNI over the available years to ensure a comprehensive understanding of income distribution across countries.

Disaster Data:

Data on mass disasters from the Centre for Research on the Epidemiology of Disasters (2025) was utilized, detailing over 26,000 disaster events worldwide. This dataset includes various attributes such as disaster type, event counts, and economic losses. We merged this data with GNI and GDP to analyze the economic impacts of disasters across different regions.

Education Expenditure Data:

Government expenditure on education as a percentage of GDP from 2000 to 2023 was obtained from UNESCO (UIS Data Browser, 2025). This dataset enables exploration of the relationship between education funding and economic performance.

Worldwide Governance Indicators:

WGI data, which encapsulates perceptions of governance quality, was collected from the World Bank (2024). We focused on the percentile rank of each country, indicating governance quality from 0 (lowest) to 100 (highest). Average WGI scores were calculated to provide a comprehensive view of governance trends.

Data Processing

Data Cleaning:

All datasets were meticulously cleaned to remove missing values and ensure data integrity. This process involved filtering out incomplete records to maintain the robustness of the analysis.

Data Integration:

Multiple CSV files were combined into single DataFrames to facilitate analysis. For instance, GDP and GNI data were merged with relevant variables such as income groups and regions, creating a comprehensive dataset for further examination.

Analytical Methods

Descriptive Analysis:

Initial exploratory data analysis was conducted to summarize key statistics and visualize distributions of GDP, GNI, and governance indicators. Bar charts, density plots, and scatter plots were utilized to illustrate relationships and distributions across different countries and regions.

Correlation Analysis:

Spearman's Rank Correlation coefficient was calculated to assess the strength and direction of relationships between variables, particularly GNI and GDP.

Regression Analysis:

Simple linear regression was employed to examine the relationship between GNI and WGI, followed by a multiple regression model to explore how WGI and regional factors predict GNI. The significance of predictors was evaluated using p-values and R-squared values to assess model fit.

Chi-Square Test:

The Pearson Chi-Square test was utilized to examine the association between income groups and regions, providing insights into the distribution of income categories across geographic contexts.

Visualization

Various plots were generated to visually represent the findings, including bar charts, density plots, scatter plots, and coefficient plots. These visualizations aid in interpreting complex data relationships and enhancing the overall understanding of economic disparities and governance quality.

6.2 Research Procedure

In this study, we obtained global GDP (Gross Domestic Product) data from the World Bank (Our World in Data, 2025), covering the years from 1990 to 2023. GDP represents the total value of all goods and services produced within a country's borders during a specific period, serving as a key measure of economic activity and output. Rather than focusing solely on the data from 2023, we opted to analyze the entire dataset to calculate the average GDP. This approach mitigates the influence of short-term factors, such as economic shocks and policy changes, which can skew results when examining a single year. By utilizing all available data, we aim to provide a more comprehensive view of long-term economic trends while reducing the impact of outlier years and short-term fluctuations. Additionally, we removed all missing data to ensure the integrity of our analysis. As illustrated in Figure 1, the DataFrame of GDP is presented in descending order, highlighting that the top three countries with the highest GDP are Luxembourg (LUX), Qatar (QAT), and Bermuda (BMU).

	Code	Mean_GDP
109	LUX	119833.39
152	QAT	103262.63
25	BMU	94768.08
29	BRN	94260.46
111	MAC	92005.94
159	SGP	86032.92
7	ARE	82842.93
134	NOR	79486.02
46	CYM	77871.81
34	CHE	71727.36
163	SMR	70137.56

Figure 1: DataFrame of GDP

We obtained global GNI (Gross National Income) data from the World Bank, covering the years up to 2023 (World Bank, 2025). GNI measures the total income earned by a country's residents, including income from abroad, divided by the population. It serves as a more effective indicator of poverty than GDP, as it accounts for income generated both domestically and internationally, providing a clearer understanding of the resources available to the population. Rather than relying solely on data from 2023, we utilized the entire dataset to calculate the average GNI, ensuring a more comprehensive analysis. To facilitate this, we combined two CSV files into one DataFrame, incorporating key variables: income group and average GNI for each country. As shown in Figure 2, we merged the DataFrames of GDP and GNI into a single DataFrame for further analysis, which includes the ISO code, average GDP, average GNI, income group, and region for each country.

	Code	Mean_GDP	Mean_GNI	IncomeGroup	Region
1	ABW	37577.873	21171.3889	High income	Latin America & Caribbean
2	AFG	2393.391	437.7273	Low income	South Asia
3	AGO	7709.619	1621.2821	Lower middle income	Sub-Saharan Africa
4	ALB	9967.503	2666.5000	Upper middle income	Europe & Central Asia
5	AND	57195.717	45942.5000	High income	Europe & Central Asia
6	ARE	82842.932	34296.9231	High income	Middle East & North Africa
7	ARG	24377.324	5747.6667	Upper middle income	Latin America & Caribbean
8	ARM	9362.556	2507.1875	Upper middle income	Europe & Central Asia

Figure 2: DataFrame of GDP and GNI

According to the World Bank, in 2023, low-income countries are defined as those with a GNI per capita of $1,145 or less; lower-middle-income countries fall between $1,146 and $4,515; upper-middle-income countries range from $4,516 to $14,005; and high-income countries are those with a GNI per capita exceeding $14,005 (Roser et al., 2023). The bar chart in Figure 3 illustrates the distribution of countries by income group as defined by the World Bank in 2023. High-income countries represent the largest portion, followed by upper middle-income nations. Lower middle-income countries show a moderate count, while low-income countries constitute the smallest group.

The distribution of countries by income group reveals critical insights into global economic inequalities. High-income countries dominate the chart, indicating a substantial concentration of wealth and resources, while low-income countries represent a much smaller segment. This disparity is particularly concerning given that low-income countries, defined as having a GNI per capita of $1,145 or less, face persistent economic challenges. The relatively small number of low-income countries suggests that targeted policies are essential to uplift these nations and address the systemic factors contributing to poverty. Additionally, the presence of lower middle-income countries highlights the ongoing struggle for many nations striving to improve their economic status. Understanding these dynamics is vital for formulating effective strategies to promote equitable growth and reduce poverty on a global scale.

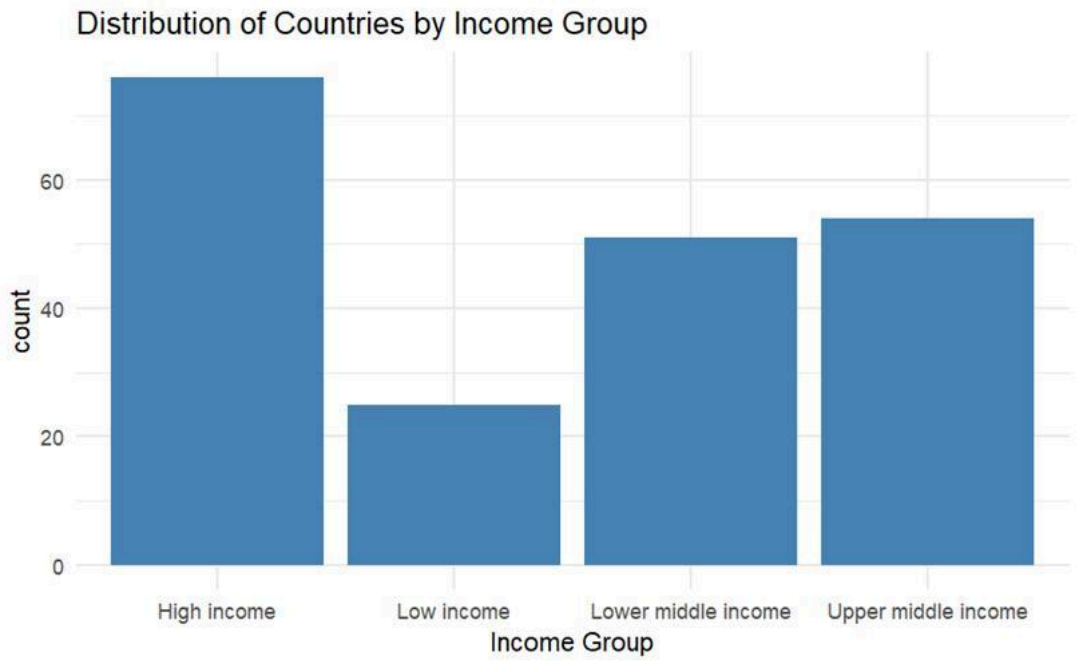

Figure 3: Bar chart of Distribution of Countries by Income Group

The density plot in Figure 4 demonstrates the distribution of GNI (Gross National Income) in current USD across countries, characterized by a pronounced right-skewed distribution. The red dashed line represents the median GNI, highlighting that a significant proportion of countries have GNI values below this median. Most countries cluster around lower income levels, with very few reaching the higher GNI thresholds.

This density plot illustrates the substantial income disparities present in the global economy, with the median indicating that more than half of the countries have a GNI below this level. The concentration of GNI values at lower levels suggests that many nations face significant economic challenges, which may perpetuate cycles of poverty. The right skew of the distribution further emphasizes the small number of countries with very high GNI, indicating that wealth is not evenly distributed. Understanding this distribution is crucial for policymakers, as it highlights the need for targeted interventions aimed at boosting the incomes of lower GNI countries and addressing the structural factors that contribute to economic inequality.

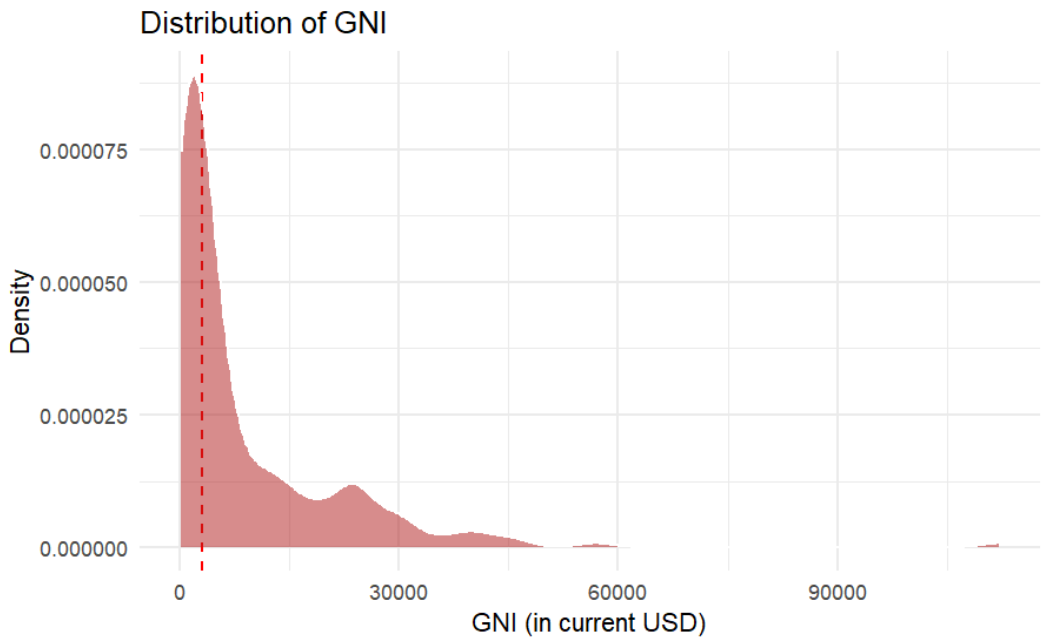

Figure 4: Density Plot of Distribution of GNI

The scatter plot (Figure 5) illustrates a robust positive correlation between GNI and GDP, with a Spearman's Rank Correlation coefficient of 0.868065 indicating a strong association. This relationship suggests that countries with higher economic output (GDP) tend to provide greater income for their residents (GNI). High-income countries are predominantly clustered along the upper end of the plot, while low- and lower-middle-income countries are more dispersed across the lower GNI and GDP values.

However, the dispersion of low- and lower-middle-income countries reveals a critical point: while these nations may have moderate GDP, their GNI does not always reflect similar income levels. This disparity could be attributed to factors such as high levels of income inequality, reliance on foreign investments, or remittances that do not translate into local economic growth. The clustering of high-income countries at the upper end emphasizes the concentration of wealth and economic stability in these nations. Understanding this relationship is crucial for policymakers, as it highlights the need for strategies that not only drive economic growth but also ensure that this growth translates into increased income for all citizens, thereby addressing poverty and inequality.

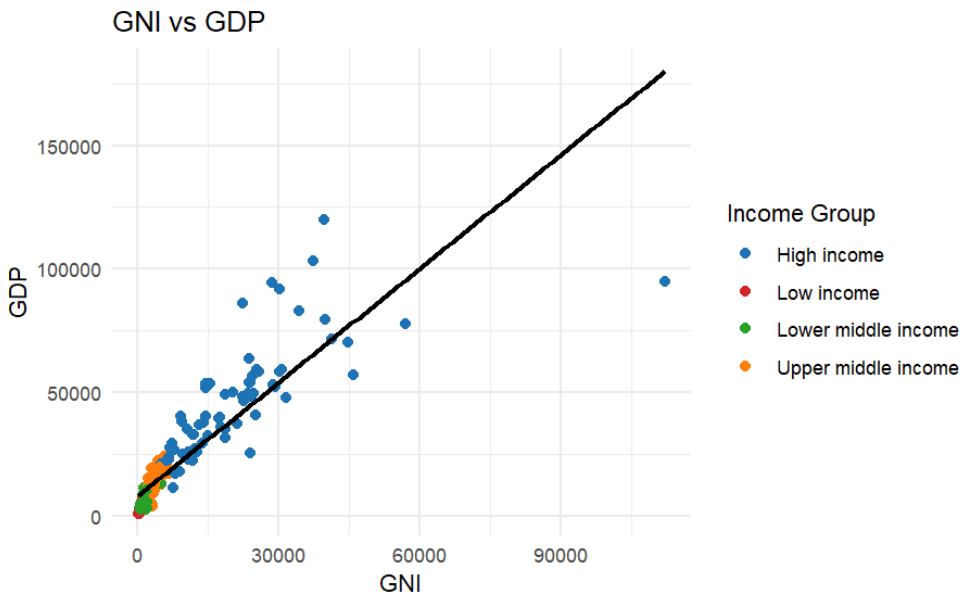

Figure 5: Scatter plot for GNI and GDP

In examining geographic factors, the first aspect we considered was the regional classification of countries. According to the dataset provided by the World Bank (2025), the world is divided into seven regions: East Asia & Pacific, Europe & Central Asia, Latin America & Caribbean, Middle East & North Africa, North America, South Asia, and Sub-Saharan Africa. To assess the association between income groups and regions, we employed the Pearson Chi-Square test, which yielded a p-value smaller than 2.2e-16, indicating a statistically significant relationship.

The bar plot in Figure 6 presents the average GNI by region, revealing that North America (NA) has a significantly higher average GNI compared to other regions. The Middle East & North Africa (MENA) follows at a lower level, while East Asia & Pacific (EAP), Europe & Central Asia (ECA), and Latin America & Caribbean (LAC) show moderate GNI values. South Asia (SA) and Sub-Saharan Africa (SSA) exhibit the lowest average GNI among the regions.

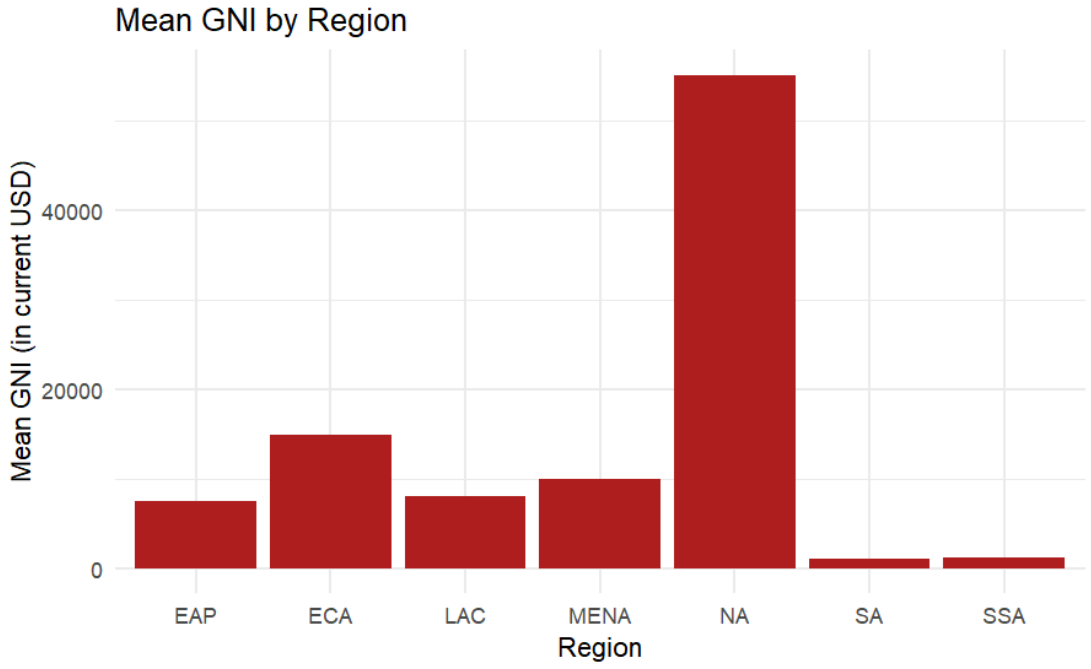

Figure 6: Bar plot of Average GNI by Region

Figure 7 presents the distribution of income groups across the same regions. It shows that North America has a predominantly high-income population, while regions like Sub-Saharan Africa and South Asia have a higher proportion of low-income and lower-middle-income categories. The distribution in Europe & Central Asia (ECA) and Latin America & Caribbean (LAC) reflects a more balanced mix of income groups, including significant upper-middle-income populations.

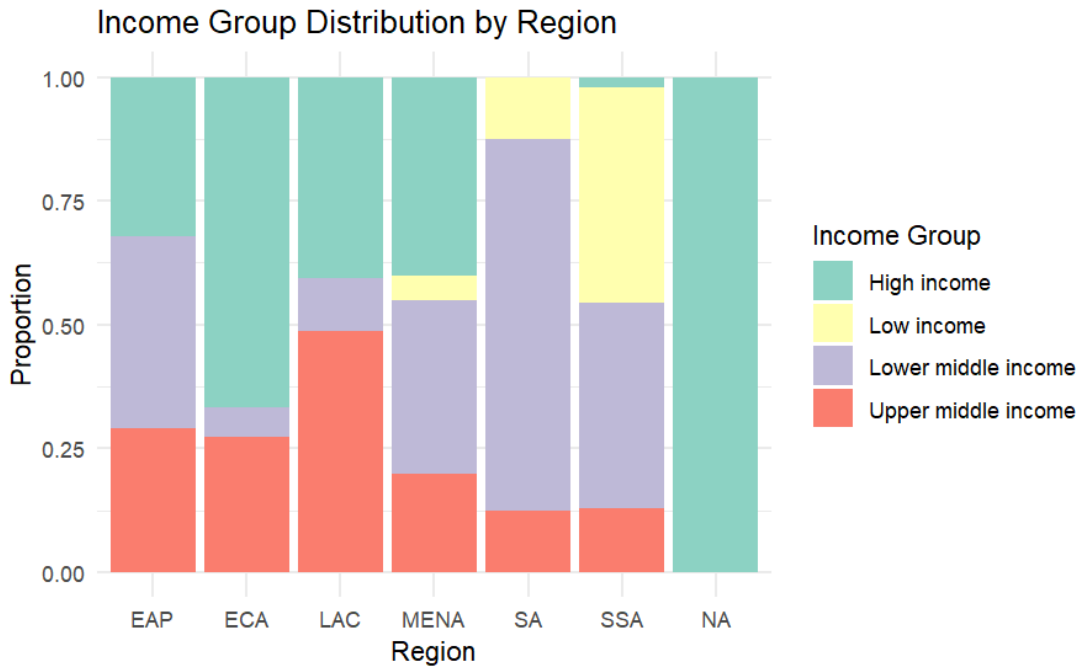

Figure 7: Bar plot of Income Distribution by Region.

The findings from Figure 6 and Figure 7 underscore significant economic disparities across regions as reflected in average GNI. North America's outstanding average GNI highlights its economic strength and prosperity, allowing for higher living standards and better access to resources. This elevated income level correlates with its classification as a high-income region.

Conversely, regions like South Asia and Sub-Saharan Africa face considerable economic challenges, as indicated by their low average GNI. The higher proportions of low-income and

lower-middle-income groups in these regions suggest persistent poverty and limited economic opportunities. Factors such as inadequate infrastructure, lack of access to quality education, and systemic inequalities contribute to these challenges, hindering overall development.

The varied income distribution in Europe & Central Asia and Latin America & Caribbean indicates a more complex economic landscape. These regions contain a mix of income groups, suggesting that while some countries may experience growth and prosperity, others may lag behind, experiencing economic difficulties.

The statistically significant relationship identified through the Pearson Chi-Square test further emphasizes the importance between regional context and income disparities. With a p-value smaller than 2.2e-16, it is clear that income groups are closely associated with geographic regions. The income levels are not uniformly distributed across different regions, indicating underlying economic disparities that are influenced by geographic factors.

By addressing the unique challenges faced by different regions, policymakers can work towards fostering equitable economic growth, improving living standards, and reducing poverty on a global scale. Understanding these dynamics is essential for developing effective strategies that promote inclusive economic development tailored to the specific needs of each region.

Another geographic factor we examined is the impact of disasters. The Centre for Research on the Epidemiology of Disasters (2025) reports that their dataset encompasses information on over 26,000 mass disasters worldwide from 1900 to the present. This comprehensive database is compiled from diverse sources, including UN agencies, non-governmental organizations, reinsurance companies, research institutes, and press agencies, and is made available in open access for non-commercial use. It includes various attributes such as disaster type, disaster group, event counts, total individuals affected, total fatalities, and economic losses. For our analysis, we summarized all records up to 2025 into a single DataFrame organized by each country's ISO code. This consolidated DataFrame was then merged with average GNI and GDP data to enable more in-depth analysis.

	ISO	Total_Event	Total_Affected	Total_Deaths	Total_Damage	Mean_GDP	Mean_GNI
1	AFG	155	34586968	13729	236461019	2393.391	437.7273
2	AGO	51	8985951	963	20180787	7709.619	1621.2821
3	ALB	25	1130370	797	928904522	9967.503	2666.5000
4	ARE	3	534	4	0	82842.932	34296.9231
5	ARG	84	2781373	525	12127150286	24377.324	5747.6667
6	ARM	10	416615	5	257635686	9362.556	2507.1875
7	ATG	3	32600	1	328375041	25733.289	10874.2222
8	AUS	127	741146	909	70468397617	48266.534	24187.4194
9	AUT	30	64479	1336	9409606215	56694.099	24396.1290
10	AZE	9	132293	50	285740982	13315.557	3101.8750

Figure 8: DataFrame containing Disaster Data, Average GDP and GNI

The bar plot in Figure 9 illustrates the distribution of disaster events categorized by disaster type. Hydrological disasters, such as floods and landslides, account for the highest frequency of occurrences, followed by meteorological events, including storms and extreme temperatures. Climatological disasters, such as droughts, also show significant counts, while biological, geological, and extra-terrestrial events are considerably less frequent. This distribution highlights the predominant types of disasters affecting various regions.

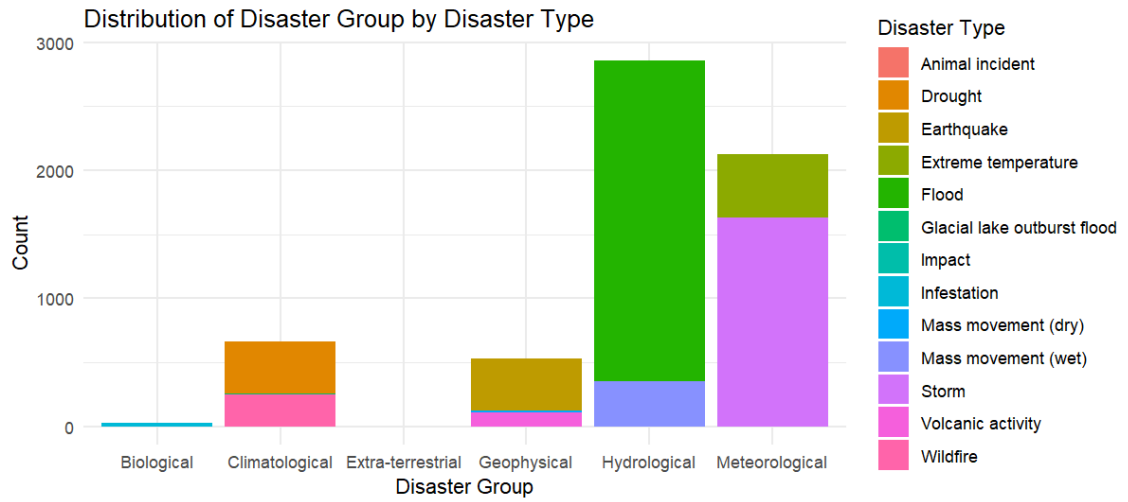

Figure 9: Bar plot of Event Frequency by Disaster Group.

The bar plot in Figure 10 displays the total damage caused by various disaster types, measured in USD. The data reveals that certain disasters, such as tropical cyclones and ground movement, result in the highest total damages, reaching into the trillions of dollars. In contrast, other disaster types, like storm surge and collision, show significantly lower damage totals. This stark contrast highlights the economic impact associated with specific disaster types.

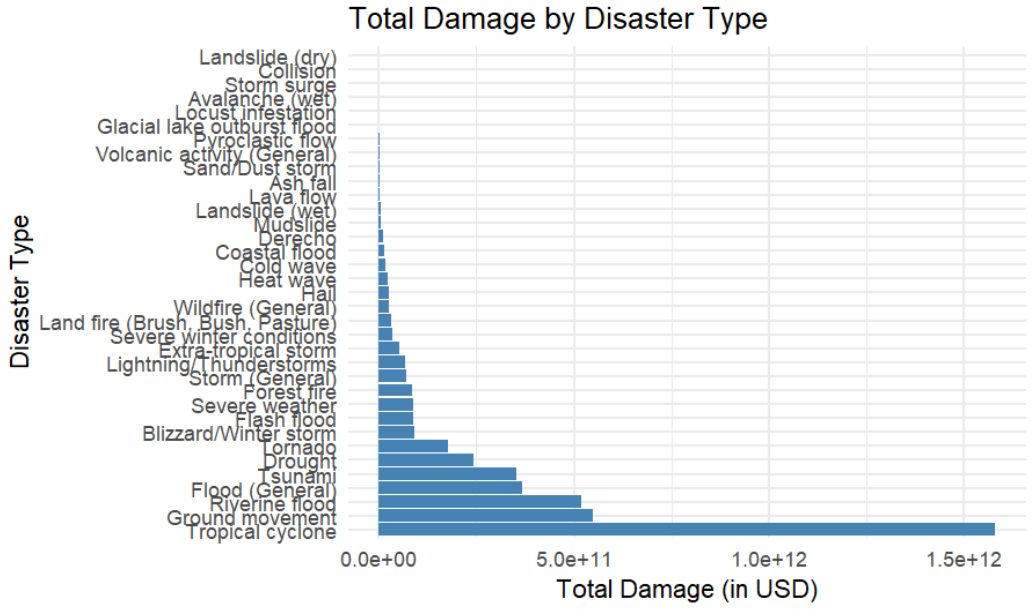

Figure 10: Bar plot of Total Damage by Disaster Type

The analysis of total damage by disaster type underscores the severe financial implications of natural disasters. The prominence of tropical cyclones and floods in terms of total damage indicates that these events not only occur frequently but also have devastating economic consequences. This is particularly relevant in regions that are vulnerable to severe weather conditions, where infrastructure and livelihoods are often at risk.

The lower damage totals associated with less frequent disasters, such as geological events and biological incidents, suggest that while these disasters can be impactful, their overall economic toll is less pronounced. Understanding the financial implications of different disaster types is crucial for better risk assessment and resource allocation in disaster management.

The findings from both Figures 9 and 10 together paint a comprehensive picture of the disaster landscape, revealing that while hydrological and meteorological disasters are frequent, their potential for causing significant economic damage is substantial.

The correlation heatmap in Figure 11 presents the relationships between various disaster metrics—such as total events, total affected, total deaths, and total damage—and the average GNI. The strongest correlation is observed between total events and total affected (0.68), indicating that as the number of disaster events increases, the number of people affected also tends to rise. The correlation between average GNI and total damage is very weak (0.14), while the correlation between GNI and other disaster metrics is even lower.

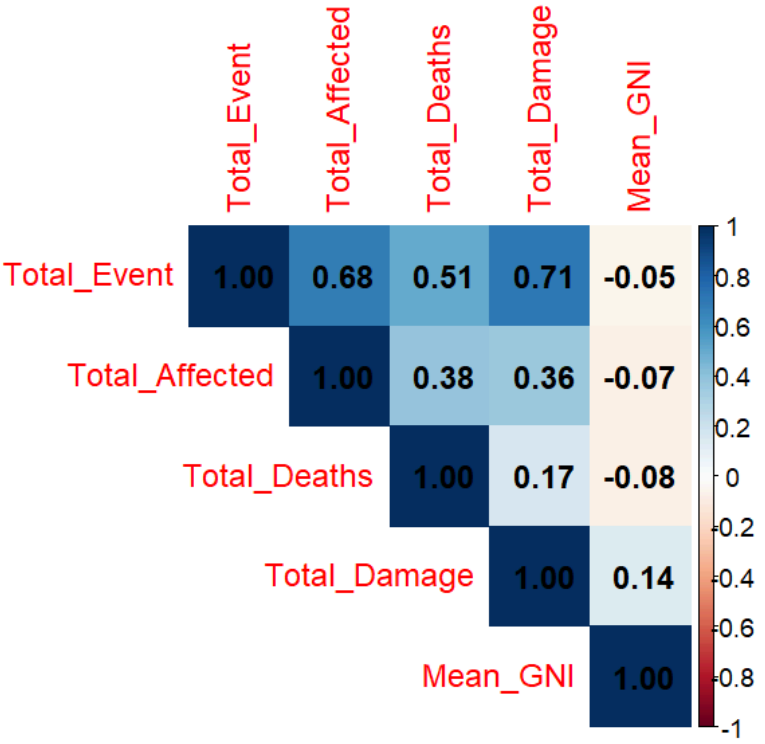

Figure 11: Correlation Heatmap by Disaster Data and GNI

The findings from the correlation heatmap reveal important dynamics between economic factors and disaster impacts, particularly concerning average GNI. The strong correlation between total events and total affected (0.68) underscores the direct relationship between the frequency of disasters and the number of individuals impacted. This suggests that as disasters

become more frequent, communities are increasingly at risk of suffering significant consequences, regardless of their economic status.

The low correlation between average GNI and total damage (0.14) suggests that higher income does not always equate to greater economic losses from disasters. This could imply that wealthier countries may have better infrastructure and resources to absorb the impacts of disasters, potentially leading to less reported damage compared to less affluent nations.

In addition, we conducted the ANOVA analysis to assess the association between income group and total damage, resulting in the p-value of 0.34. Since it is well above the conventional significance level of 0.05, we fail to reject the null hypothesis. It indicates that there is no statistically significant difference in total damage among the different income groups. This finding implies that differences in total damage incurred by disasters are not meaningfully associated with income group classifications.

Furthermore, the analysis indicates that poverty does not exhibit a strong relationship with disaster impacts when considering geographic factors. Geographic location can significantly influence vulnerability to disasters, often overshadowing economic indicators like GNI. For instance, regions prone to natural hazards may experience severe consequences regardless of their income levels. Factors such as population density, infrastructure resilience, and preparedness play crucial roles in determining the extent of disaster impacts. Thus, while GNI provides insight into economic capacity, it is not a definitive predictor of disaster outcomes, emphasizing the importance of geographic context in understanding and addressing disaster risks.

In examining the social factors, we utilized a dataset detailing government expenditure on education as a percentage of GDP from 2000 to 2023, provided by the United Nations Educational, Scientific and Cultural Organization (UNESCO) (UIS Data Browser, 2025). Instead of focusing solely on the most recent data from 2023, we calculated the average expenditure across the entire dataset. This approach allows for a more comprehensive analysis, reducing the influence of short-term fluctuations. As illustrated in Figure 12, the DataFrame presents the ISO codes alongside the average government expenditure on education as a percentage of GDP, organized in descending order. Notably, the countries with the highest expenditure percentages are Tuvalu (TUV), American Samoa (ASM), and Kiribati (KIR).

geoUnit	Mean_Education
TUV	14.775506
ASM	14.717050
KIR	12.834712
FSM	11.533602
MHL	11.086387
CUB	10.123635
LSO	9.212743
BWA	8.745798
NAM	8.430945
SLB	7.890265

Figure 12: DataFrame of Average Government Expenditure on Education as a Percentage of GDP

The density plot in Figure 13 illustrates the distribution of education expenditure as a percentage of GDP across various countries. The peak of the distribution is concentrated around 4% to 6%, indicating that most countries allocate a moderate level of their GDP to education. The plot also reveals a long tail extending towards higher expenditure percentages, suggesting that while the majority of countries spend within this moderate range, a smaller number allocate significantly higher percentages.

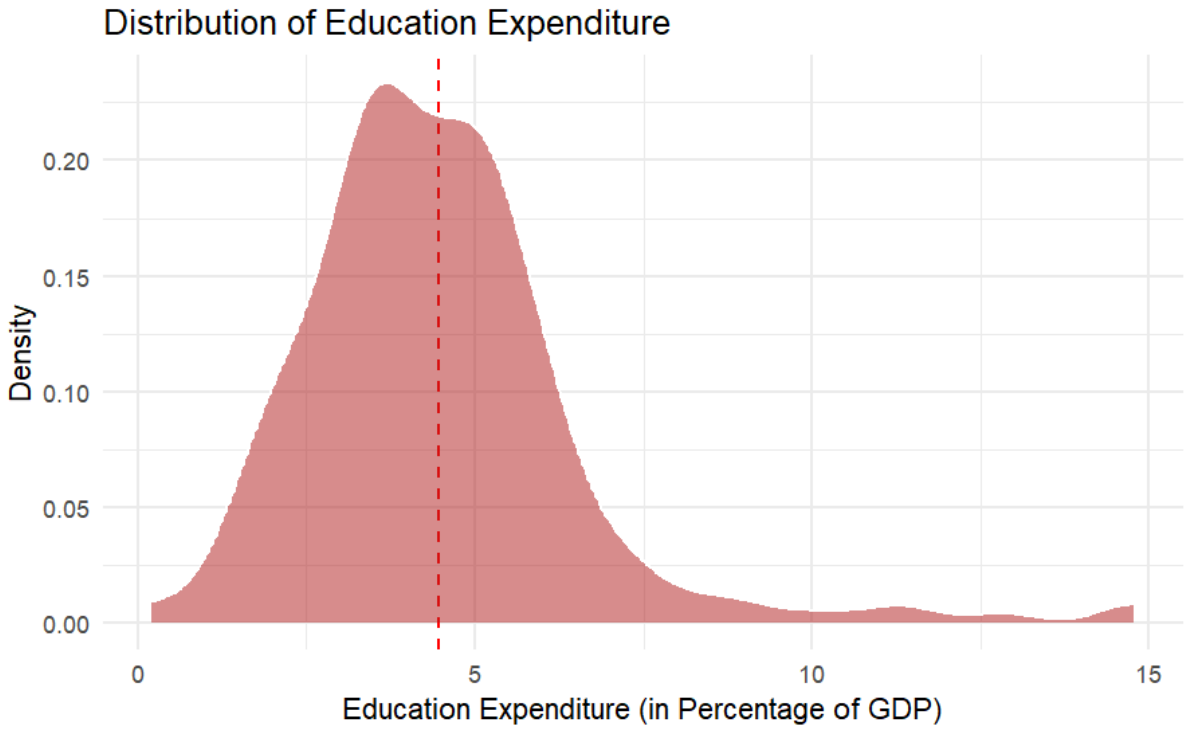

Figure 13: Density plot of Education Experiture Distribution

The density plot provides insightful perspectives on global education expenditure patterns. The concentration of data around 4% to 6% indicates a common threshold for many countries, reflecting a general consensus on the level of investment deemed necessary for education. This range may represent a balance between educational needs and economic constraints faced by governments.

However, the presence of a long tail in the distribution highlights disparities in education funding. A small subset of countries allocates a much larger percentage of their GDP to education, which could be indicative of different national priorities, economic capacities, or educational policies. Countries at the higher end of the expenditure spectrum may be investing heavily in their educational systems to enhance quality and accessibility, suggesting a commitment to long-term socio-economic development.

The scatter plot in Figure 14 illustrates the relationship between Gross National Income (GNI) and education expenditure as a percentage of GDP across various countries. The correlation coefficient calculated is approximately 0.0055, indicating a very weak positive correlation between the two variables. Most data points cluster around lower values of education expenditure, with a few countries reporting higher expenditure, but no clear linear relationship is evident.

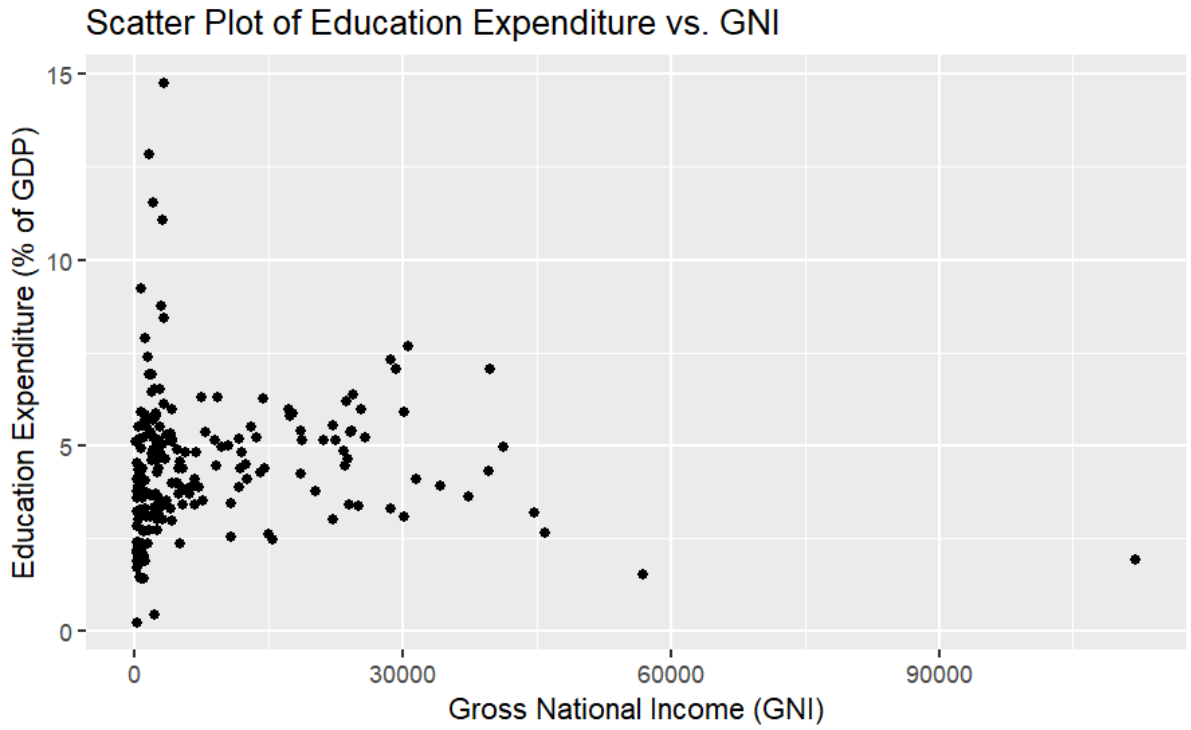

Figure 14: Scatter Plot of Education Expenditure vs. GNI

The scatter plot reveals that there is little to no discernible relationship between GNI and education expenditure. The very weak correlation coefficient of 0.0055 supports this observation, suggesting that changes in GNI are not associated with significant changes in education spending. This indicates that factors influencing education expenditure may be independent of economic performance as measured by GNI.

The clustering of data points around lower percentages of education expenditure implies that many countries invest a relatively small portion of their GDP in education, irrespective of their income levels. This could suggest a common threshold for education investment, where countries might prioritize other areas of expenditure, regardless of their economic capacity.

Additionally, the presence of a few outliers with high education expenditure might indicate countries that prioritize education significantly, potentially due to different socio-economic policies or cultural values. However, these outlier cases do not influence the overall weak relationship evidenced in the majority of the data.

Overall, the findings from the scatter plot highlight the complexity of education funding, suggesting that while GNI may provide some context, it is not a reliable predictor of education expenditure.

Next, we utilized data from the Worldwide Governance Indicators (WGI) to analyze social factors. This dataset provides insights into broad trends in perceptions of governance quality across countries and over time. According to the World Bank (2024), the WGI is derived from a variety of data sources produced by over 30 think tanks, international organizations, non-governmental organizations, and private firms worldwide. These sources (1) generate original primary data using well-defined methodologies, (2) capture respondents' perceptions of relevant governance dimensions, (3) encompass multiple countries, and (4) are regularly updated to ensure accuracy and relevance.

From the original dataset, we focused on the variable 'pctrank', which indicates the percentile rank of each country, ranging from 0 (lowest) to 100 (highest). We first removed all missing values to ensure data integrity. Consistent with our approach throughout this project, rather than concentrating solely on the most recent data from 2023, we calculated the average across the entire dataset. This method provides a more comprehensive analysis and mitigates the impact of short-term fluctuations. Subsequently, we merged this DataFrame with GNI data for further analysis, as shown as Figure 15.

code	Mean_WGI	Mean_GDP	Mean_GNI	IncomeGroup	Region
ABW	85.719425	37577.873	21171.3889	High income	Latin America & Caribbean
AFG	5.452119	2393.391	437.7273	Low income	South Asia
AGO	14.932595	7709.619	1621.2821	Lower middle income	Sub-Saharan Africa
ALB	42.713212	9967.503	2666.5000	Upper middle income	Europe & Central Asia
ARE	66.904411	82842.932	34296.9231	High income	Middle East & North Africa
ARG	44.496949	24377.324	5747.6667	Upper middle income	Latin America & Caribbean
ARM	41.865229	9362.556	2507.1875	Upper middle income	Europe & Central Asia
ATG	70.514200	25733.289	10874.2222	High income	Latin America & Caribbean
AUS	92.326815	48266.534	24187.4194	High income	East Asia & Pacific
AUT	92.275414	56694.099	24396.1290	High income	Europe & Central Asia

Figure 15: DataFrame of WGI and GNI

The density plot in Figure 16 illustrates the distribution of the Worldwide Governance Indicators (WGI) across countries. The shape of the distribution is approximately bell-shaped, indicating a relatively normal distribution centered around the mean. The red dashed line represents the mean WGI, which is positioned around the midpoint of the scale, suggesting that a significant number of countries have governance indicators clustered around this average value.

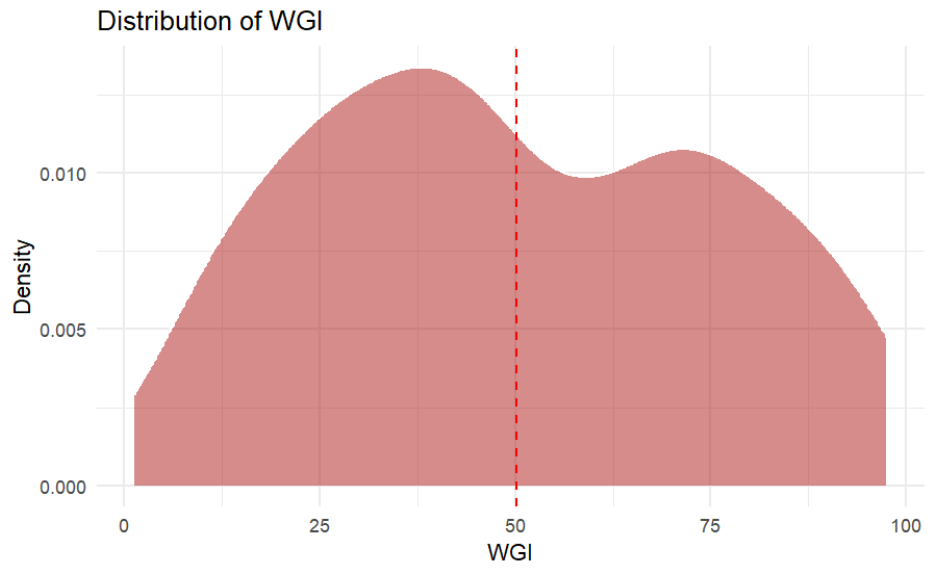

Figure 16: Density plot of WGI

The density plot reveals key insights into the governance quality perceptions across countries. The concentration of values around the mean suggests that many countries experience governance quality that is perceived as average, with fewer countries at both the high and low extremes.

The presence of a slight skewness or flattening in the distribution indicates that while there are countries with very high governance quality, the majority remain close to the mean. This observation highlights potential disparities in governance quality, where a minority of countries excel significantly compared to the average.

The scatter plot in Figure 17 reveals a strong positive relationship between Gross National Income (GNI) and the Worldwide Governance Indicators (WGI), as evidenced by the Spearman rank correlation coefficient of **0.8102**. This high correlation indicates that countries with higher GNI tend to report better governance quality.

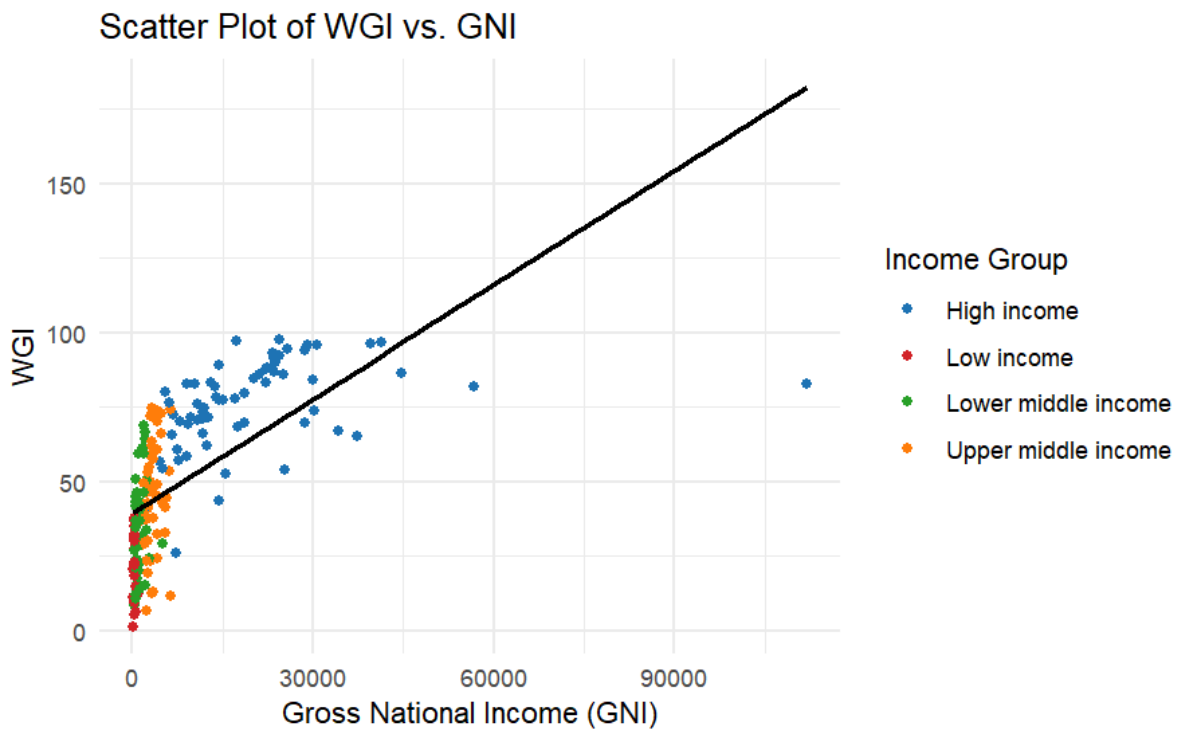

Figure 17: Scatter plot of WGI vs. GNI

The simple linear regression model in Figure 18 provides further insights into this relationship:

The Multiple R-squared value of 0.4108 indicates that approximately 41.1% of the variability in GNI can be explained by WGI. The Adjusted R-squared value of 0.4076 suggests a reasonably good fit for the model, accounting for the number of predictors used.

The Intercept is estimated at -7632.77, which represents the expected GNI when WGI is zero, although a WGI score of zero is not practically applicable.

The coefficient for Mean_WGI is 321.63 with a highly significant p-value (< 2e-16). This indicates that for every unit increase in WGI, GNI increases by approximately 321.63 units, demonstrating a strong positive relationship between governance quality and economic performance.

The residuals range from a minimum of -13036 to a maximum of 93030, indicating significant variability in the data. The residual standard error is 9947, suggesting some degree of prediction error in the model.

```
Call:
lm(formula = Mean_GNI ~ Mean_WGI, data = wgi)

Residuals:
   Min     1Q Median     3Q    Max
-13036  -4999  -1030   3187  93030

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -7632.77    1592.15  -4.794 3.34e-06 ***
Mean_WGI      321.63      28.24  11.388  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 9947 on 186 degrees of freedom
Multiple R-squared:  0.4108,    Adjusted R-squared:  0.4076
F-statistic: 129.7 on 1 and 186 DF,  p-value: < 2.2e-16
```

Figure 18: Simple regression model to predict WGI based on GNI

The findings from this simple regression model highlight a significant and positive relationship between WGI and GNI. The strong coefficient for Mean_WGI suggests that improvements in governance quality are associated with substantial increases in economic performance. This aligns with existing literature that posits effective governance enhances economic outcomes through mechanisms such as improved institutional integrity, reduced corruption, and better public services.

The R-squared value of 41.1% indicates that while WGI explains a significant portion of the variability in GNI, nearly 59% of the variability remains unexplained. This suggests that other factors, such as economic policies, historical contexts, social dynamics, and external economic conditions, also play important roles in determining GNI.

The residual analysis reveals a wide range of prediction errors, indicating that while the model captures a general trend, individual country variations can be substantial. Some countries may experience high GNI despite lower WGI scores, possibly due to unique economic conditions or external factors such as foreign investment or resource wealth.

In conclusion, this simple regression analysis underscores the importance of governance as a critical driver of economic performance. Policymakers should focus on enhancing governance quality as a strategy to foster economic growth. However, recognizing the limitations of the model is crucial, as a multi-faceted approach considering additional variables will likely yield a more comprehensive understanding of the factors influencing GNI.

In our research, we identified that Gross National Income (GNI) exhibits the strongest associations with Worldwide Governance Indicators (WGI) in social factors and region in geographical factors. Consequently, we decided to develop a multiple regression model to explore and illustrate the relationships among these variables more comprehensively.

```
Residuals:
   Min     1Q Median     3Q    Max
-33575  -4121   -913   2599  57911

Coefficients:
                                          Estimate Std. Error t value
(Intercept)                              -7810.46    2281.99  -3.423
Mean_WGI                                   273.99      29.05   9.431
factor(Region)Europe & Central Asia       4641.11    2008.07   2.311
factor(Region)Latin America & Caribbean   -442.82    2158.01  -0.205
factor(Region)Middle East & North Africa  6815.34    2563.95   2.658
factor(Region)North America              39242.16    5285.35   7.425
factor(Region)South Asia                  -319.48    3492.68  -0.091
factor(Region)Sub-Saharan Africa           574.71    2165.00   0.265
                                          Pr(>|t|)
(Intercept)                              0.000768 ***
Mean_WGI                                  < 2e-16 ***
factor(Region)Europe & Central Asia      0.021953 *
factor(Region)Latin America & Caribbean  0.837648
factor(Region)Middle East & North Africa 0.008565 **
factor(Region)North America              4.35e-12 ***
factor(Region)South Asia                 0.927220
factor(Region)Sub-Saharan Africa         0.790963
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 8611 on 180 degrees of freedom
Multiple R-squared:  0.5727,    Adjusted R-squared:  0.5561
F-statistic: 34.47 on 7 and 180 DF,  p-value: < 2.2e-16
```

Figure 19: Multiple Regression Model to predict GNI

The Multiple R-squared value of 0.5727 indicates that approximately 57.3% of the variability in GNI can be explained by WGI and the region. The Adjusted R-squared value of 0.5561 suggests a good model fit, considering the number of predictors included.

The Intercept is estimated at -7810.46, suggesting a baseline value for GNI when WGI is zero, although a WGI of zero is not practically relevant.

The coefficient for Mean_WGI is 273.99 with a highly significant p-value (< 2e-16). This indicates that for every unit increase in WGI, GNI increases by approximately 273.99 units, demonstrating a strong positive relationship between governance quality and economic performance.

Among the regions:

- North America has a significant positive coefficient of 39,242.16 ($p < 4.35e-12$), indicating that this region has the highest GNI relative to the reference group.
- Middle East & North Africa shows a significant positive coefficient of 6815.34 ($p = 0.008565$), suggesting a higher GNI compared to the reference region.
- Europe & Central Asia also exhibits a positive coefficient of 4641.11 ($p = 0.021953$), indicating a favorable GNI.
- In contrast, regions such as Latin America & Caribbean, South Asia, and Sub-Saharan Africa show non-significant coefficients, indicating their GNI is not significantly different from the reference region.

The findings from this regression model highlight the significant role that governance quality (as measured by WGI) and regional factors play in determining GNI. The strong positive coefficient for WGI suggests that improvements in governance quality are closely linked to economic performance. This relationship underscores the notion that effective governance can lead to better economic outcomes, likely through enhanced institutional frameworks, improved public services, and increased investor confidence.

The regional analysis reveals stark disparities in GNI, particularly with North America standing out as the region with the highest GNI, followed by the Middle East & North Africa and Europe & Central Asia. These regions benefit from a combination of higher WGI scores and favorable economic conditions.

Conversely, the lack of significant findings for regions such as Latin America & Caribbean, South Asia, and Sub-Saharan Africa suggests that these areas may face unique challenges that impede their economic performance, despite potential governance improvements. This could be due to a variety of factors, including historical economic structures, ongoing political instability, or social issues that affect economic growth.

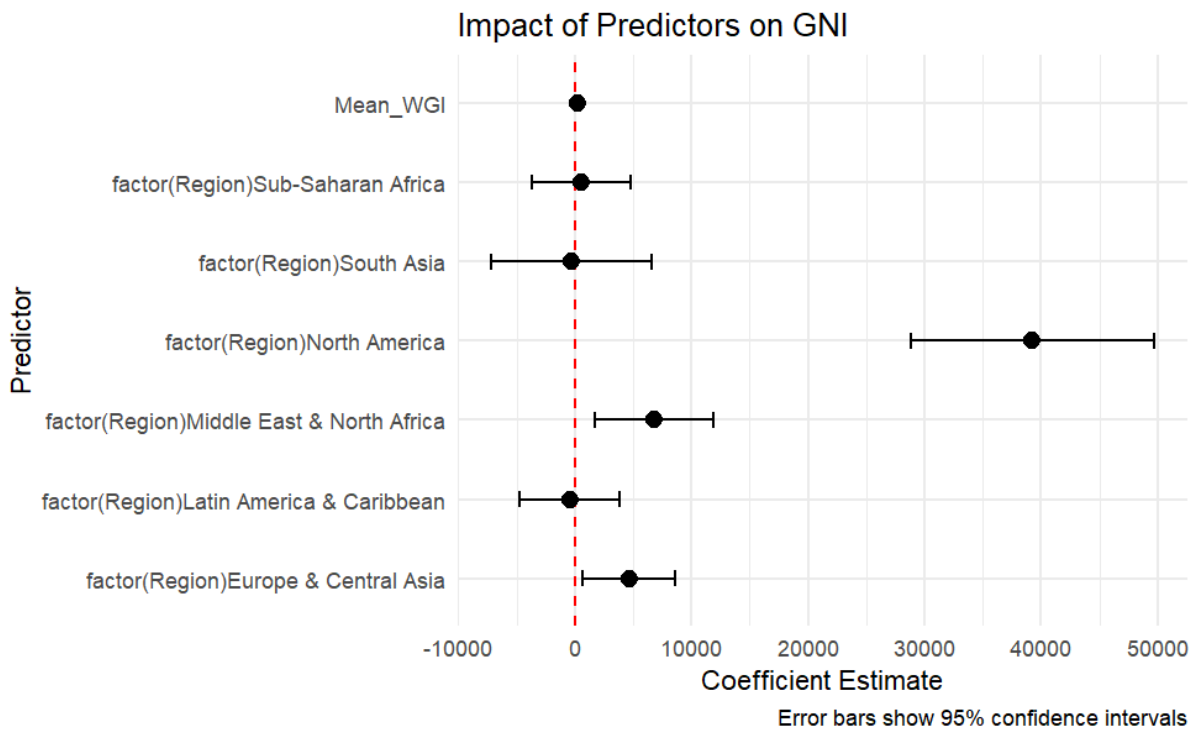

Figure 20: Coefficient Plot for GNI

The coefficient plot in Figure 20 effectively illustrates the impact of each predictor on GNI, with error bars representing 95% confidence intervals. The clear differentiation of coefficients indicates varying impacts of governance and region on GNI. Regions like North America have a substantial positive impact, while others such as South Asia and Sub-Saharan Africa do not show significant differences from the reference category.

In conclusion, this analysis emphasizes the importance of governance quality in driving economic performance, while also highlighting the significant variations in GNI across different regions. Policymakers should focus on enhancing governance as a pathway to economic growth, while also addressing the specific challenges faced by underperforming regions to foster broader economic development.

6.3 Ethical Considerations

It is concerned that our research focus is only the impacts towards the economy of different independent variables, and giving the illusion that other criterias are not worthed to consider. For example, in the heatmap of figure 11, it shows that the fatality number caused by disasters has no correlation to GNI level and economic losses, while it is not a significant predictor, it is a significant issue to deal with as it involves people's lives. Moreover, although education spending is not correlated to GNI level in our findings, it does not mean that public education spending is not necessary, because it involves values that cannot be measured by money such as social welfare and human rights.

7. Summary

Based on the above findings, certain recommendations are proposed to address the issue of geospatial poverty:

7.1 The defense of meteorological disasters

As the figure 10 shows, typical cyclones are the kind of disaster that cause the most economic loss among all, and it is way beyond the others. Tornadoes, droughts and severe weather have a significant impact in economic destruction either. Also, figure 9 shows that meteorological disasters are the second most frequently occurred type of disasters, just behind hydrological disasters. Therefore, effectively defending the damages from meteorological disasters can reduce the economic loss due to disasters to a large extent. It is suggested that shared weather information of satellites, and building weather stations are essential for preventing tremendous economic loss caused by meteorological disasters. A lot of developing areas have less advanced weather stations, and no satellites for weather monitoring. The international

community can help those regions by sharing the meteorological information from satellites. Furthermore, we can help developing areas to build more advance and quantity of weather stations, so they collect more accurate and larger spatial coverage of the weather data , enhance their ability to predict meteorological disasters, so the governments can send alerts timely to the society to prepare and protect their business and properties before the disasters. Apart from that, increasing the resilience of infrastructures is also important. According to the report of (The Straits Times, 2024), every year, disasters cause economic losses between US$732 billion to US$845 billion on average in infrastructure and buildings, which is equivalent to 14% of global GDP in 2022. The governments of related regions should emphasize on the robustness of infrastructures, and the international community can help them by providing labour, advanced architectural technologies and low-cost building materials, so they can increase the resilience of infrastructure and reduce the risk of being destroyed by the meteorological disasters.

7.2 The defense of hydrological disasters

Apart from meteorological disasters, hydrological disasters are another type of disaster that occurs the most, and causes the largest economic losses. As shown in figure 9, hydrological disasters is the most occured disaster group among all. In figure 10, it can be seen that hydrological disasters such as flood, tsunami, and ground movement cause significant economic losses. To minimize the damages to the economy from them, it is recommended to have better water resource reservation to address drought, which includes developing an efficient irrigation system and harvesting rainwater during droughts. According to the research of (Bwambale et al., 2021), efficient strategies for agricultural irrigation are crucial for water reservation. Harvesting rainfall can also increase water conservation and address

droughts. For tsunami and ground movement, It is recommended to regulate and restrict developments around high-risk areas, which includes floodplains, wetlands, and unstable slopes. Those areas are dangerous when floods, landslides, or soil erosion occur, limiting economic activities in those areas can significantly prevent lives and economic losses. Moreover, by limiting development in those areas can stabilize soil, prevent soil erosion, and increase the capacity of land to absorb water, which can prevent the occurrence risk of hydrological disasters, and mitigate the economic losses. Practicing contour farming and maintaining vegetative cover are also crucial for sustainability, because these agricultural practices will not cause land degradation and reduce runoff, so the risk of floods and landslides will be decreased.

7.3 Government expenditure on education

From our findings in Figure 13, there is no strong correlation observed between education expenditure and GNI level, the correlation coefficient is only around 0.0055, which means that based on our model, government spending on education contributes only a little to people's wealth. In other words, spending on education might not be the most effective way to increase people's wealth. However, it is necessary to be aware that the situation might depend on specific regions. For example, in figure 12, a small number of countries spend a tremendous percentage of GDP in education. It might be the social welfare policies for those countries, but it is also possible that those countries are knowledge-based economies, which are the economies that highly rely on industries that require a high level of professional knowledge. In that case, education expenditure might be very significant for GNI, therefore, the correlation between education expenditure and GNI level depends.

7.4 Focus on WGI

The result shows a strong correlation between WGI and GNI level, with a correlation coefficient of 0.8102. It is recommended that instead of focusing on the previous independent variables, governments of poor regions should focus on governance quality including the level of voice and accountability, political stability and absence of violence/terrorism, government effectiveness, regulatory quality, rule of law, and control of corruption. The research from (Tsegaw, 2020) echoes with our findings, it found that WGI is highly correlated with HDI (Human Development Index), it is interpreted that doing well on WGI means an appropriate place for economics activities in terms of safety, rule of law, government efficiency, and corruption level. These factors are crucial for attracting foreign capital to the economies. Therefore, in order to get rid of poverty, governments of poor regions should enhance their governance quality by forming independent organizations to improve the six dimensions of WGI, such as tackling corruption by establishing organizations like ICAC in Hong Kong.

8. Reference

Akkoyunlu, S., & Ramella, D. (2018). Corruption and Economic Development. Proceedings of the 10th Economics & Finance Conference, Rome. https://doi.org/10.20472/efc.2018.010.001

Akter, S., & Mallick, B. (2013). The poverty–vulnerability–resilience nexus: Evidence from Bangladesh. Ecological Economics, 96, 114–124. https://doi.org/10.1016/j.ecolecon.2013.10.008

Bangalore, M., Hallegatte, S., Vogt-Schilb, A., & Rozenberg, J. (2016). Unbreakable: building the resilience of the poor in the face of natural disasters. In Washington, DC: World Bank eBooks. https://doi.org/10.1596/978-1-4648-1003-9

Barro, R. J. (2013). Education and economic growth. Annals of economics and finance, 14(2), 301-328.

Botzen, W. J., Deschenes, O., & Sanders, M. (2019). The economic impacts of natural disasters: A review of Models and empirical studies. Review of Environmental Economics and Policy, 13(2), 167–188. https://doi.org/10.1093/reep/rez004

Bwambale, E., Abagale, F. K., & Anornu, G. K. (2022). Smart Irrigation Monitoring and control strategies for improving water use efficiency in precision agriculture: A Review. *Agricultural Water Management*, *260*, 107324. https://doi.org/10.1016/j.agwat.2021.107324

Castaño, M., Méndez, M., & Galindo, M. (2015). The effect of social, cultural, and economic factors on entrepreneurship. Journal of Business Research, 68(7), 1496–1500. https://doi.org/10.1016/j.jbusres.2015.01.040

Centre for Research on the Epidemiology of Disasters. (2025). *EM-DAT - The international disaster database*. https://www.emdat.be/

CollegeNP. (2023). Urban vs rural development: An economic perspective. CollegeNp.com. https://www.collegenp.com/article/urban-vs-rural-development-an-economic-perspective

Gallup, J. L., Sachs, J., & Mellinger, A. (1998). Geography and economic development. https://doi.org/10.3386/w6849

GATAWA, G. (2022). *The Effect of Social Factors to Economic Growth*. https://doi.org/10.31124/advance.19397081.v1

Hanushek, E., & Wößmann, L. (2010). Education and economic growth. In Elsevier eBooks (pp. 245–252). https://doi.org/10.1016/b978-0-08-044894-7.01227-6

Jin, X., Luan, W., Yang, J., Yue, W., Wan, S., Yang, D., Xiao, X., Xue, B., Dou, Y., Lyu, F., & Wang, S. (2023). From the coast to the interior: global economic evolution patterns and mechanisms. Humanities and Social Sciences Communications, 10(1). https://doi.org/10.1057/s41599-023-02234-4

Liu, Y., Liu, J., & Zhou, Y. (2017). Spatio-temporal patterns of rural poverty in China and targeted poverty alleviation strategies. Journal of Rural Studies, 52, 66–75. https://doi.org/10.1016/j.jrurstud.2017.04.002

OECD. (2024). Education Economic and Social Outcomes | OECD. https://www.oecd.org/en/topics/education-economic-and-social-outcomes.html

Our World in Data. (2025). *GDP per capita – World Bank – In constant international-$* [Dataset]. Data compiled from multiple sources by World Bank. https://ourworldindata.org/grapher/gdp-per-capita-worldbank

Roser, M., Rohenkohl, B., Arriagada, P., Hasell, J., Ritchie, H., & Ortiz-Ospina, E. (2023). *Data page: World Bank income groups*. In *Economic growth*. Our World in Data. https://ourworldindata.org/grapher/world-bank-income-groups

Rozenberg, J., & Hallegatte, S. (2016). Model and methods for estimating the number of people living in extreme poverty because of the direct impacts of natural disasters. In World Bank, Washington, DC eBooks. https://doi.org/10.1596/1813-9450-7887

Scherr, S. J. (2000). A downward spiral? Research evidence on the relationship between poverty and natural resource degradation. Food Policy, 25(4), 479–498. https://doi.org/10.1016/s0306-9192(00)00022-1

Skidmore, M., & Toya, H. (2002). DO NATURAL DISASTERS PROMOTE LONG‐RUN GROWTH? Economic Inquiry, 40(4), 664–687. https://doi.org/10.1093/ei/40.4.664

Srinivasu, B., & Rao, P. S. (2013). Infrastructure development and economic growth: Prospects and perspective. Journal of business management and Social sciences research, 2(1), 81-91.

The Straits Times. (2024, November 13). *Climate change damage to infrastructure poses "huge" financial risks: Report*.

https://www.straitstimes.com/asia/south-asia/climate-change-damage-to-infrastructure-poses-huge-financial-risks

This is what the climate crisis is costing economies around the world. (2024, September 10). World Economic Forum. https://www.weforum.org/stories/2023/11/climate-crisis-cost-global economies/

UIS Data Browser. (2025). *Government expenditure on education as a percentage of GDP*. UNESCO Institute for Statistics. https://databrowser.uis.unesco.org/view#indicatorPaths=UIS-SDG4Monitoring%3A0%3AXGDP.FSGOV&geoMode=countries&geoUnits=&browsePath=EDUCATION%2FUIS-SDG4Monitoring&timeMode=range&view=table&chartMode=multiple&tableIndicatorId=XGDP.FSGOV&chartIndicatorId=XGDP.FSGOV&chartHighlightSeries=&chartHighlightEnabled=true

United Nations. (2024). Transforming our world: The 2030 agenda for sustainable development | department of economic and social affairs. https://sdgs.un.org/2030agenda

World Bank. (2025). *GNI per capita, Atlas method (current US$)* [Dataset]. https://data.worldbank.org/indicator/NY.GNP.PCAP.CD

World Bank. (2024). *Worldwide governance indicators*. https://www.worldbank.org/en/publication/worldwide-governance-indicators

Zhou, L., & Xiong, L. (2017). Natural topographic controls on the spatial distribution of poverty-stricken counties in China. Applied Geography, 90, 282–292. https://doi.org/10.1016/j.apgeog.2017.10.006