

CALIFORNIA STATE UNIVERSITY, NORTHRIDGE

Data-Driven Insights on Economic Growth, Public Debt, and Social Development

A graduate project submitted in partial fulfilment of the requirements

For the degree of Master of Science in Business Analytics

By

Sajid Khan Hyder

Masthani Sai Sharanya Tiruveedhi

Rahul Dipak Talreja

Ruchira Kanawade

May 2025

The graduate project is approved:

---

Dr. Akash Gupta

---

Date

---

Dr. Qiuhua Sheng

---

Date

---

Dr. Pouyan Eslami, Chair

---

Date

California State University, Northridge

## Table of Contents

Signature Page	ii
Abstract	iv
Chapter 1: Introduction	1
Chapter 2: Research Objectives and Question	2
Chapter 3: Methodology	3
Chapter 4: Results	6
Chapter 5: Discussion	17
Chapter 6: Conclusion	18

## Abstract

### Data-Driven Insights on Economic Growth, Public Debt, and Social Development

By

Sajid Khan Hyder

Masthani Sai Sharanya Tiruveedhi

Rahul Dipak Talreja

Ruchira Kanawade

Master of Science in Business Analytics

Foreign direct investment (FDI) is a major contributor for economic development, particularly in emerging and developing economies. However, the contributing factors that influence the magnitude of FDI received by a nation remain complex. Indicators that include determinants and socioeconomic indicators such as education levels, employment rates, public debt levels, social

spending, and trade balances exert significant influence on investor confidence and capital inflows. The primary objective of this research is to analyze the impact of a nation's socioeconomic indicators on the volume of FDI it receives and subsequently assess the consequent effects of FDI on its gross domestic product (GDP) growth. This study is a well-rounded approach to analyze the most relevant socioeconomic factors that affects the FDI inflow across different countries combining several methods such as, Multiple Linear Regression (MLR), general regression models, time series analysis, and classification techniques. This will help in understanding how these investments affect the GDP growth over time. Moreover, classification models such as decision trees and logistic regression will be applied to categorize countries based on their FDI trends and their GDP growth. This analysis will also help compare patterns across countries that are grouped based on income levels, low, middle, and high to emphasize regional differences and trends. For consistency and reliability, all the data will be taken from World Bank database, focusing on the years 2021 to 2023. The main research objective of this study is to analyze the most influential indicators, in this case, socioeconomic factors that attract foreign investment and quantify their effects on GDP growth. It is anticipated to contribute to policymaking by highlighting specific areas, such as education, infrastructure development, and economic stability, where enhancements could lead to increased FDI inflows. Furthermore, the research may provide insights into how diverse country profiles respond to investment inflows, enabling international organizations and investors to make informed decisions.

Keywords: Foreign Direct Investment (FDI), Gross Domestic Product (GDP) Growth, Socioeconomic Indicators, Economic Development, Regression Analysis

## **Chapter 1: Introduction**

### **1.1 Background of the Study**

In the globalized economic landscape, nations actively seek to attract Foreign Direct Investment (FDI) as a catalyst for economic growth, infrastructure development, and technological advancement. Simultaneously, a country's macroeconomic stability—shaped by indicators such as public debt, trade balance, inflation, and unemployment—plays a crucial role in determining its investment attractiveness. Developing and emerging economies face a delicate balancing act between borrowing for development and maintaining fiscal sustainability. Understanding how these socioeconomic indicators influence FDI inflows and, in turn, affect GDP growth is vital for shaping data-driven economic policy.

### **1.2 Statement of the Problem**

Despite the growing availability of economic data, the causal and correlative relationships between a nation's debt structure, trade dynamics, social investment, and GDP growth remain complex and context dependent. In particular, the impact of external debt on economic growth variability, and the predictive power of economic indicators on FDI inflows, is not well understood. There is a need for empirical analysis to clarify whether high debt discourages investment or if trade performance and fiscal behavior can offset such concerns. This research seeks to bridge that gap by analyzing multi-year data across developing and emerging economies to uncover the structural dynamics at play.

### **1.3 Importance and Significance of the Research**

This study is significant for policymakers, economists, and development agencies aiming to enhance economic performance through better management of debt, investment incentives, and

fiscal policy. By identifying the key predictors of FDI and GDP growth, the research offers actionable insights for countries seeking to improve their global investment profile and sustain long-term development. The study's findings also contribute to academic discourse by combining traditional statistical methods with machine learning models to understand complex economic systems.

## 1.4 Structure of the Report

Chapter 2: Research Objectives and Research Questions

Chapter 3: Methodology

Chapter 4: Results

Chapter 5: Discussion

Chapter 6: Conclusion

## **Chapter 2: Research Objectives and Research Question**

### 2.1 Research Objectives

This study aims to explore the interrelationship between socioeconomic indicators, foreign direct investment (FDI), and GDP growth. The specific objectives of this research are:

To evaluate how key socioeconomic indicators influence the inflow of foreign direct investment (FDI) into developing and emerging economies.

To assess the impact of external and public debt levels on GDP growth and its variability.

To identify the most significant predictors of FDI inflows using statistical and machine learning techniques.

To investigate whether trade balance, inflation, and public debt have measurable effects on GDP per capita growth.

To categorize countries based on economic profiles through clustering and classification models for deeper insight into policy-driven economic patterns.

## 2.2 Research Question

How do a nation's socioeconomic indicators influence the volume of foreign direct investment (FDI) it receives, and what is the subsequent impact on its gross domestic product (GDP) growth?

# Chapter 3: Methodology

## 3.1 Database and Data

The study utilizes data obtained from World Bank Data (<https://data.worldbank.org>), a comprehensive and reputable global database. The dataset includes annual country-level observations spanning several decades and contains a wide array of macroeconomic and development indicators. Key indicators used in this study include GDP growth (%), GDP per capita growth, public debt (% of GDP), foreign direct investment (FDI) inflows, inflation rate, trade balance, tax revenue, unemployment rate, social spending, exchange rates (PPP), and government expenditure on debt servicing.

Each record in the dataset consists of the country name, indicator (series) name, year of observation, and corresponding numerical value. Data was filtered to focus on developing and emerging economies, as these represent regions with dynamic economic changes and high



relevance for FDI-related insights. To ensure consistency and integrity, only countries and years with complete records for key indicators were retained.

Log transformation was applied to the FDI inflow data to normalize the distribution and reduce the influence of extreme values. Categorical variables such as country classifications (e.g., "Emerging" vs. "Developing") were encoded numerically for compatibility with machine learning algorithms. The final sample included a curated set of countries and years selected based on indicator completeness, economic relevance, and data quality, enabling robust statistical and predictive modeling.

### 3.2 Analytical Methods

To analyze the relationship between socioeconomic indicators, foreign direct investment (FDI), and GDP growth, we employed a combination of statistical tests, machine learning models, and clustering algorithms. The analyses were conducted using Python-based tools, including libraries such as pandas, scikit-learn, matplotlib, and seaborn, within a Jupyter Notebook environment for interactive data exploration and model development.

We began with descriptive and exploratory data analysis to understand the distribution and trends of key variables such as GDP growth, unemployment, inflation, trade balance, and various forms of debt. Visual tools like line graphs, bar charts, and correlation heatmaps were used to highlight differences across income groups and regions.

To determine whether external debt influences GDP growth variability, we conducted an independent two-sample T-test comparing countries with high versus low external debt levels. The resulting p-value (0.0023) indicated a statistically significant difference, confirming that higher external debt is associated with greater volatility in GDP growth.

Next, we employed multiple linear regression to evaluate how inflation, trade balance, and public debt influence GDP per capita growth. The regression analysis revealed that inflation and public debt had significant negative effects, while trade balance did not show statistical significance. The R-squared value for this model was 0.026, suggesting limited explanatory power but meaningful directional insights.

To further examine growth classification and prediction, we applied logistic regression to classify countries as high-growth or low-growth economies. For predicting GDP growth more accurately, we compared linear regression, random forest regression, and neural networks (multilayer perceptron). Among these, the random forest model showed the best performance with an R-squared value of approximately 0.15, outperforming the linear and neural network models in capturing non-linear relationships.

For predicting FDI inflows, we applied both random forest and linear regression models to the log-transformed FDI data. While the linear regression model appeared to fit perfectly ( $R^2 = 1.0$ ), it exhibited clear signs of overfitting. The random forest model, with an  $R^2$  of 0.62, was more robust and identified the external balance on goods and services as the top predictor of FDI inflows.

A decision tree classifier was used to understand the key factors distinguishing emerging from developing countries. This interpretable model highlighted tax revenue and unemployment rate as the most influential variables for classification.

Lastly, K-means clustering was employed to segment countries into three distinct groups based on GDP growth, public debt, and trade balance. The optimal number of clusters ( $k = 3$ ) was

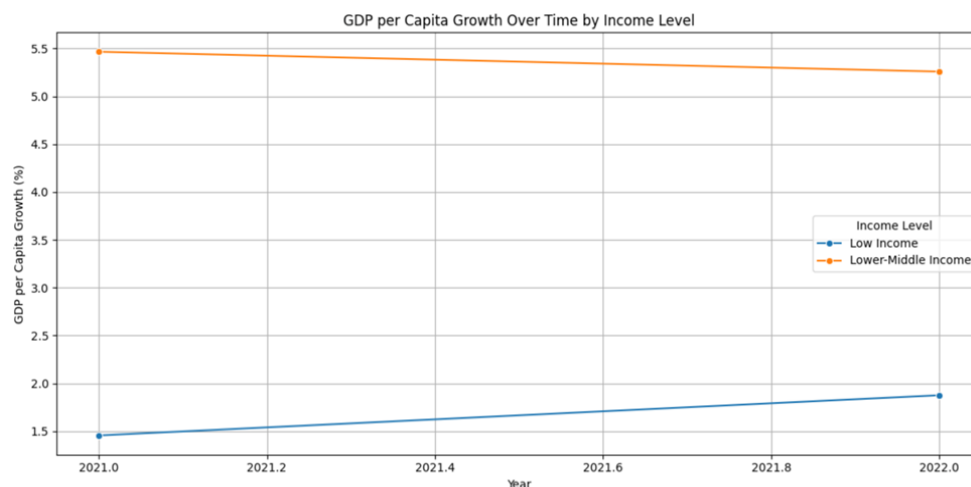
determined using the Elbow Method. This clustering provided insights into common economic patterns and informed our understanding of investment trends and policy implications.

Additional tests, such as F-tests and correlation analysis, were conducted to assess the relationship between social spending and GDP stability, as well as between trade balance and GDP growth. Although the correlation between trade balance and GDP growth was weak ( $r = 0.197$ ), it still suggested a consistent global trend.

## Chapter 4: Results

### 4.1 Economic Growth Trends by Income Level

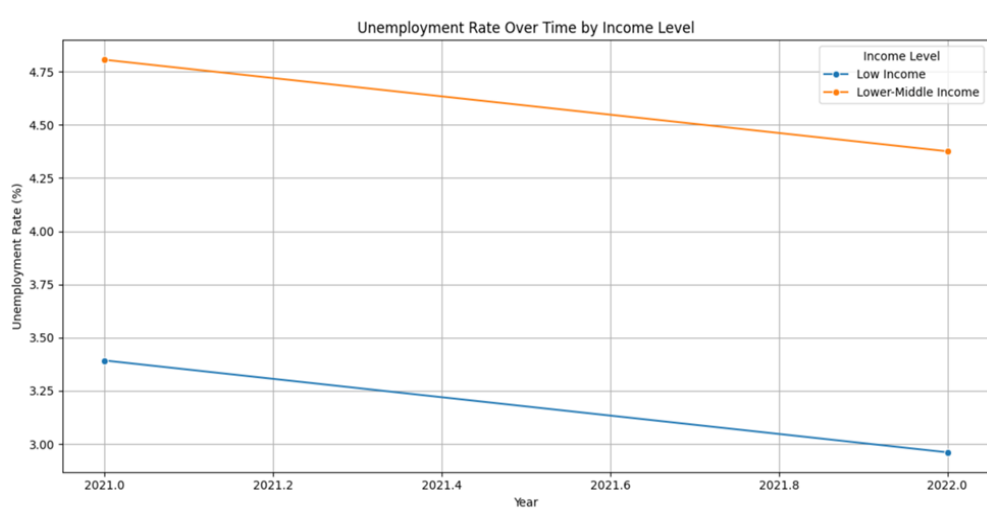
This study identifies clear distinctions in economic performance between income groups. For low-income countries, GDP per capita growth exhibited a modest upward trend, increasing from approximately 1.5% in 2021 to around 2.0% in 2022. This suggests a gradual economic recovery following pandemic-induced disruptions. In contrast, lower-middle-income countries demonstrated higher growth rates overall but with a slight decline from about 5.5% to 5.0%, pointing to a strong yet stabilizing economic condition. These patterns highlight the disparity in resilience and recovery speed across income tiers.



*Fig 1: GDP per capita over time by income level*

## 4.2 Unemployment Rate Analysis

Labor market trends mirrored the GDP dynamics to some extent. Low-income countries saw a decline in unemployment rates from 3.3% to 3.0%, reflecting a slow but positive improvement in employment conditions. In lower-middle-income countries, the decline was more significant, from approximately 4.75% to 4.25%, suggesting better labor market performance. These figures indicate that as GDP improves, employment prospects tend to follow, though the relationship varies across income levels.

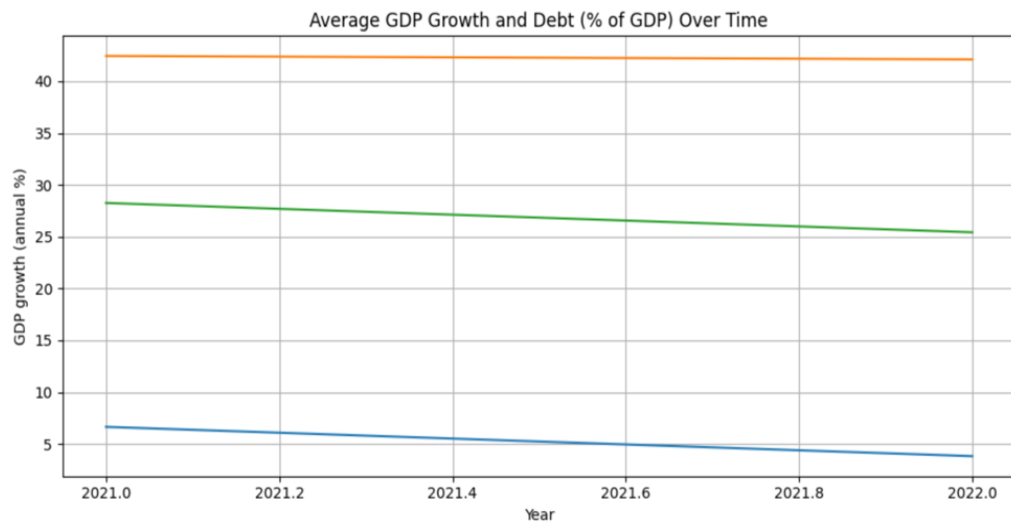


*Fig 2: Unemployment rate over time by income level*

## 4.3 Impact of Domestic and External Debt on GDP Growth

An assessment of debt profiles revealed that domestic debt remained consistently high, averaging around 43% of GDP. This suggests a continued reliance on domestic borrowing, potentially due to limited access to international credit or intentional fiscal strategy. Conversely, external debt declined slightly, from about 29% to 27%, pointing to a reduced dependence on foreign creditors. However, GDP growth dropped from 6% to below 5%, which could indicate that high

domestic debt burdens are exerting pressure on national growth potential. These findings underscore the delicate balance between leveraging debt for growth and maintaining fiscal sustainability.



*Fig 3: Average GDP growth and debt over time*

#### 4.4 GDP Growth Classification: Logistic Regression Analysis

Using logistic regression, countries were classified into high-growth and low-growth economies. High-growth countries, visualized in blue on the world map, were primarily emerging markets across Asia, parts of Europe, and select regions in Africa and Latin America. These countries are typically attractive for foreign direct investment (FDI) due to their growth potential but also pose higher investment risks. Low-growth countries, shown in yellow, included advanced economies in North America and Australia. These markets offer more stability and predictability, albeit with lower returns. The classification framework provides insights into investment strategies and country-specific economic trajectories.

World map showing the distribution of high and low levels of a variable. The legend indicates: High (blue) and Low (yellow).

#### 4.5 GDP Growth Prediction: Modeling Approaches

9

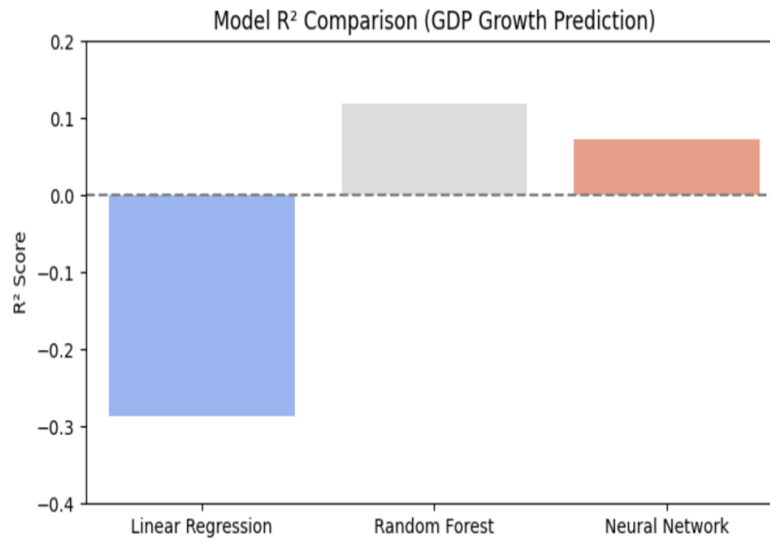


Fig 5: Model R² comparison for GDP growth prediction

#### 4.6 Key Predictors of FDI

Random forest and linear regression models were applied to identify significant predictors of FDI inflows. The linear regression model, though achieving an R-squared of 1.0, showed signs of overfitting. In contrast, the random forest model provided a more realistic R-squared of 0.62. Among the variables, the external balance on goods and services (% of GDP) emerged as the most influential predictor, followed by trade (% of GDP) and unemployment rates. These findings emphasize the importance of macroeconomic stability and trade performance in attracting foreign investment.

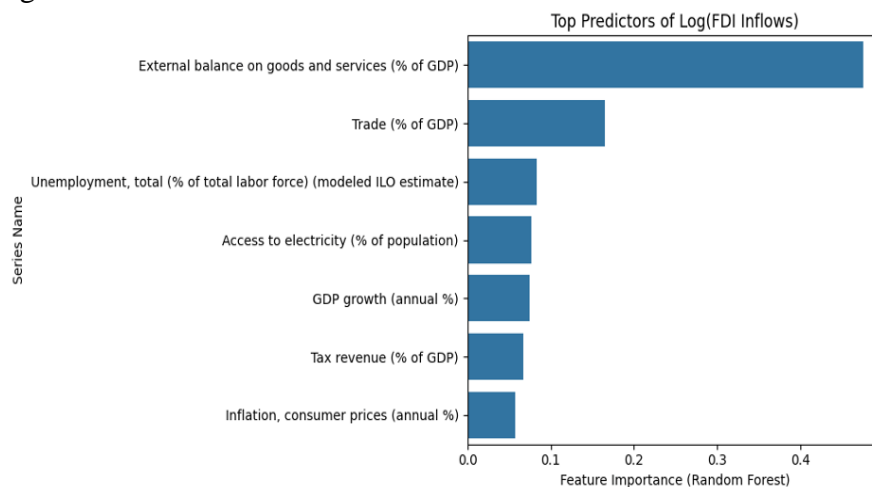
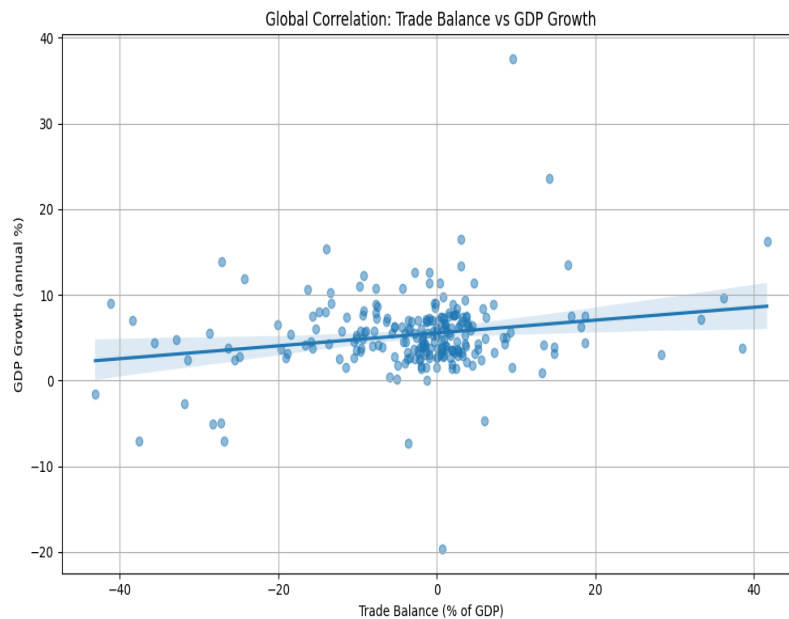


Fig 6: Top predictors of FDI

#### 4.7 Trade Balance vs. GDP Growth

A correlation analysis between trade balance and GDP growth revealed a weak but positive relationship, with a correlation coefficient of 0.197. While not statistically strong, the trend suggests that countries with healthier trade balances generally achieve better growth. Most developing countries clustered around a trade balance near zero and exhibited moderate GDP growth, reinforcing the notion that balanced trade supports economic stability.



*Fig 7: Trade balance vs GDP growth*

#### 4.8 Emerging vs. Developing Country Classification

A decision tree model was used to classify countries as emerging or developing based on economic indicators. The results indicated that countries with low tax revenue and low GDP growth are more likely to be classified as developing. In contrast, those with higher tax revenue and lower unemployment rates tend to fall under the emerging category. This classification helps



policymakers understand the critical economic factors that differentiate development stages and provides guidance for targeted policy interventions.

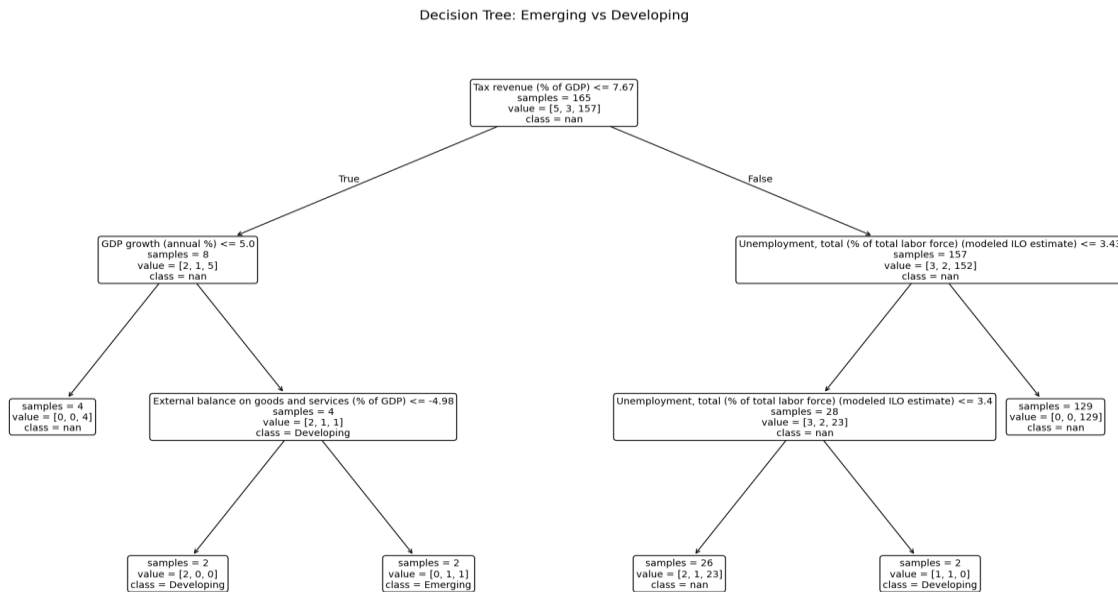


Fig 8: Classification of emerging and developing countries

#### 4.9 Significance of External Debt on GDP Growth (T-Test)

A two-sample T-test was conducted to evaluate whether there is a significant difference in GDP growth between countries with high and low external debt. The resulting p-value of 0.0023 indicated statistical significance at the 5% level. The analysis revealed that countries with high external debt experienced greater variability in GDP growth, while those with lower external debt enjoyed more consistent and stable growth. This finding supports the assertion that managing external debt is essential for economic stability.

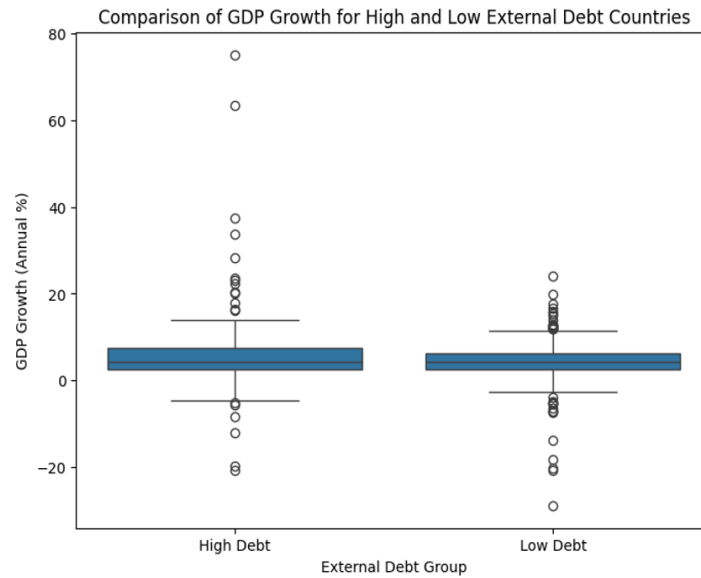


Fig 9: Comparison of GDP growth for high and low external debt countries

#### 4.10 Impact of Economic Factors on GDP per Capita Growth

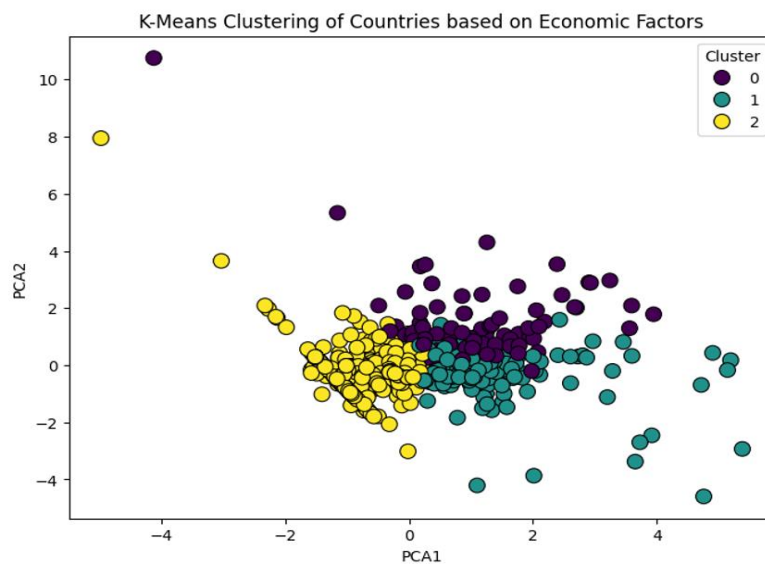
A multiple linear regression model examined the impact of inflation, trade balance, and public debt on GDP per capita growth. The model had a low R-squared value of 0.026, explaining only 2.6% of the variance. Both inflation (coefficient = -0.034) and public debt (coefficient = -0.118) had negative and statistically significant effects on GDP growth, while trade balance was not statistically significant. These results suggest that while inflation and public debt are influential, the relationship between trade and growth requires further exploration.

OLS Regression Results						
=====						
Dep. Variable:	GDP_per_Capita_Growth		R-squared:	0.026		
Model:	OLS		Adj. R-squared:	0.022		
Method:	Least Squares		F-statistic:	6.997		
Date:	Sat, 19 Apr 2025		Prob (F-statistic):	0.000120		
Time:	07:17:04		Log-Likelihood:	-2569.8		
No. Observations:	798		AIC:	5148.		
Df Residuals:	794		BIC:	5166.		
Df Model:	3					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
const	5.9832	0.570	10.502	0.000	4.865	7.101
Inflation	-0.0342	0.011	-3.007	0.003	-0.056	-0.012
Trade_Balance	1.642e-05	0.004	0.004	0.996	-0.007	0.007
Public_Debt	-0.1180	0.033	-3.605	0.000	-0.182	-0.054
=====						
Omnibus:	660.219	Durbin-Watson:	2.123			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	41955.376			
Skew:	3.252	Prob(JB):	0.00			
Kurtosis:	37.922	Cond. No.	259.			

Fig 10: OLS regression results

#### 4.11 Clustering of Countries Based on Economic Indicators

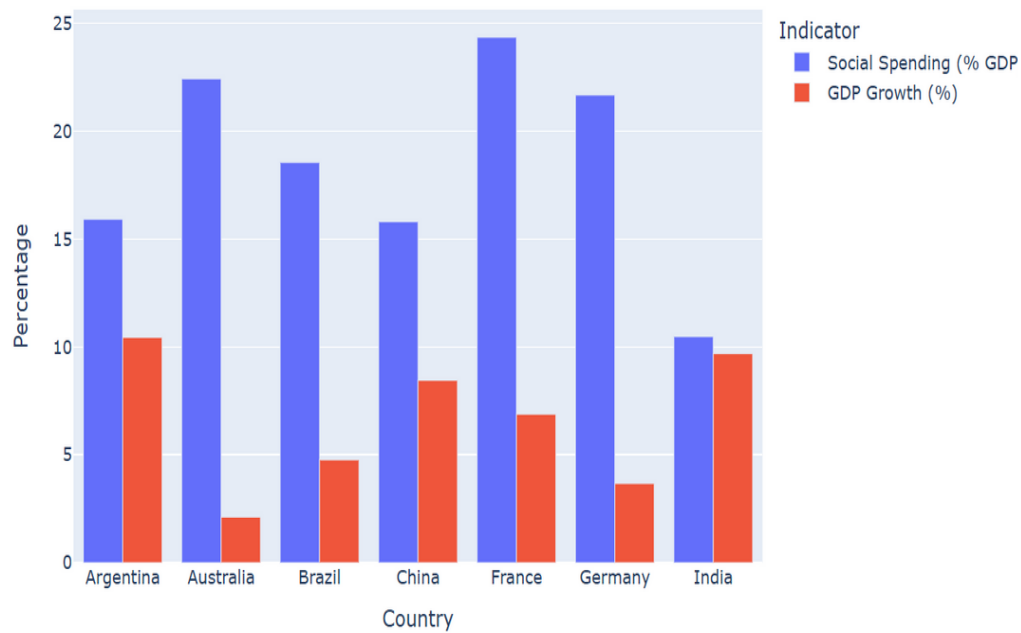
Using K-means clustering, countries were grouped into three distinct clusters based on GDP growth, public debt, and trade balance. This approach revealed that countries with similar economic profiles tend to exhibit similar economic behaviors. Clustering provided valuable insights for investment strategy and policy formulation, suggesting that economic similarity can predict similar patterns in investment behavior and growth performance.



*Fig 11: K-Means clustering of countries based on economic factors*

#### 4.12 Social Spending and Economic Stability

An analysis was conducted to determine whether higher social spending leads to better economic stability. The results were mixed. While some countries with high social spending displayed strong GDP growth, others did not, indicating inconsistency. An F-test revealed that high social spending did not necessarily reduce the variance in GDP growth. Argentina, for example, stood out with 10.44% GDP growth and 15.92% social spending, potentially driven by post-COVID fiscal policies. Overall, the analysis concluded that social spending alone is not a reliable predictor of economic stability.



*Fig 12: Social spending and GDP growth across countries*

#### 4.13 Debt Servicing, Exchange Rates, and FDI

This section explored how debt servicing costs and exchange rates affect FDI. Countries with higher debt servicing burdens, as a percentage of GDP, typically experienced lower FDI inflows. Additionally, exchange rate variability was linked to investor confidence. Clustering analysis using the Elbow Method identified three economic groupings based on GDP growth, public debt, and trade balance, revealing that similar economic structures often attract similar investment patterns. This insight can support strategic policy decisions to enhance investment appeal.

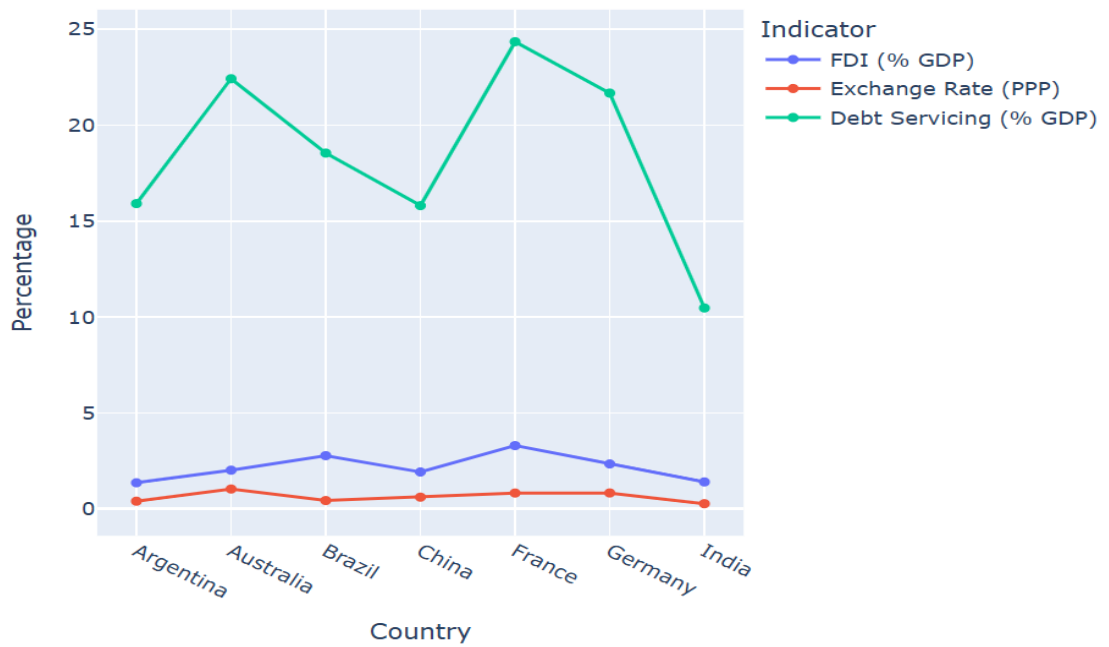
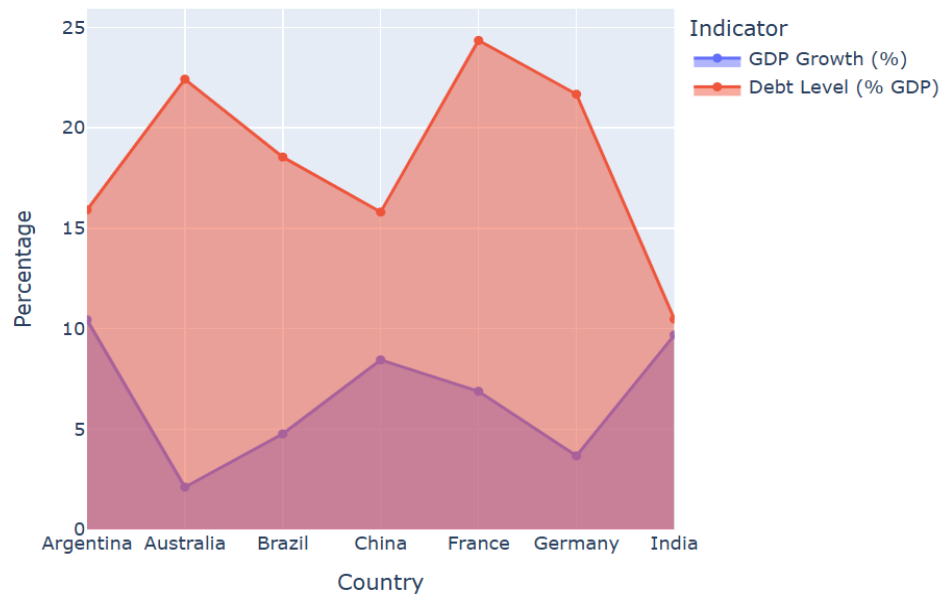


Fig 13: Percentages of FDI, exchange rate and debt servicing across countries

#### 4.14 Clustering Based on Debt and Growth Indicators

Finally, an area chart and correlation analysis were used to explore the relationship between debt levels and GDP growth. A potential pattern emerged, suggesting that moderate debt levels may be associated with higher growth. For instance, India exhibited high growth (9.69%) with moderate debt (17.95%), while France had lower growth (6.88%) and higher debt (24.35%). These contrasts highlight diverse fiscal strategies and suggest that debt level metrics may serve as useful predictors in future machine learning models aimed at forecasting growth.

### Predicting GDP Growth with Debt Levels



*Fig 14: Predicting GDP growth with debt levels*

## Chapter 5: Discussion

The results suggest a clear pattern: countries with low external debt tend to experience more stable GDP growth. This finding aligns with the theory that high debt can cause economic volatility. Our regression analysis also confirmed that inflation and public debt negatively affect GDP growth, which is consistent with existing literature.

However, there were some unexpected findings, such as the lack of statistical significance of the trade balance on GDP per capita growth. This warrants further research into how trade dynamics might impact economic outcomes, especially in countries with different economic structures.

The study is limited by its focus on developing nations and the absence of data for high-income economies. Future research could expand to include a broader range of countries and a longer timeframe to deepen the understanding of these relationships.

## **Chapter 6: Conclusion**

The key findings of this study highlight the significant impact of external debt on GDP growth, with countries carrying lower levels of external debt exhibiting more stable growth patterns. Additionally, inflation and public debt were found to negatively affect GDP per capita growth, indicating that managing these factors could contribute to greater economic stability. While social spending and exchange rates were initially expected to play a larger role, they proved important but not in the anticipated manner.

In terms of research significance, this study provides valuable insights into the relationship between debt and economic growth, particularly in the context of developing nations. The results can help inform policy decisions, offering guidance for creating balanced strategies that manage debt while fostering sustainable growth.

## References

[1] World Bank, “World Development Indicators,” The World Bank Group, 2023. [Online].

Available: <https://data.worldbank.org>