REGIONAL ECONOMIC PERFORMANCE ANALYSIS BASED ON MACROECONOMIC INDICATORS ACROSS 200 COUNTRIES (2010–2024)

RICHARD MAWUKO DAGBA

Dissertation submitted to International Business School
for the partial fulfilment of the requirement for the degree
of
MASTER OF SCIENCE IN IT FOR BUSINESS DATA
ANALYTICS

December 2025

DECLARATION

This dissertation is a product of my own work and is not the result of anything done in collaboration. I consent to the university's free use including online reproduction, including electronically, and including adaptation for teaching and education activities of any whole or part item of this dissertation

Dagba Richard Mawuko

Word length: 10,557 words

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to all those who supported and guided me throughout the completion of this dissertation.

First and foremost, I would like to thank my supervisors for their invaluable guidance, expertise, and continuous support throughout this research. I am deeply grateful to Dénes Jurányi, my supervisor with a focus on data analytics, for his insightful feedback, methodological guidance, and encouragement in applying advanced forecasting techniques and analytical rigor to this study. I am equally thankful to Péter Bárczy, my supervisor with a focus on the related business field, for his practical insights, constructive advice, and support in aligning the analytical findings with real-world business and economic implications.

I would also like to extend my appreciation to Gyarmathy Zsófia, the Module Leader and Seminar Leader for Advanced Forecasting Techniques for Data Scientists, whose lectures and seminars provided a strong theoretical and practical foundation for the forecasting methods applied in this dissertation.

Furthermore, I would like to thank the faculty and staff of the International Business School for providing a supportive academic environment and access to learning resources that contributed significantly to my studies.

Finally, I am grateful to my family and friends for their patience, encouragement, and understanding throughout the duration of my studies. Their constant support and motivation played an important role in enabling me to complete this dissertation successfully.

REGIONAL ECONOMIC PERFORMANCE ANALYSIS BASED ON MACROECONOMIC INDICATORS ACROSS 200 COUNTRIES (2010–2024)

DAGBA RICHARD MAWUKO

EXECUTIVE SUMMARY

This project looks at how economies perform across 200 countries from 2010 to 2025. It focuses on three main economic indicators which are GDP in current US dollars, GDP per capita in current US dollars and annual GDP growth. The study responds to a common problem in global economic research. Although economic data is widely available few studies combine regional comparisons with reliable forecasts using the same method. Many studies only describe past trends without looking ahead while others focus on forecasting without placing results in a regional context. This makes it harder for policymakers, investors and organizations to assess past performance and plan for the future.

The data used in this study was obtained from Kaggle and originally collected from the World Bank Open Data API. It covers more than 200 countries between 2010 and 2025. The data was carefully prepared by cleaning errors, checking consistency and handling missing values. Missing values for GDP, GDP per capita and GDP growth were filled using calculated values and additional sources so the dataset could be used with confidence for analysis.

The exploratory data analysis examined the three key indicators to understand patterns and trends. GDP in current US dollars is highly uneven across countries with a few large economies such as the United States China and Japan producing most of the global output. GDP per capita shows how wealthy people are on average and highlights high income microstates like Monaco Liechtenstein and Luxembourg. GDP growth rates differ widely across countries. Fast growing economies such as Guyana Djibouti and Tajikistan show strong growth but also high instability. These differences reflect factors such as level of industrial development technology, population size and global trade participation. Time series plots also captured the effect of major global events including the slowdown around 2014 and 2015 and the COVID pandemic in 2020 which caused short term declines in economic activity across many regions.

Forecasting was carried out for five major economies which are the United States China Japan Germany and the United Kingdom. Three forecasting models were used which are Theta ARIMA and Bayesian Structural Time Series. The models were trained using data from 2010 to 2019 and tested using data from 2020 to 2024. ARIMA and BSTS were carefully tuned to improve accuracy. ARIMA produced the most accurate point forecasts across most countries. BSTS added value by showing uncertainty ranges which helps with planning under risk. Theta provided a simple and stable baseline. The forecasts suggest that all

five economies are likely to continue growing up to 2030 although the speed and certainty of growth differ between countries due to their economic structure and past performance.

Overall, the results show the benefit of combining regional economic comparison with strong forecasting methods. The insights from this study can support government policy decisions, guide investment planning and inform development strategies.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

Chapter 1 INTRODUCTION

1.1 Background and Context of the Project

In a highly interconnected global economy, economies do not operate in isolation. Events in one country often affect others very quickly through trade, money flows, and financial markets. A policy change, an economic shock, or a financial problem in one region can spread across borders and influence many economies at once (Bationo et al, 2025). Because of this strong connection, understanding how different regions perform economically has become very important for governments, investors, and international organizations. Looking closely at key macroeconomic indicators helps decision makers track economic stability, judge growth prospects, and spot potential risks early (Ullah et al, 2024). Clear and reliable economic analysis supports better planning and faster responses when economic conditions begin to change.

Over the last fifteen years, the global economy has gone through major changes and disruptions (Papetti, 2021). After the 2008 to 2009 global financial crisis, countries recovered at very different rates (Kohler & Stockhammer, 2021). Wealthy economies kept interest rates low for a long time, while many developing countries struggled with debt and budget problems. Then came the COVID-19 pandemic, which caused sharp drops in economic output, higher unemployment, and record levels of government spending (Tetteh & Ntsiful, 2023). More recently, rising inflation and tighter monetary policies have created new challenges for keeping economies stable (Jorda & Nechio, 2023). Because of all these events, the years from 2010 to 2024 provide a good period to study long-term economic trends, how regions respond to shocks, and how their growth paths differ.

Macroeconomic indicators offer a clear way to understand how economies are performing. Measures like gross domestic product, Gross Domestic Product (GDP) growth, inflation, unemployment, public debt, and interest rates are commonly used to evaluate economic health (Ullah et al, 2024). Each indicator shows one part of the economy, but looking at them together gives a better picture of growth, government finances, and job market conditions (Tetteh & Ntsiful, 2023). Studying these indicators across different countries and over time helps reveal regional patterns, structural differences, and shared weaknesses that are not obvious when looking at just one country.

Even with large amounts of economic data available from sources like the World Bank, comparing countries in a meaningful way is still difficult because economies differ in size and structure and the quality of data varies (Maromarco, 2022). Looking at data alone is not enough for decision makers who need to plan for the future, so there is a growing need for methods that not only review past performance but also offer predictions and insights ahead (Buturac, 2021). This project fits into that need and aims to combine regional economic analysis with forecasting to give a clearer understanding and support better planning.

1.2 Business and Technical Significance

Regional economic analysis is very useful for both business and policy decisions. Policymakers use macroeconomic indicators to understand economic conditions, plan fiscal and monetary policies, and manage public debt. Seeing how these indicators change compared to other regions helps governments measure performance and spot early signs of trouble (Ullah et al, 2024). For investors, trends in macroeconomic data affect how they assess country risks, diversify portfolios, and make long-term investment choices (Eldomiaty et al, 2024). International organizations rely on regional comparisons to guide development programs, decide where to allocate resources, and monitor imbalances in the global economy (Shon et al, 2024).

However, decision makers often face incomplete or inconsistent analyses. Many studies look at single countries or focus on individual indicators which makes it hard to assess the wider regional picture. This project tackles that problem by using a consistent approach across a large dataset covering many countries which makes it possible to compare regions meaningfully and draw clear insights.

From a technical perspective, this project shows how data analytics and time series modeling can be applied to macroeconomic data. Working with economic data from many countries brings challenges such as missing values, inconsistent reporting, and differences in how indicators are measured. Careful data cleaning, validation, and transformation are needed to make sure the analysis is reliable. Forecasting economic indicators also requires models that can handle uncertainty, changing structures, and evolving trends. This project uses ARIMA (AutoRegressive Integrated Moving Average), a widely used method for predicting future values based on past trends, and Bayesian Structural Time Series (BSTS). These approaches provide interpretable and probabilistic predictions that are well suited to economic data.

1.3 Problem Statement and Project Objectives

Even though global macroeconomic data is becoming more widely available, there are few studies that combine regional economic comparisons with reliable forecasting using consistent methods. Many existing studies simply describe past trends without looking ahead, while others focus on forecasting without providing a clear comparison across regions. This makes it harder for decision makers to understand past performance and plan for the future.

The main problem this project tackles is the lack of a unified data-driven approach that both analyses regional economic performance and forecasts key macroeconomic indicators. The primary goal of the project is to study and compare economic performance across regions using macroeconomic indicators and to build models that predict future economic trends.

To reach this goal, the project follows several specific objectives. These include examining past trends and relationships among important indicators like GDP growth, GDP per capita, comparing economic performance across regions using consistent measures, developing and testing a baseline model, ARIMA and Bayesian Structural Time Series forecasting models for selected indicators; and turning the results into insights that help support better economic decision making.

1.4 Scope of the Study

The scope of this study is designed to balance detailed analysis with practical usefulness. The dataset includes over 200 countries. This allows for wide regional comparisons and reduces geographic bias. The study focuses on yearly data from 2010 to 2024 which makes it possible to look at long-term trends rather than short-term changes. While the dataset covers many macroeconomic indicators, the focus is on those most relevant to economic performance and stability, such as GDP growth, iGDP per capita and GDP Current.

The forecasting part of the study looks at short- to medium-term predictions using ARIMA and BSTS. The study does not try to prove cause-and-effect or evaluate specific policies. Instead, it takes a predictive and comparative approach with the aim to find patterns and likely future paths based on historical data. These limits help ensure the results are consistent with the methods used and the quality of the data.

1.5 Expected Contributions

This project contributes both analytically and practically. On the analytical side, it shows a clear workflow for working with large macroeconomic datasets including preparing the data, exploring trends, and forecasting using time series models. It also shows how Bayesian Structural Time Series models can be useful for economic forecasting because they capture uncertainty and make results easier to interpret. On the practical side, the project provides insights that matter to policymakers, investors, and international organizations. Regional comparisons reveal economic strengths and weaknesses and the forecasting results help with planning and managing risks. By combining past analysis with future predictions, the project improves understanding of regional economic performance and supports decisions based on data.

Chapter 2 LITERATURE REVIEW

2.1 Theoretical Foundations of Macroeconomic Forecasting

Macroeconomic forecasting depends on understanding how economies work and how key indicators like GDP, inflation, unemployment, and public debt relate to each other. Economic theories show that these variables are linked through production, consumption, and financial markets (Dou et al, 2020). Classical growth theory focuses on how capital, labor, and technology drive GDP growth while Keynesian ideas emphasize changes in overall demand and the role of policy (Gali, 2020). Modern methods such as dynamic stochastic general equilibrium models try to include expectations and unexpected shocks in forecasting (Čapek et al, 2023)

Time series forecasting has become a main tool for turning these theories into practical predictions. The idea is that past patterns in economic data can help predict future trends. Methods range from simple models that look at one variable to more complex approaches that include multiple factors. These techniques help analysts identify trends and cycles in economic indicators and provide a basis for evaluating policies and planning investments (Buturac, 2022).

2.2 Empirical Studies on GDP Forecasting

Many studies have used time series models to forecast GDP in different countries, showing both the benefits and limits of these methods. Bheemanna and Megeri (2024) applied the Bayesian Structural Time Series (BSTS) model to predict India's GDP. Their research showed that BSTS can capture underlying patterns in economic data and provide forecasts that include uncertainty. The model achieved a low Mean Absolute Percentage Error (MAPE) of 12.9966 for GDP, demonstrating good accuracy and usefulness for medium- and long-term predictions especially in emerging economies. BSTS also helps policymakers see a range of possible outcomes instead of relying on a single number.

Similarly, Hassan and Mirza (2021) used a combination of ARIMA and Prophet models to forecast India's GDP growth. The ARIMA model showed a mean error of 0.07557, a root mean square error of 3.3146, a mean absolute error of 2.2993, a mean absolute percentage error of 103.1148, and a mean absolute scaled error of 0.8723, indicating a good fit when historical trends were stable. The Prophet model complemented ARIMA by capturing seasonal patterns and non-linear trends. Their study highlighted the advantage of combining models, using ARIMA for linear trends and Prophet for handling sudden changes and more flexible patterns.

Furthermore, Cao Dingshan (2025) used ARIMA models to forecast real GDP in the United Kingdom, Germany, and France with quarterly data from 2015 to 2024. The model was able to capture both linear trends and economic cycles effectively. In Germany, ARIMA achieved a MAPE below 2 percent, showing stable economic performance. In the UK, the forecasts predicted a 2.3 percent drop in GDP because of post-Brexit trade issues and rising inflation. The study showed that ARIMA is easy to interpret and works well for stable economies, but it can be sensitive to sudden structural changes.

Liu (2025) compared ARIMA and Linear Regression to forecast China's GDP over five years. For 2028, ARIMA predicted 1,031,800 billion CNY while Linear Regression forecasted 1,663,300 billion CNY. Both models offered useful insights into structural trends, showing that no single method works best in every case and that model choice depends on country-specific dynamics and the forecast period.

Also, McCloskey and Remor (2024) tested ARIMA, Vector Autoregression (VAR), and Linear Regression for the UAE. ARIMA was the most reliable for long-term forecasts with MAPE between 3.8 and 9 percent, while Linear Regression worked better for short-term scenario-based predictions. VAR gave insights across multiple variables but was less flexible during sudden structural changes, highlighting the importance of considering post-crisis or post-pandemic shifts when forecasting.

Andrianady (2023) also compared ARIMA, VAR, and MIDAS for Madagascar and found ARIMA performed best with a mean absolute error of 49.79, MAPE of 4.38, and RMSE of 58.03. Bäurle et al. (2020) applied VAR(4) to sectoral GDP in Switzerland and the eurozone, achieving RMSE values between 1.86 and 2.25. This showed that VAR is strong for capturing inter-sectoral relationships, although it is more complex to implement.

Finally, Jonathan (2024) compared VAR and N-HiTS for US GDP forecasts from 2024 to 2029. N-HiTS slightly outperformed VAR in handling complex non-linear trends in high-dimensional data, with mean square error 2.8606, mean absolute error 0.7298, and RMSE 1.6913 for N-HiTS, compared to 2.9270, 0.7111, and 1.7109 for VAR.

2.3 Advanced Data Analytics Techniques in Macroeconomic Forecasting

Beyond traditional models like ARIMA and Linear Regression, Bayesian Structural Time Series (BSTS) has become popular for economic forecasting because it uses a probabilistic approach and is flexible. BSTS breaks down time series into trend, seasonal, and regression parts and provides a range of possible outcomes for predictions. This is especially helpful for countries experiencing structural changes, policy shifts, or sudden economic shocks. Bheemanna and Megeri (2024) showed that BSTS can forecast GDP while giving a clear measure of uncertainty and this can help policymakers understand a range of likely growth scenarios instead of relying on a single fixed forecast.

Other advanced forecasting methods include VAR models, which look at how multiple economic indicators influence each other. VAR has been widely used in European economies to study sectoral contributions to GDP and to test different policy scenarios (Bäurle et al., 2020). Mixed Data Sampling (MIDAS) methods extend VAR and ARIMA to work with variables collected at different frequencies.

Kamolthip (2021) used Long Short-Term Model (LSTM) with U-MIDAS to forecast Thailand's GDP. This shows its usefulness for short-term predictions and nowcasting with mixed-frequency data. Although LSTM performs well with high-frequency data, its complexity and difficulty in interpretation often make probabilistic methods like BSTS better for long-term strategic forecasting.

Neural hierarchical models such as N-HiTS have also been suggested for macroeconomic forecasting. They can capture complex non-linear patterns in high-dimensional time series and can be more accurate than traditional methods, especially in volatile economies or when many leading indicators are included (Jonathan, 2024). However, these models need large amounts of training data and computing power, and they are harder to interpret than ARIMA or BSTS.

2.4 Comparative Insights from Existing Studies

Research consistently shows that no single model works best for all GDP forecasting. ARIMA works well in stable economies with linear trends and predictable cycles, while BSTS provides forecasts that show uncertainty and can handle structural changes. Linear Regression and VAR are useful for scenario analysis and looking at multiple variables, but they depend heavily on the quality of the input data. Newer neural models like N-HiTS and LSTM are promising for short-term, high-frequency forecasts, yet they are often too complex and hard to interpret for policy purposes.

Empirical studies highlight that ARIMA remains the most widely tested and reliable baseline model for GDP forecasting (Cao, 2025; Hassan & Mirza, 2021; Andrianady, 2023). BSTS adds value by offering uncertainty estimates and breaking the time series into interpretable components, which helps policymakers understand possible outcomes (Bheemanna & Megeri, 2024). Performance varies across countries and forecast periods, so combining models according to the forecast horizon and economic context can improve accuracy (McCloskey & Remor, 2024). It is also important to account for structural breaks, such as economic shocks or pandemics, because standard models may otherwise underestimate volatility and produce biased forecasts (McCloskey & Remor, 2024; Cao, 2025).

2.5 Gaps in Literature

Despite extensive research, several gaps still exist. Many studies focus on a single country, which limits understanding of cross-country or regional differences. While ARIMA and BSTS models have been studied on their own, few works combine these approaches to compare historical trends with probabilistic forecasts across multiple countries at the same time. In addition, few studies translate their findings into policy guidance or practical business insights, which means decision makers often lack actionable information.

This project tackles these gaps by using a consistent dataset covering more than 200 countries from 2010 to 2024 to enable meaningful cross-country comparisons. Forecasting focuses on the top five countries by average GDP, combining ARIMA and Bayesian Structural Time Series to produce interpretable and probabilistic predictions for key macroeconomic indicators. The analysis is linked to practical implications for policymakers, investors, and international organizations, ensuring the results go beyond purely academic insights.

Chapter 3 METHODOLOGY

This study takes a structured approach to examine regional economic performance using macroeconomic indicators from 2010 to 2024. The methodology is designed to maintain data accuracy, ensure careful preprocessing, and support reliable model results. It follows standard practices in economic data analysis and forecasting and focuses mainly on GDP, GDP per capita, and GDP growth as the key indicators.
The methodology follows three main phases. The first is data collection, which involves gathering comprehensive macroeconomic information from reliable sources. The second is data preprocessing, where the data is cleaned, transformed, and missing values are addressed to make it consistent. The third is data validation, which checks that the data is reliable and suitable for analysis and forecasting. Each of these phases is essential to ensure that the following analyses, including ARIMA, and BSTS forecasting, produce accurate and meaningful results.

3.1 Data Collection

The primary dataset for this project comprises year-wise macroeconomic indicators for over 200 countries spanning 2010–2025. The dataset is sourced from Kaggle, originally extracted from the World Bank Open Data API. The dataset includes crucial economic variables such as:

- GDP (current USD): total value of goods and services produced. It provides a snapshot of economic size.
- GDP per capita: GDP divided by population offering insight into average individual wealth and living standards.
- GDP growth (%): annual percentage change in GDP which indicates economic momentum.
- Inflation (CPI %): consumer price index that measures price changes in a basket of goods and services.
- Unemployment rate: proportion of the labor force without work. This shows labor market health.
- Real interest rate: the cost of borrowing adjusted for inflation. It is relevant for investment and policy analysis.
- Public debt (% of GDP): government debt as a share of economic output that reflects fiscal sustainability.
- Government expenditures and revenues: this also reflects fiscal policy and government activity.
- Current account balance: net trade and capital flows which indicates international competitiveness.

- Gross national income: total income earned by a country's residents.
- Tax revenue: financial resources collected by the government.

The dataset's comprehensiveness allows for regional and cross-country comparisons and serves as a reliable base for forecasting GDP growth, inflation, and related indicators. The Kaggle dataset can be accessed at https://www.kaggle.com/datasets/tanishksharma9905/global-economic-indicators-20102025

3.2 Data Preprocessing

3.2.1 Initial Data Review

An initial look at the dataset showed missing values in several macroeconomic indicators, which is common in large datasets covering many countries over long periods. Although many variables are available, this study focuses on three main indicators: GDP in current USD, GDP per capita, and GDP growth. These indicators are chosen for several reasons.

First, they are highly relevant to economic performance (Alif, 2024). GDP measures the total output of a country and is a widely accepted indicator of national economic strength. GDP per capita adjusts for population size and provides insights into average wealth and living standards, which is important for comparing regions. GDP growth shows how fast an economy is expanding or contracting each year, indicating economic momentum and stability. Together, these indicators capture both the absolute size of an economy and its relative development and they provide a strong foundation for cross-country analysis (Adewole, 2024).

Also, these indicators are complete and more reliable than others. Variables like government spending, current account balances, or tax revenue are often missing or inconsistently reported, especially for smaller or less-developed countries. Focusing on GDP, GDP per capita, and GDP growth ensures the analysis uses data with broad coverage and high reliability.

Furthermore, these indicators are important for forecasting. Predictive models of regional economic performance need to capture the main trends and dynamics of key economic measures. GDP and GDP growth provide essential inputs for projecting future economic trends, while GDP per capita allows meaningful comparisons across countries. Concentrating on these indicators helps models to focus on the most informative variables without adding noise from less reliable data.

Finally, the selection aligns with the research objectives. The study aims to compare economic performance across countries and forecast future trends. GDP, GDP per capita, and GDP growth directly relate to understanding regional growth, economic disparities, and predicting future outcomes. Focusing on these indicators ensures that data preprocessing and modeling remain closely connected to the project goals.

Focusing on these three indicators helps the study stay methodically sound while dealing with practical problems in large and varied datasets. This focus also makes it easier to handle missing data accurately and lowers the chance of bias from less important or incomplete variables.

3.2.2 Handling Missing Values

<u>Removal of 2025 Data</u>

The dataset had entries for 2025 which were mostly empty because the official macroeconomic data for that year were not available when the data was collected. Using 2025 data could have added incomplete or uncertain information. Therefore, all records for 2025 were left out of the analysis.

<u>Handling Missing GDP and GDP Per capita</u>

For GDP (current USD) and GDP per capita, missing values were filled in using extra data from the World Bank DataBank. This made the historical data for each country as complete as possible. Any remaining missing values were replaced with the country's average across the years that had data, which keeps the country's trends without adding random numbers.

Some countries needed special attention because of naming differences. For example, Korea was listed differently in the dataset compared to the World Bank, so the name had to be changed. The British Virgin Islands and Gibraltar did not have enough historical data for GDP and GDP per capita, so they were removed from the analysis.

<u>Handling Missing GDP Growth</u>

Annual GDP growth rates were recalculated from the GDP (current USD) values for each country to address any remaining gaps. The formula used is:

$$GDP\ Growth\ (\%) = \frac{GDP_r - GDP_{t-1}}{GDP_{t-1}} \times 100$$

This method makes sure that growth rates show real economic changes and are not distorted by estimated values to make sure that the analysis is accurate and reliable.

3.3 Data Validation

Several checks were done to make sure the cleaned dataset was reliable and ready for forecasting.

<u>Cross-Referencing</u>

GDP and GDP per capita values were checked against data from the World Bank and IMF to identify any differences or discrepancies. This ensured the numbers used were accurate and aligned with trusted sources.

<u>Temporal Consistency</u>

Year-on-year growth rates were reviewed to ensure economic changes were realistic. Any sudden or implausible jumps in the data were flagged and corrected to maintain logical trends over time.

Correlation Checks

Relationships between key indicators, such as GDP and GDP per capita or GDP growth and unemployment, were examined to ensure they remained logical after missing values were filled. This helped confirm that the data maintained its internal consistency.

These validation steps together confirmed that the dataset was complete, consistent, and suitable for both descriptive analysis and forecasting using models like ARIMA and BSTS.

Chapter 4 EXPLORATORY DATA ANALYSIS

The exploratory data analysis (EDA) in this study concentrates on the three macroeconomic indicators selected for the project: GDP (Current USD), GDP per Capita (Current USD), and GDP Growth (% Annual). The purpose is to investigate patterns, trends, and correlations within these indicators across more than 200 countries from 2010 to 2024.

4.1 Correlation Analysis

A correlation matrix was created to see how the selected indicators relate to other key economic variables like unemployment. The results show that GDP and GDP per capita have only a weak connection, meaning that a large economy does not always give high income per person. GDP growth is mostly independent of total GDP and per-capita income, which means short-term changes in growth can happen regardless of an economy's size or average wealth. Negative correlations with unemployment support well-known economic ideas: higher unemployment usually comes with lower output and slower growth. These results show that GDP, GDP per capita, and GDP growth measure different aspects of economic performance. Each captures a unique dimension, which highlights the need to consider all three in both analysis and forecasting.

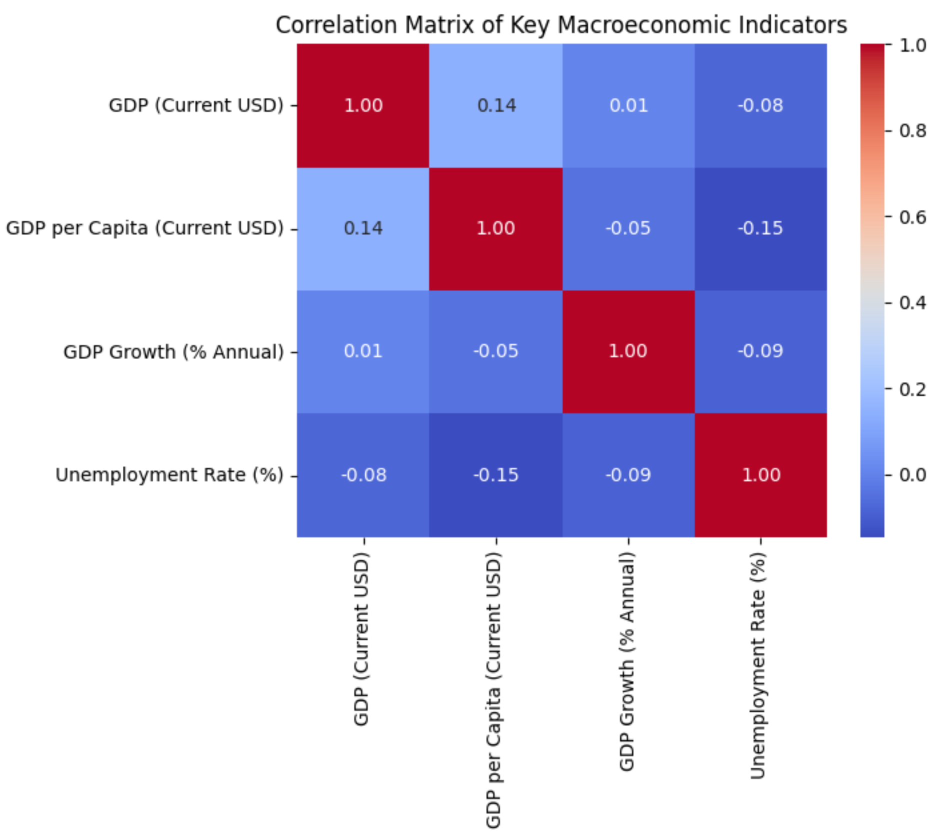

Figure 4-1 Correlation Heatmap of Key Macroeconomic indicators

4.2 Distributional Insights

GDP (current USD) demonstrates a highly skewed distribution, with a few countries making up most of the world's total output. Looking at 3,219 country-year observations, the average GDP is about 403 billion USD, while the median is much lower at 26.3 billion USD. This shows that large economies like the United States, China and so on dominate global output while most other countries contribute far less. The smallest GDP which was around 32 million USD is found in very small or emerging economies, while the largest of over 29 trillion USD, is the United States which alone represents a huge share of world GDP. This pattern reflects the concentration of industry, technology, and financial markets in a few nations. It also shows that comparing total GDP without considering per capita can give a misleading picture of individual economic prosperity.

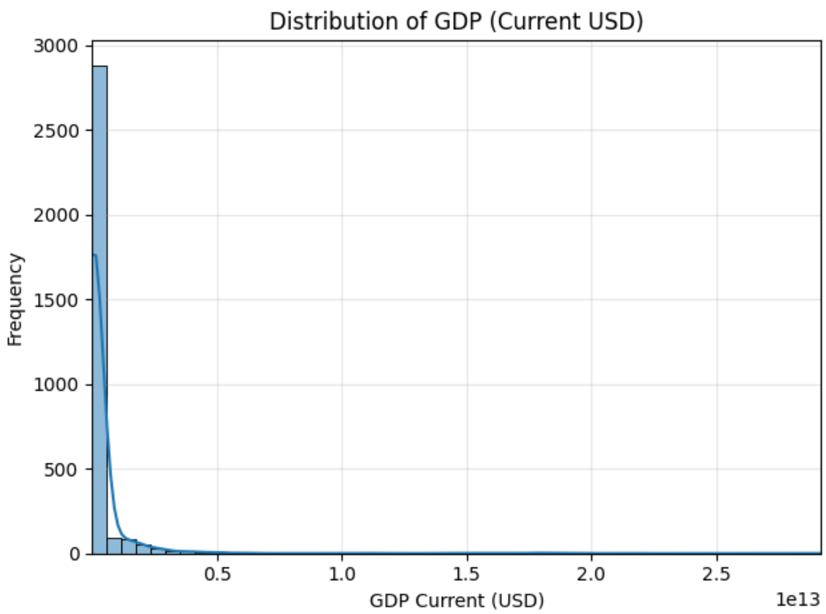

Figure 4-2 Distribution of GDP (Current USD)

GDP per capita (Current USD) shows the differences in individual economic well-being across countries. The average GDP per person is about $18,700, while the median is much lower at $7,051. This insight shows that the distribution is skewed by very wealthy small countries and developed nations. The lowest per-capita GDP was around $154. This represents low-income countries with limited industry and services, while the highest exceeds $256,000 representing microstates like Monaco. This wide range highlights major differences in living standards and wealth, reflecting global income inequality reported by the World Bank. Factors such as education and skills, technology use, natural resources, and quality of governance help explain these gaps.

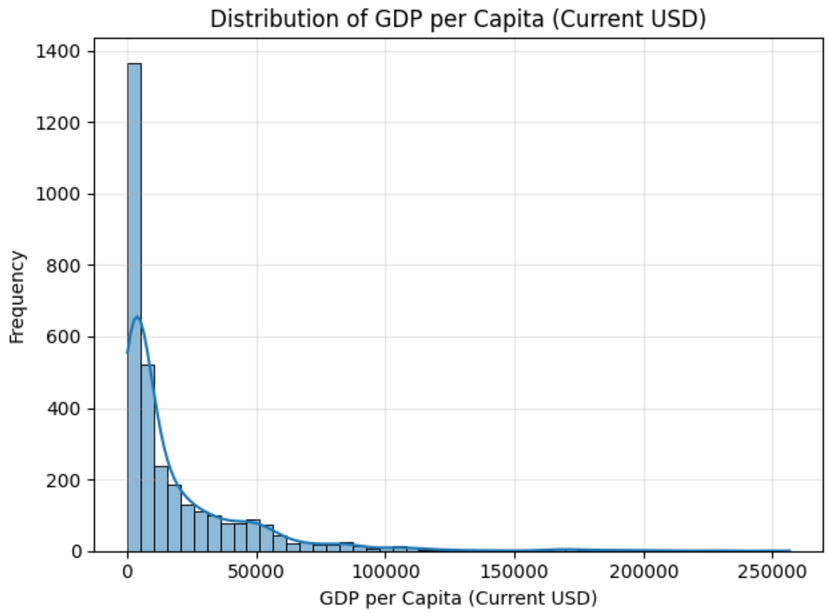

Figure 4-3 Distribution of GDP per capita

GDP growth (%) measures the annual percentage change and shows how economies expand or contract over time. The analysis of this indicator reveals that most countries grow at a moderate pace with an average rate of about 2.92% and a median of 3.13%. Some countries experience extreme changes such as a drop of 54.34% during severe crises or a surge of 104.16% in short-term booms often linked to resource discoveries or unusual data. These wide swings indicate that while mature economies tend to grow steadily and emerging markets can face more volatility but also have the potential for higher growth. External shocks, natural disasters, and dependence on commodities contribute to these differences across countries.

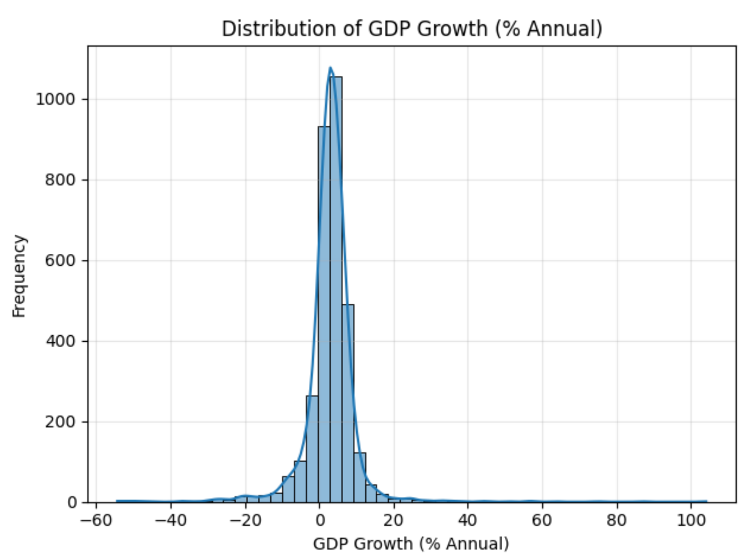

Figure 4-4 Distribution of GDP growth

Table 4-1 Descriptive Statistics Table

Statistic	GDP (Current USD)	GDP per Capita (Current USD)	GDP Growth (% Annual)
count	3,219	3,219	3,219
mean	403,323,500,000.00	18,699.74	2.92
std	1,786,532,000,000.00	27,447.05	7.09
min	32,105,410.00	153.93	-54.34
25%	6,216,871,000.00	2,307.35	0.81
50%	26,344,570,000.00	7,051.20	3.13
75%	191,136,400,000.00	24,284.16	5.55
max	29,184,890,000,000.00	256,580.52	104.16

Overall, these patterns show that most of the world's economic output comes from a few large countries, wealth is unevenly spread among individuals, and growth rates differ widely because of both structural factors and external events. Looking at GDP, GDP per capita, and GDP growth together gives a clear picture of the size of economies, the well-being of people, and how economies change over time.

4.3 Temporal Trends

GDP (Current USD) from 2010 to 2024 shows a steady upward trend that reflects overall global economic growth. There were slowdowns in 2014 and 2015, likely caused by changes in global commodity prices and a sharp drop in 2019 and 2020 due to the COVID-19 pandemic. After 2020, global GDP recovered quickly as fiscal support, trade resumption, and higher consumer activity boosted economic activity.

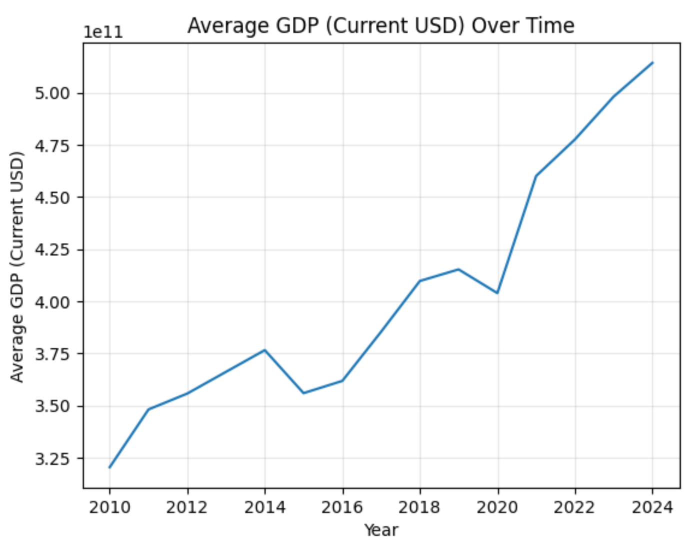

Figure 4-5 Average GDP Over Time

GDP per capita (current USD) also rises overall but shows sharp drops in 2014–2015 and in 2020 during the pandemic. The quick recovery after 2020 highlights the strength of advanced economies which benefited from strong policies and diverse economic structures. Changes in per-capita income show how global crises can affect individual wealth differently across countries.

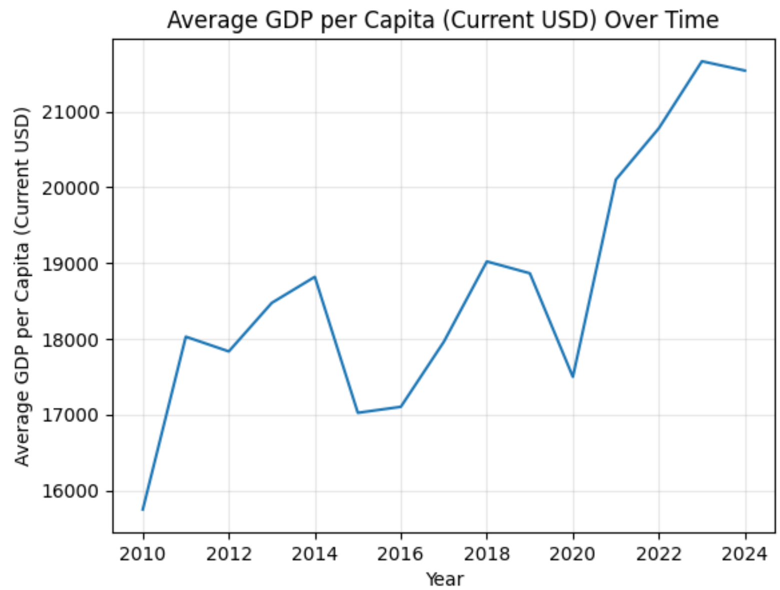

Figure 4-6 Average GDP per capita over time

GDP growth (%) per year is more volatile than total GDP or per-capita measures. From 2010 to 2019, growth was fairly steady. This steadiness indicates long-term global expansion. In 2020, the pandemic caused a sharp decline that matches worldwide contractions reported by international organizations. After 2020, growth rebounded quickly in a "V-shaped" pattern as it rrose above pre-pandemic levels with differences across countries depending on their economic structure and policy responses.

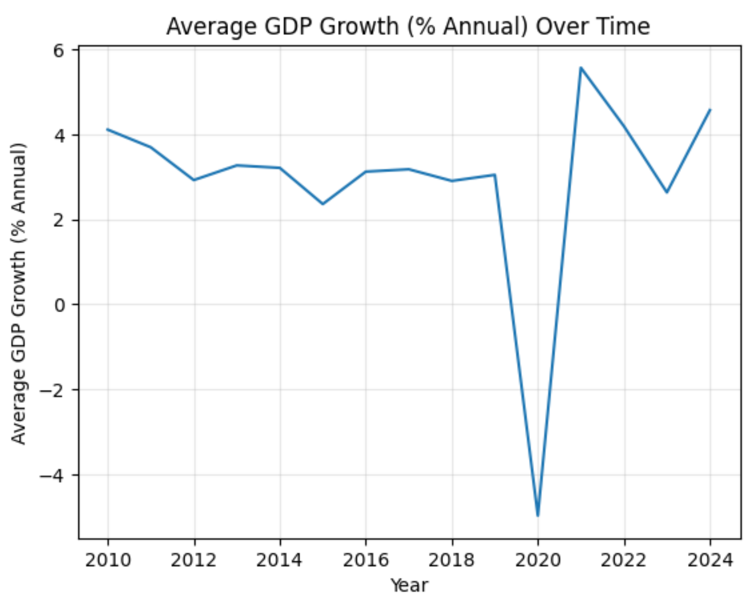

Figure 4-7 Average GDP growth over time

4.4 Country-Wise Analysis

Country-level patterns show large differences in economic performance and highlight how factors like structure, population, and policy affect outcomes. Small, wealthy countries such as Monaco, Liechtenstein, and Luxembourg have the highest GDP per capita. Their high productivity, advanced financial services, and concentrated wealth was reflected in this insight. Their GDP growth is moderate, which is typical for mature and stable economies with limited room to expand. Strong institutions, low population growth, and integration into global finance help maintain high living standards even with slower growth.

Table 4-2 Analysis Results of Top countries by GDP per capita

Country	GDP (Current USD)	GDP per Capita (Current USD)	GDP Growth (% Annual)
Monaco	7,050,607,000.00	190,528.21	2.67
Liechtenstein	6,616,728,000.00	173,625.63	2.35
Luxembourg	70,697,590,000.00	118,670.67	2.68
Bermuda	7,111,666,000.00	111,089.63	0.22
Switzerland	743,561,600,000.00	88,150.10	2.12

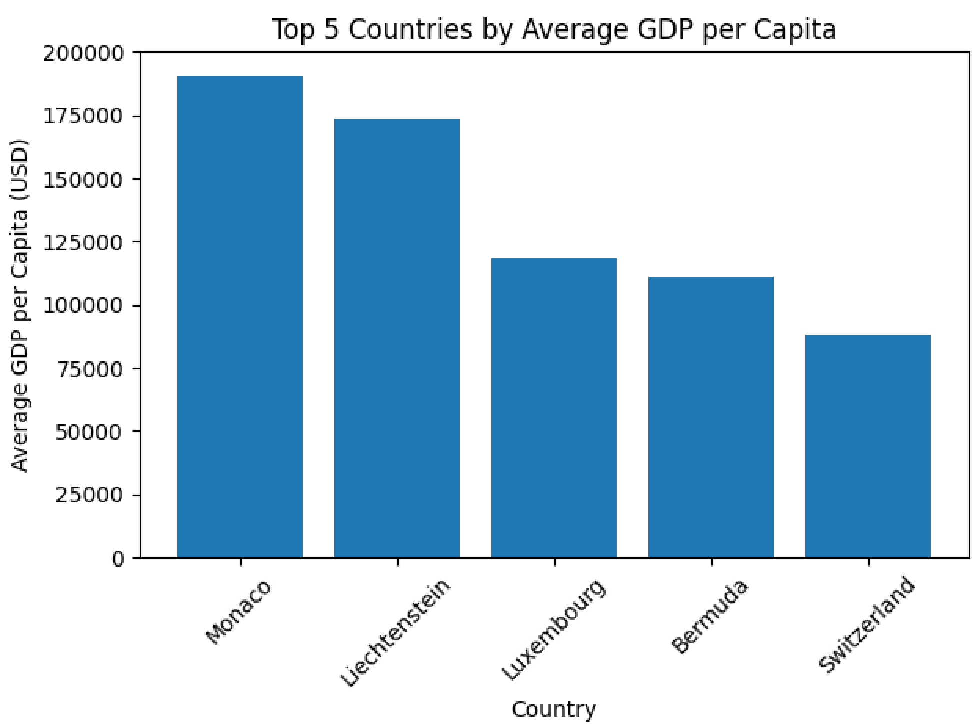

Figure 4-8 Top 5 countries by average GDP per capita

Large economies like the United States, China, Japan, Germany, and the United Kingdom lead in total GDP because of their big populations, strong industrial bases, and advanced technology. Income per person varies across these countries showing that a large economy does not always mean high individual wealth. For example, China's high total GDP contrasts with a lower per-capita GDP due to its large population. Growth in these nations is relatively steady as supported by different economies that can handle shocks from crises, trade disruptions, and geopolitical events. Policies, investment in innovation, and economic diversification help these countries remain resilient.

Table 4-3 Analysis Results of Top countries by GDP

Country	GDP (Current USD)	GDP per Capita (Current USD)	GDP Growth (% Annual)
United States	20,549,850,000,000.00	63,062.27	2.57
China	12,795,060,000,000.00	9,180.34	6.73
Japan	5,032,637,000,000.00	39,729.59	0.59
Germany	3,935,740,000,000.00	47,823.70	1.56
United Kingdom	2,913,646,000,000.00	44,173.00	2.06

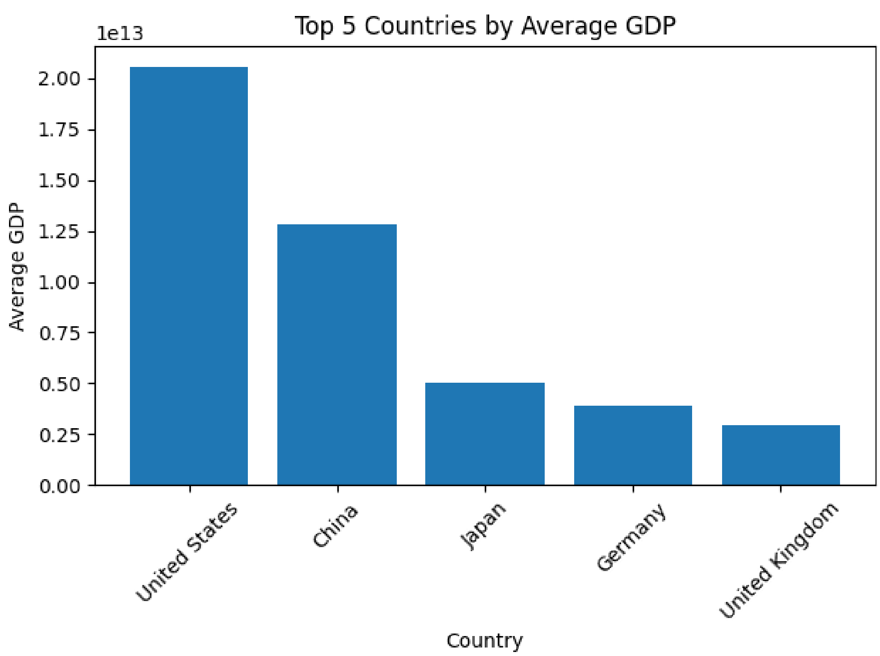

Figure 4-9 Top 5 countries by average GDP

Fast-growing countries such as Guyana, Djibouti, Tajikistan, Turkmenistan, and Mongolia see high GDP growth which is often driven by resource use, foreign investment, infrastructure projects, or economic reforms. These countries usually rely on one main sector like mining, energy, or agriculture, which allows rapid expansion but also creates vulnerability to price swings, climate events, or political instability. Despite strong growth, their total GDP and income per person remain lower than in developed nations, showing that fast growth does not automatically lead to broad prosperity. Long-term development requires structural reforms, investment in skills, and economic diversification to reduce risks from external shocks.

Table 4-4 Analysis Results of Top countries by GDP growth

Country	GDP (Current USD)	GDP per Capita (Current USD)	GDP Growth (% Annual)
Guyana	7,545,425,000.00	9,423.08	16.23
Djibouti	2,791,052,000.00	2,588.18	8.81
Tajikistan	8,685,038,000.00	946.59	7.89
Turkmenistan	43,017,490,000.00	6,511.49	7.85
Mongolia	13,734,950,000.00	4,317.27	6.93

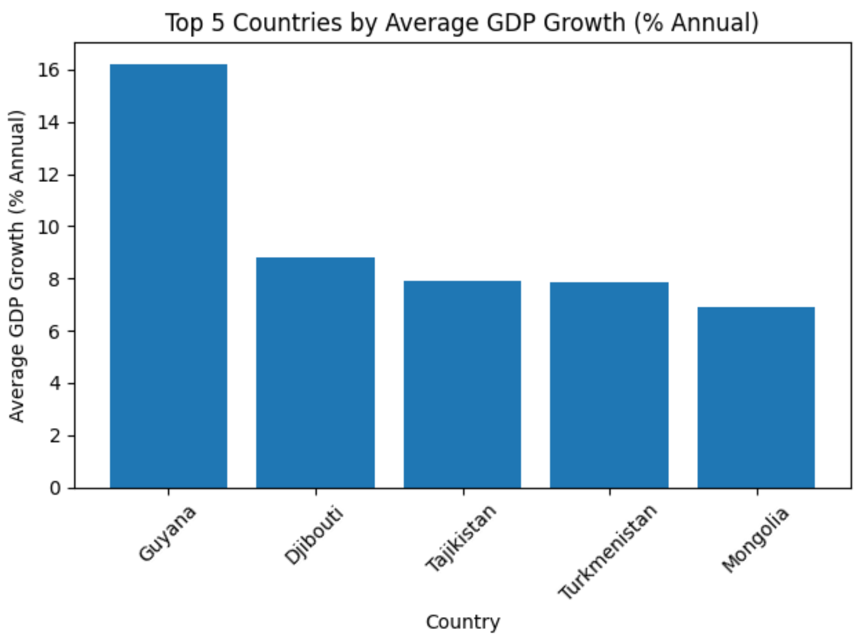

Figure 4-10 Top 5 countries by average GDP growth

4.5 Insights and Contextual Interpretation

The exploratory analysis highlights several important patterns supported by external sources. The sharp drop in global GDP in 2020 matches the pandemic-related declines reported by the IMF and World Bank while the quick recovery reflects the impact of worldwide government and monetary support.

GDP per capita shows persistent global inequality as confirmed by the findings from the OECD and World Bank. Smaller countries face larger swings in GDP growth and this shows their sensitivity to global markets and external shocks. In contrast, large and diversified economies tend to grow steadily and remain resilient. It is indicating their structural and institutional strengths.

The weak links between GDP, GDP per capita, and growth show that each measure tells a different story. A large economy does not necessarily have high income per person or rapid growth. Examining all three together gives a clearer and more accurate view of regional economic performance.

Overall, this exploratory analysis gives a clear picture of trends, differences, and relationships in the selected indicators. It sets the stage for predictive modeling and supports the use of reliable forecasting methods so that later analyses reflect real economic patterns. The insights from this analysis provide both descriptive and contextual understanding and create a strong foundation for the modeling and forecasting parts of the project.

Chapter 5 IMPLEMENTATION OF ALGORITHMS AND MODELS

5.1 Overview

This stage of the project focuses on building and testing time series models to predict future economic performance using GDP in current US dollars. The analysis covers five major economies: the United States, China, Japan, Germany, and the United Kingdom. Forecasts are made for each country on its own. Each national GDP series is treated as a separate time series rather than being combined with others. This approach keeps the unique features of each economy such as size growth patterns and how countries react to shocks which is important for macroeconomic forecasting.

For each country three forecasting methods are applied: the Theta method, Autoregressive Integrated Moving Average also called ARIMA, and Bayesian Structural Time Series or BSTS. Using the same set of models for every country makes it possible to compare how well each method performs within a single economy. It also avoids problems that can arise when very different economies are grouped together.

The modeling has two main goals. First each model is tested on how well it predicts GDP during a holdout period from 2020 to 2024. This period includes major economic disruptions especially the COVID 19 shock which makes it a strong test of model reliability. Second, the model that performs best for each country is trained again using all available historical data and then used to produce GDP forecasts up to 2030.

Model training uses yearly GDP data from 2010 to 2019 while testing uses data from 2020 to 2024. Performance is measured using Mean Absolute Percentage Error Mean Absolute Error and the R squared value. This setup ensures that model choice is based on real forecasting accuracy for each country rather than on theory alone.

By using country level forecasts and testing different types of models this chapter provides a clear and reliable base for long term GDP projections and for comparing results across the world's largest economies.

5.2 Algorithm and Model Development

5.2.1 Theta Model

The Theta method is a simple time series forecasting method that works with one variable at a time. It breaks a time series into several trend lines called theta lines. Each line shows a different view of the trend. These lines are projected forward on their own and then combined to form the final forecast. The method is well known for performing well in forecasting competitions and for capturing long term trends and short-term movements without much tuning (Petropoulos et al, 2021).

The Theta method is used here as a baseline model. Its simple structure, strong stability and low computing cost make it a good reference point when comparing results with more advanced models. For GDP forecasting this method fits well because GDP data usually follows a clear trend and changes smoothly over time.

The model was applied using Python and was fitted separately to each country's GDP data. The default Theta settings were used since the method is designed to work well across many types of economic time series without the need for complex adjustments.

5.2.2 Autoregressive Integrated Moving Average (ARIMA)

The ARIMA model is a well-known time series forecasting method. It predicts future values by learning from past values and from past shocks in the data. This makes it a good fit for economic data where what happened before often affects what happens next (Cao, 2025).

ARIMA is described using three parts. The first part p shows how many past values are used. The second part d shows how many times the data is different to remove trends. The third part q shows how many past errors are taken into account (Alif, 2025).

To find the best ARIMA setup for each country a structured search was carried out. Values of p were tested from zero to five. Values of d were tested from zero to four. Values of q were tested from zero to five. This range is wide enough to capture trends and short-term movements while keeping the model simple.

Each ARIMA version was trained on earlier data and then tested on later years. The final choice was based on how well the model predicted new data rather than how well it fit past data. This helps avoid overfitting and improves real world performance.

ARIMA was included because it is easy to understand and explain. It clearly shows how past GDP values affect future forecasts and it also acts as a strong reference point when comparing results with more flexible models like BSTS.

5.2.3 Bayesian Structural Time Series (BSTS)

BSTS is a time series forecasting method that focuses on uncertainty as well as trends. It breaks a data series into clear parts such as long-term trend seasonal movement and random noise. It uses Bayesian reasoning which means it shows how confident the forecasts are instead of giving only one fixed number (Bheemanna & Megeri, 2024). Unlike ARIMA it does not rely on fixed relationships and can adjust when the data pattern changes.

This flexibility is important for economic forecasting. Economies often change direction after events like financial crises, pandemics or major policy decisions. BSTS is better at adjusting to these shifts than traditional linear models (Bheemanna & Megeri, 2024).

Two ways of estimating the model were tested. The first was Kalman filtering which is faster and more direct. The second was Markov Chain Monte Carlo sampling which explores uncertainty in more depth. The tests used either Kalman or MCMC with 500 or 1000 simulation steps. This setup balances accuracy with reasonable computing time.

Each country was modeled on its own using historical data. The results include both forecast values and ranges that show possible outcomes. These ranges are useful for policy and business decisions because they show risk and uncertainty rather than a single prediction.

5.3 Model Training and Forecasting Pipeline

For each country the modeling steps followed the same process. First the data from 2010 to 2019 was used to train the Theta ARIMA and BSTS models. Each model then produced GDP forecasts for the years 2020 to 2024. These forecasts were compared with the actual GDP values to see how well each model performed.

After this comparison the best model for each country was chosen based on overall accuracy. That model was then trained again using all available data from 2010 to 2024. The retrained model was used to generate GDP forecasts up to 2030.

This two-step approach makes sure the final forecasts use as much historical information as possible while still relying on models that proved they could predict unseen data well.

5.4 Optimization and Hyperparameter Tuning

Hyperparameter tuning was done for the ARIMA and BSTS models to improve forecast accuracy and to make sure the results were reliable for each country. For ARIMA a grid search was used to test different model settings in a clear and structured way. The autoregressive term p was tested from 0 to 5. The differencing term d was tested from 0 to 4. The moving average term q was tested from 0 to 5. Each p d q setting was trained using data from 2010 to 2019 and then tested on forecasts for 2020 to 2024. Forecast results were checked using MAPE MAE and R squared. This process helped identify the ARIMA structure that best matched the trend stability and volatility of each country's GDP series.

For BSTS tuning focused on how the model estimates its internal states and how much simulation is used. Two options were tested. Kalman filtering and Markov Chain Monte Carlo. For the MCMC approach the number of simulation steps was set to either 500 or 1000 to balance accuracy and computing time. Kalman filtering was faster and gave fixed updates while MCMC captured uncertainty more clearly by sampling many possible outcomes. These options were compared using out of sample forecast results to find the best balance between accuracy, clarity and efficiency.

The Theta model did not need tuning because it uses a standard setup that works well for many economic time series. It was used as a baseline so the results from ARIMA and BSTS could be compared against a simple and stable reference.

5.5 Model Evaluation Metrics

Model performance was measured using three different metrics: Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and the coefficient of determination (R2).

MAPE shows how accurate the forecasts are compared to the actual GDP values. It expresses errors as a percentage, which makes it easier to compare model performance across countries with very different GDP sizes (Andrianady, 2023). A lower MAPE means the forecast is closer to the real values.

$$MAPE = \frac{1}{n}\sum_{n=i}^{n} \left|\frac{ytrue_i - ypred_i}{ytrue_i}\right| \times 100$$

Where

n = total number of data points

$$ytrue_i = actual\ GDP\ at\ point\ i$$

$$ypred_i = predicted\ GDP\ at\ point\ i$$

MAE gives the average size of the forecast errors in absolute terms, expressed in currency units. This makes it straightforward to understand how far the model predictions are from the actual GDP in real economic terms. It is especially useful for seeing the practical impact of forecast errors on economic analysis and planning (Hassan & Mirza 2021)

$$MAE = \frac{1}{n}\sum_{n=i}^{n} |ytrue_i - ypred_i|$$

R2 measures how well the model explains the variation in observed GDP values. A higher R^2 indicates that the model captures more of the patterns and trends in the actual data, showing how much of the real-world variation the model can account for (Adewale et al, 2024).

$$R2 = 1 - \frac{\sum_{i=1}^{n} (ytrue_i - ypred_i)^2}{\sum_{i=1}^{n} (ytrue_i - ymean)^2}$$

Using all three metrics together provides a complete view of model performance. It ensures that forecasts are evaluated not just for relative accuracy but also for absolute error and explanatory power. This approach is important in economic forecasting because decision-makers need reliable information in both relative and actual monetary terms.

5.6 Benchmarking and Model Comparison

The Theta model acts as a simple benchmark, giving a reference point based on trend-focused forecasting without relying on complex assumptions. ARIMA models generally perform better than this baseline because they capture patterns from past values and short-term economic shocks. BSTS models often improve forecasts even more, especially during periods of structural change like the economic disruptions caused by COVID-19.

Evaluating all models over the same testing period keeps the comparison fair and consistent. Differences in how well models perform across countries also show that factors such as economic structure and volatility affect how easy or difficult it is to forecast GDP accurately.

5.7 Summary

This chapter explained how time series forecasting models were used to predict GDP for five major economies. By combining a simple baseline model with classical and Bayesian approaches, the analysis balances ease of understanding, clarity, and flexibility. The careful process of training, tuning, and testing ensures that forecasts up to 2030 are based on actual performance and a solid methodology.

The next chapter presents the results of these models and compares their performance, showing how accurate the forecasts are and what they mean for economic analysis.

Chapter 6 RESULTS AND ANALYSIS

The GDP trends for the five countries including the United States, China, Japan, Germany, and the United Kingdom show different economic paths from 2010 to 2024. Each country follows its own pattern influenced by national policies, population changes, and global events, especially the COVID-19 pandemic in 2020. Using Theta, ARIMA, and BSTS models to forecast provides a view of expected GDP paths through 2030 and shows how the accuracy of predictions differs between models.

6.1 Models Performances

United States

The United States shows steady GDP growth from 2010 to 2019, rising from about 15 trillion dollars to 21.5 trillion dollars. In 2020, GDP fell temporarily due to the global economic impact of COVID-19, but it quickly recovered and reached nearly 29.2 trillion dollars by 2024.

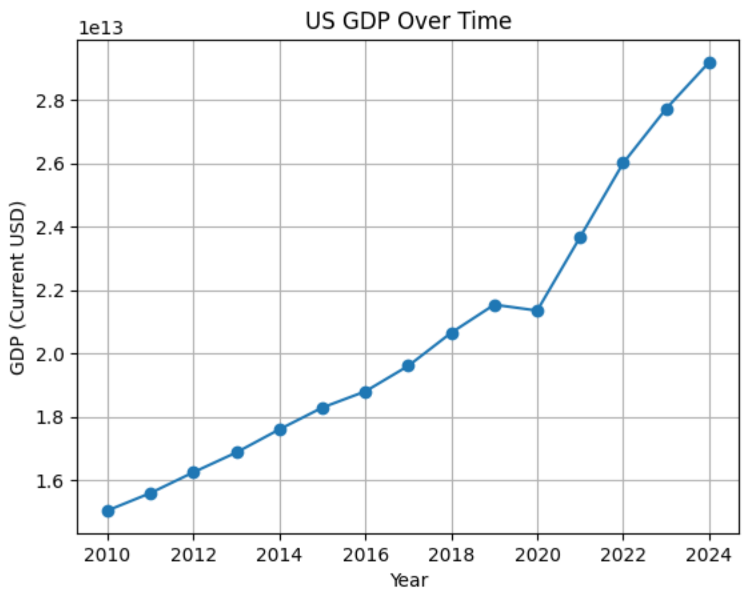

Figure 6-1 USA GDP Over Time

Forecast performance differs among models. The ARIMA model with order 2,1,3 achieves the lowest MAPE of 2.35% and an R2 of 0.94, showing strong predictive accuracy. The Theta model is stable but has higher error with a MAPE of 11.79% and R2 of -0.81, showing it struggles with sharp economic shocks. The BSTS model captures uncertainty well with wide prediction intervals but underperforms in point forecasts, reaching a MAPE of 15.13% and R2 of -2.10. The Forecasts to 2030 show steady GDP growth, but BSTS intervals suggest possible fluctuations, highlighting long-term uncertainty.

Table 6-1 USA Forecasted GDP

Year	Theta Model (Trillion USD)	ARIMA (Trillion USD)	BSTS (Trillion USD)
2025	21.89	30.49	29.10
2026	22.25	31.65	29.10
2027	22.60	32.67	29.10
2028	22.96	33.57	29.10
2029	23.31	34.36	29.10
2030	23.67	35.06	29.10

China

China's GDP grows quickly from around 6 trillion dollars in 2010 to 13 trillion dollars in 2019, driven by expansionary policies, industrial output, and trade. The 2020 slowdown reflects COVID-19 impacts, and recovery afterward shows volatility due to changes in global demand and domestic regulations.

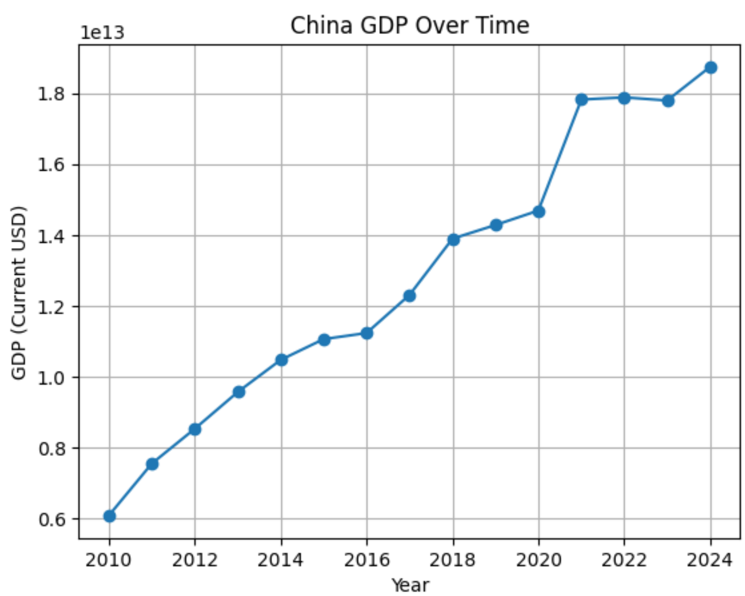

Figure 6-2 China GDP Over time

The ARIMA model with order 4,1,4 performs best with a MAPE of 2.97% and R2 of 0.667, capturing underlying trends despite some outliers in forecasts for 2026 and 2030 caused by extreme historical fluctuations. Theta shows a MAPE of 10.08% and R^2 of -1.16 while BSTS has a MAPE of 17.32% and R2 of -4.98, struggling with rapid shifts. Long-term forecasts suggest continued GDP growth but with cyclical adjustments and uncertainty in global demand.

Table 6-2 China Forecasted GDP

Year	Theta Model (Trillion USD)	ARIMA (Trillion USD)	BSTS (Trillion USD)
2025	14.71	25.72	18.73
2026	15.14	18.83	18.73
2027	15.58	25.28	18.73
2028	16.01	22.27	18.73
2029	16.44	41.90	18.73
2030	16.87	5.01	18.73

<u>Japan</u>

Japan shows a stable but low-growth GDP path. GDP rose from about 5.76 trillion dollars in 2010 to 5.12 trillion dollars in 2019 followed by slight contractions from 2020 to 2024 which was influenced by demographic stagnation and low domestic consumption.

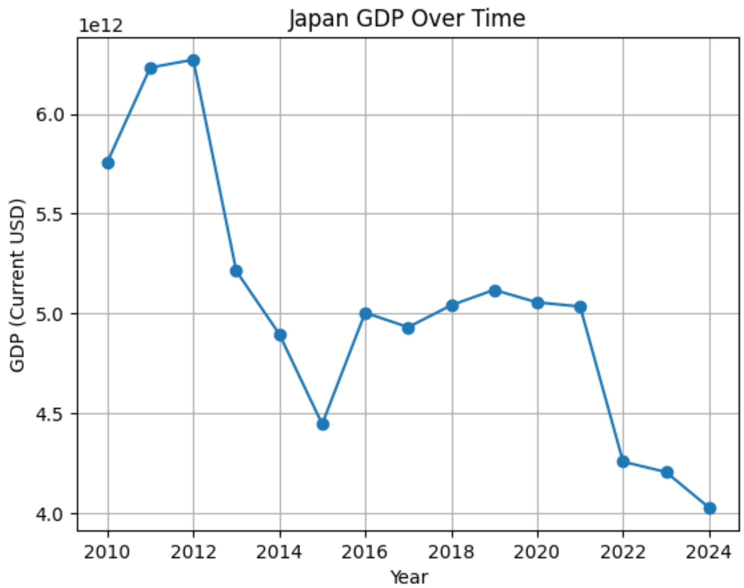

Figure 6-3 Japan GDP over time

ARIMA with order 4,1,4 performs best with a MAPE of 2.97% and R2 of 0.667, capturing small variations effectively. Theta has a MAPE of 10.18% and R2 of -0.49 while BSTS reaches 13.13% and R2 of -1.55 with wide intervals showing potential GDP ranges under uncertainty. Forecasts indicate slow growth through 2030 consistent with Japan's mature economy and limited expansion.

Table 6-3 Japan Forecasted GDP

Year	Theta Model (Trillion USD)	ARIMA (Trillion USD)	BSTS (Trillion USD)

2025	5.05	3.98	4.10
2026	4.99	4.34	4.10
2027	4.92	4.31	4.10
2028	4.85	4.19	4.10
2029	4.79	4.02	4.10
2030	4.72	4.01	4.10

Germany

Germany shows steady growth from 2010 to 2019, increasing GDP from about 3.47 trillion dollars to 3.96 trillion dollars. Minor fluctuations reflect trade cycles and external shocks. COVID-19 caused a small dip in 2020 with recovery reaching roughly 4.66 trillion dollars by 2024.

Figure 6-4 Germany GDP over time

ARIMA with order 3,2,0 produces the most accurate forecasts with a MAPE of 2.47% and R2 of 0.756. Theta reaches a MAPE of 8.29% and 2 of -1.96, and BSTS has a MAPE of 11.62% and R2 of -4.08, reflecting wider uncertainty intervals. Forecasts suggest moderate and steady GDP growth through 2030 supported by Germany's strong industrial base and export-oriented economy.

Table 6-4 Germany Forecasted GDP

Year	Theta Model (Trillion USD)	ARIMA (Trillion USD)	BSTS (Trillion USD)
2025	3.92	4.74	4.60
2026	3.94	4.91	4.60
2027	3.96	5.10	4.60
2028	3.97	5.21	4.60
2029	3.99	5.36	4.60
2030	4.01	5.53	4.60

United Kingdom

The United Kingdom shows moderate GDP growth from 2.49 trillion dollars in 2010 to 2.85 trillion dollars in 2019. The economy contracts in 2020 due to COVID-19 but recovers to about 3.64 trillion dollars by 2024. This pattern reflects the impact of global shocks and the slower recovery of a mature service-oriented economy.

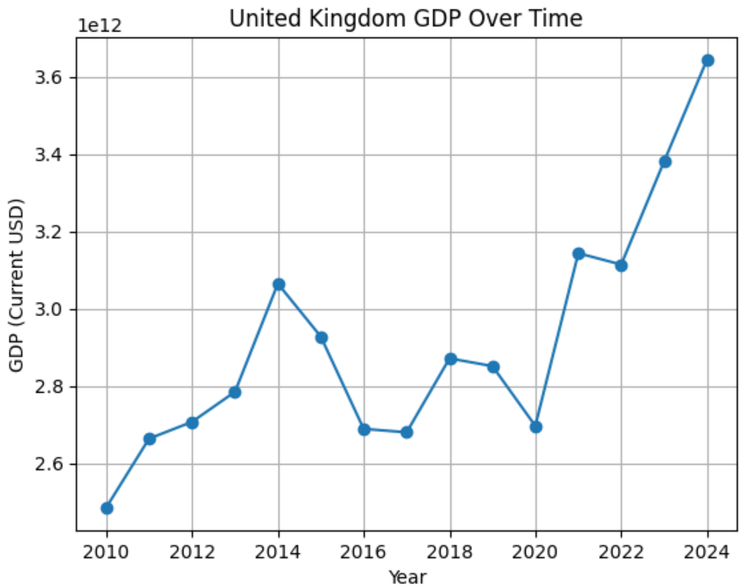

Figure 6-5 United Kingdom GDP Over time

ARIMA with order 3,2,0 performs best with a MAPE of 2.47% and R2 of 0.756, accurately tracking historical trends. Theta has a MAPE of 11.21% and R2 of -0.84, providing a stable baseline but failing to capture sudden deviations. BSTS shows higher uncertainty with a MAPE of 13.26% and R2 of -1.69. Forecasts indicate gradual GDP growth through 2030 consistent with the moderate expansion of high-income economies.

Table 6-5 United Kingdom Forecasted GDP

Year	Theta Model (Trillion USD)	ARIMA (Trillion USD)	BSTS (Trillion USD)
2025	2.86	3.84	3.53
2026	2.88	4.00	3.53
2027	2.89	4.25	3.53
2028	2.90	4.47	3.53
2029	2.91	4.66	3.53
2030	2.93	4.87	3.53

6.2 Comparative Insights Across Models and Countries

Across all five countries, ARIMA gives the most accurate point forecasts with low MAPE and high R2 values. Theta models are simpler and provide a baseline while highlighting periods of forecast instability, especially during sharp shocks like COVID-19. BSTS adds value by showing uncertainty ranges which are useful for policymakers evaluating possible risks and volatility.

Table 6-6 Models Performances Across the countries

Country	Model	MAPE (%)	R2
United States	Theta	11.7913	-0.8118
	ARIMA	2.3533	0.9385
	BSTS	15.1292	-2.0971
Japan	Theta	10.1820	-0.4917
	ARIMA	13.1293	1.4667
	BSTS	13.1315	-1.5455
United Kingdom	Theta	11.2119	-0.8442
	ARIMA	14.4025	-2.4452
	BSTS	13.2598	-1.6966
China	Theta	10.0803	-1.1564
	ARIMA	2.9708	0.6666
	BSTS	17.3164	-4.9840
Germany	Theta	8.2945	-1.9572
	ARIMA	2.4691	0.7557
	BSTS	11.6216	-4.0773

Time trends show that major economies with large industrial bases such as the United States, China, and Germany dominate global output. Mature, high-income economies like Japan and the United Kingdom grow more slowly. The 2020 pandemic acted as a global shock, showing why forecasting models must account for sudden external events.

Forecasts up to 2030 indicate continued growth in all five countries at different speeds. The United States and China show strong upward trends, Germany and the UK moderate growth, and Japan slow but steady increases. These predictions follow historical patterns and reflect structural economic factors and policy influences confirmed by international sources such as the IMF and World Bank.

6.3 Discussion

6.3.1 Alignment of the Results with Problem Statement and Objectives

The results of this project directly address the problem of limited combined analysis and forecasting in regional macroeconomic studies. By examining GDP per capita and GDP growth for more than two hundred countries between 2010 and 2024 the project offers a clear and structured comparison of economic performance across regions. At the same time, it produces reliable forecasts for major economies. This combined approach helps to close a common gap in existing research where many studies focus only on past trends while others generate forecasts without placing them in a broader regional context.

The exploratory data analysis revealed clear patterns over time and across countries that match well with economic theory and major global events. For instance, the strong skew in GDP measured in current USD shows that most global output is produced by a small number of very large economies. This agrees with findings reported by the World Bank and the IMF. When looking at GDP per capita the analysis showed a clear difference between the size of an economy and the standard of living of its citizens. Several small high-income countries perform better on a per person basis than much larger industrial economies. GDP growth trends also reflected periods of instability especially the sharp drop in 2020 caused by the COVID 19 pandemic followed by a recovery in later years. Together these findings meet the goal of examining past trends and relationships between key economic indicators and they provide a strong basis for understanding regional differences.

Country level analysis further strengthened the comparative view. High income microstates such as Monaco and Luxembourg show steady growth alongside very high GDP per capita which reflects mature economies with limited room for rapid expansion. In contrast, fast growing economies such as Guyana and Djibouti display much higher volatility. This is often linked to external shocks to natural resources or major policy changes. These patterns confirm that using consistent indicators makes it possible to compare countries effectively and they support the objective of assessing regional economic performance using standardized measures.

The forecasting models build on the earlier analysis by looking ahead and meeting the project's prediction goals. Among the models ARIMA performed best across most countries because it captured trends and past patterns very well. This was especially clear in large industrial economies such as the United States and Germany. The Theta model worked well as a simple and reliable baseline mainly in more stable economies. BSTS was useful for showing uncertainty around forecasts but it was less accurate when the historical data were short or when the economy experienced high volatility. The performance measures such as MAPE and R squared show that ARIMA can provide dependable future projections. This makes it useful for policymakers and investors who need support for economic planning.

By forecasting GDP for the United States China Japan Germany and the United Kingdom up to 2030 the project meets its goal of developing and testing ARIMA and Bayesian Structural Time Series models. The forecasts reflect expected paths based on current trends and they help identify possible periods of growth or slowdown. For example, the United States shows steady growth following the COVID 19 shock in 2020. China's forecast points to moderate but uneven growth which reflects ongoing structural changes. Japan displays slow but stable increases that are typical of a mature economy. These results offer practical insights that support better economic decision making.

The comparison of model performance across countries also shows why forecasting methods should be chosen with care. ARIMA performs very well in industrial economies while BSTS tends to perform worse in this setting. This highlights the importance of selecting models that match the economic context. Such context aware model choice is often missing in earlier studies. Overall, the project succeeds in combining comparative analysis with reliable forecasting and presents a practical data driven framework for assessing regional economic performance.

6.3.2 Comparison of the insights with Existing Studies

The forecasting results from this project are in line with earlier studies and also add new insights by applying ARIMA BSTS and Theta models to individual countries. Across the five countries studied ARIMA gave the best results with the lowest MAPE and the highest R squared values. This confirms that ARIMA is a reliable model for large and stable economies. These findings agree with Cao 2025 who showed that ARIMA performed very well for Germany and the United Kingdom with very low error levels. They also support Liu 2025 who found that ARIMA is suitable for capturing GDP trends in China over a medium-term period.

BSTS produced forecasts that included uncertainty ranges which is one of its main strengths. However its accuracy measured by MAPE and R squared was lower than ARIMA for several countries. This matches the findings of Bheemanna and Megeri 2024 who showed that BSTS works better in emerging economies but struggles when the data history is short or when the economy is very unstable. In this project BSTS was still useful because it provided confidence intervals for GDP forecasts especially for China and the United States. This shows that BSTS is more useful for understanding possible scenarios than for making exact point predictions.

The Theta model acted as a strong baseline and delivered reasonable results across all countries. This supports earlier research which shows that simple time series methods can still perform well when economic trends are smooth and stable as discussed by Buturac 2022. In countries like the United States and Germany where recovery after the pandemic has been steady, Theta captured the general trend well although ARIMA remained more accurate overall.

When comparing these results with studies that use multiple models such as Hassan and Mirza 2021 the value of combining different forecasting methods becomes clear. Similar to their work this project shows that ARIMA is strong at modeling trends, BSTS helps explain uncertainty and Theta provides a stable reference point. Using these models together gives a better picture of future economic paths and supports the project goal of providing useful insights for policymakers.

Chapter 7 BUSINESS INSIGHTS AND RECOMMENDATIONS

The results from the exploratory analysis and the forecasting models clearly show that economic behavior differs across the five countries studied. These findings are useful for policymakers, investors and international organizations involved in economic planning. By turning the analysis into practical insights decision makers can better match their strategies with expected economic trends.

7.1 Insights from Past Trends and Future Forecasts

Looking at past data shows clear differences in economic size and growth. The United States and China lead global GDP because of their large industries, strong technology sectors and deep involvement in global trade. Japan, Germany and the United Kingdom have smaller total GDP values but show stable economies with steady growth. High GDP per capita in countries like Germany and the United Kingdom points to strong productivity and high living standards. In contrast China shows strong growth mainly due to ongoing development and industrial expansion.

Forecasting adds more understanding to these patterns. ARIMA models performed better than BSTS and Theta when making direct predictions. This suggests that models based on clear trends work well for large and stable economies. For example, forecasts for the United States and Germany show continued growth up to 2030 which reflects recovery after the pandemic and strong industrial capacity. BSTS models were less precise in exact predictions but they provided useful uncertainty ranges. These ranges help policymakers understand how outcomes may change especially in economies like China where policy shifts and structural changes are still happening. Theta models produced simple and stable forecasts which makes them useful for basic planning scenarios.

Comparing total GDP with GDP per capita highlights the need for careful decision making. While the United States ranks highest in total GDP GDP per capita better reflects individual economic well-being. This measure is important for planning jobs, social policies and consumer demand. In the same way countries with fast GDP growth may not always show equal improvements in living standards across the population.

7.2 Implications for Policymakers

Policymakers can use these findings to guide economic planning. The projected growth patterns support continuing current policies in stable economies while also showing where more focused actions may be needed. For example, the wider uncertainty in China's forecasts suggests that flexible fiscal policies are important along with close monitoring of investment and industrial activity to reduce risks from sudden changes.

The results also show the value of using multiple economic indicators together. Looking at GDP GDP per capita and growth rates at the same time helps identify areas that need social spending, infrastructure development or economic diversification. This approach helps ensure that growth improves overall living conditions and not just total output.

7.3 Implications for Investors and International Organizations

For investors these forecasts help with deciding where to invest and how to manage risk. Countries with steady growth and low uncertainty such as Germany and the United States are attractive for long term investments in areas like infrastructure technology and finance. Countries with more volatile growth such as China may offer higher returns but require careful planning and risk control. BSTS forecasts are especially useful here because they show both positive and negative scenarios.

International organizations can also benefit from these results. The findings can support decisions related to development, support trade policy and economic advice. Countries like Germany and Japan with strong productivity and stable growth can serve as reference points for best practices. Countries with fast growth but lower income levels can be supported through programs that help turn economic expansion into broader social benefits.

7.4 Recommendations for Stakeholders

Use more than one forecasting model for better decisions

Using ARIMA BSTS and Theta together gives a clearer picture of the future. ARIMA works best for economies that are stable and follow clear trends. BSTS is useful when planning for different possible outcomes especially in uncertain or fast changing economies. Theta provides a simple baseline that helps confirm overall trends.

Focus on turning economic growth into better living standards

A high GDP alone does not mean people are better off. Policies should aim to increase income per person, create more jobs and improve productivity so that economic growth benefits the wider population.

Use forecasts when making investment decisions

Investors should look at both expected growth and the level of uncertainty. Countries with stable ARIMA forecasts are better suited for long term investments. BSTS uncertainty ranges help investors manage risk in countries where economic conditions are less predictable.

Prepare for unexpected economic shocks

The forecasting results show that major events like the COVID pandemic can strongly affect economic outcomes. Governments, businesses and investors should plan ahead and keep strategies flexible so they can respond quickly to future shocks.

Keep track of structural and policy changes

Regular monitoring of key economic indicators is important. Changes in government policy, industry direction or global trade rules can shift economic paths especially in countries with fast changing economies such as China.

Use data driven insights for global comparisons

Using consistent economic indicators makes it easier to compare regions, identify strengths and weaknesses and support international cooperation and coordinated policy actions.

CONCLUSION

This project aimed to solve the problem of not having one clear data driven way to compare regional economic performance while also producing reliable forecasts for important economic indicators. The study used a large dataset covering 200 countries from 2010 to 2025 and focused on three main indicators which are GDP in current US dollars GDP per capita in current US dollars and annual GDP growth. These indicators were chosen because they show both the size of an economy and the level of individual prosperity which gives a balanced view of economic performance.

The exploratory data analysis showed clear differences between countries and regions. High income microstates such as Monaco, Liechtenstein and Luxembourg recorded the highest GDP per capita which reflects strong productivity and a dominant financial sector. Large economies like the United States China Japan Germany and the United Kingdom led in total GDP while their growth patterns differed based on economic maturity policy choices and exposure to external shocks. Fast growing countries such as Guyana, Djibouti , Tajikistan , Turkmenistan and Mongolia recorded high growth rates driven by natural resources foreign investment or economic reforms although this growth was often unstable. Overall, the results confirmed that most global economic output is concentrated in a small number of countries while many others produce much smaller outputs.

For forecasting, the study applied Theta ARIMA and Bayesian Structural Time Series models to five major economies which are the United States China Japan Germany and the United Kingdom. The models were trained using data from 2010 to 2019 and tested on data from 2020 to 2024. Model tuning was carried out for ARIMA and BSTS to improve performance for each country. ARIMA consistently delivered the most accurate point forecasts especially for stable economies with clear trends. BSTS provided useful uncertainty ranges which helped with planning under uncertainty particularly for countries like China that experience structural changes. Theta models offered simple and stable forecasts This study shows that combining regional comparisons with strong forecasting methods can produce useful insights for policymakers, investors and international organizations. Looking at both past trends and future forecasts helps guide government policies, investment choices and development planning. Probabilistic forecasts also help decision makers understand risk and uncertainty. This approach fills the gap identified in the problem statement by using one clear method to study past economic performance and to predict future trends.

However, the study also has some limitations. The forecasts are based on past data which means they may not fully capture unexpected events or major structural changes after 2025. The analysis focuses on only three economic indicators and adding others such as inflation, unemployment or government debt could provide deeper insights. Future studies could also include longer time periods or different forecasting methods. Advanced techniques such as machine learning models could help capture complex patterns and improve accuracy.

In summary the project highlights the importance of a unified data driven framework for analyzing and forecasting regional economic performance. The results offer practical insights and clear guidance that support better decision making in a complex and changing global economy.

REFERENCES

Adewale, M.D., Ebem, D.U., Awodele, O., Sambo-Magaji, A., Aggrey, E.M., Okechalu, E.A., Donatus, R.E., Olayanju, K.A., Owolabi, A.F., Oju, J.U., Ubadike, O.C., Otu, G.A., Muhammed, U.I., Danjuma, O.R. and Oluyide, O.P., 2024. Predicting gross domestic product using the ensemble machine learning method. *Systems and Soft Computing*, 6, p.200132. https://doi.org/10.1016/j.sasc.2024.200132

Andrianady, J.R. (2023) Crunching the Numbers: A Comparison of Econometric Models for GDP Forecasting in Madagascar. MPRA Paper 116916, University Library of Munich, Germany. https://ideas.repec.org/p/pra/mprapa/116916.html

Atif, D., 2025. Enhancing long-term GDP forecasting with advanced hybrid models: a comparative study of ARIMA-LSTM and ARIMA-TCN with dense regression. *Computational Economics*, 65, pp.3447–3473. https://doi.org/10.1007/s10614-024-10683-5

Bationo, F. d'A., Murinde, V. and Soumaré, I., 2025. Cross-border banking and the transmission of global shocks to credit cycles in developing economies: a commodity price cycles channel. International Review of Financial Analysis, 106, 104515. https://doi.org/10.1016/j.irfa.2025.104515

Bäurle, G., Steiner, E. and Züllig, G., 2020. Forecasting the production side of GDP. *Research Article*. https://doi.org/10.1002/for.2725

Bheemanna, & Megeri, M.N., 2024. Forecasting of population and economic variables in India using the Bayesian Structural Time Series (BSTS) Model. *Maths*, 9(6), pp.205–210. https://doi.org/10.22271/maths.2024.v9.i6c.1922

Buturac, G., 2022. Measurement of economic forecast accuracy: A systematic overview of the empirical literature. Journal of Risk and Financial Management, 15(1), 1. https://doi.org/10.3390/jrfm15010001

Cao, D., 2025. Forecasting the real GDP of United Kingdom, Germany and France in next three years. *Theoretical and Natural Science*, 109, pp.37–41.https://doi.org/10.54254/2753-8818/2025.GL23169

Čapek, J., Crespo Cuaresma, J., Hauzenberger, N. and Reichel, V., 2023. Macroeconomic forecasting in the euro area using predictive combinations of DSGE models. International Journal of Forecasting, 39(4), pp.1820-1838. https://doi.org/10.1016/j.ijforecast.2022.09.002

Dou, W.W., Lo, A.W., Muley, A. and Uhlig, H., 2020. Macroeconomic Models for Monetary Policy: A Critical Review from a Finance Perspective. Annual Review of Financial Economics, 12, pp.95-140. https://doi.org/10.1146/annurev-financial-012820-025928

Eldomiaty, T., Azzam, I., Fouad, M. and Said, Y., 2024. The use of economic indicators as early signals of stock market progress: Perspectives from Market Potential Index. International Journal of Financial Studies, 12(1), 21. https://doi.org/10.3390/ijfs12010021

Galí, J., 2020. The effects of a money-financed fiscal stimulus. CREI, Universitat Pompeu Fabra and Barcelona GSE, Spain. https://www.crei.cat/wp-content/uploads/2020/10/gali_jme2020-1.pdf

Hassan, M.M. and Mirza, T., 2021. Using time series forecasting for analysis of GDP growth in India. I. J. Education and Management Engineering, 3, pp.40-49. https://doi.org/10.5815/ijeme.2021.03.05

Jonathan, N.D. (2024) Enhancing GDP Forecast Accuracy: Comparing Neural Hierarchical Interpolation for Time Series (N-HiTS) Model and Vector Autoregression (VAR) Model for U.S. Economic Planning

Jordà, Ò. and Nechio, F., 2023. Inflation and wage growth since the pandemic. European Economic Review, 156, 104474. https://doi.org/10.1016/j.euroecorev.2023.104474

Kamolthip, S., 2021. Macroeconomic forecasting with LSTM and mixed frequency time series data. *PIER Discussion Papers*, 165, Puey Ungphakorn Institute for Economic Research. https://ideas.repec.org/p/pui/dpaper/165.html

Kohler, K. and Stockhammer, E., 2021. Growing differently? Financial cycles, austerity, and competitiveness in growth models since the Global Financial Crisis. Review of Political Economy, 33(6), pp.1314-1341. https://doi.org/10.1080/09692290.2021.1899035

Liu, S., 2025. Predicting China's GDP in future five years by linear regression and ARIMA model. *Theoretical and Natural Science*, 101, pp.184–192. https://doi.org/10.54254/2753-8818/2025.CH22962

McCloskey, P.J. and Malheiros Remor, R., 2024. Comparative analysis of ARIMA, VAR, and linear regression models for UAE GDP forecasting. *MPRA Paper No. 122860*, University of London. https://mpra.ub.uni-muenchen.de/122860/

Moramarco, G., 2023. Measuring global macroeconomic uncertainty and cross-country uncertainty spillovers. Econometrics, 11(1), 2. https://doi.org/10.3390/econometrics11010002

Papetti, A., 2021. Demographics and the natural real interest rate: historical and projected paths for the euro area. Journal of Economic Dynamics and Control, 132, 104209. https://doi.org/10.1016/j.jedc.2021.104209

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M.Z., Barrow, D.K., Ben Taieb, S., Bergmeir, C., Bessa, R.J., Bijak, J., Boylan, J.E., Browell, J., Carnevale, C., Castle, J.L., Cirillo, P., Clements, M.P., Cordeiro, C., Oliveira, F.L.C., De Baets, S., Dokumentov, A., Ellison, J. and Ziel, F., 2022. Forecasting: theory and practice. *International Journal of Forecasting*, 38(3), pp.705–871. https://doi.org/10.1016/j.ijforecast.2021.11.001

Shon, H., Lee, H. and Kim, B., 2024. Spatial pattern of aid allocation in the early 21st century: Evidence from 38 sub-Saharan African countries. Papers in Regional Science, 103(3), 100026. https://doi.org/10.1016/j.pirs.2024.100026

Tetteh, B. and Ntsiful, E., 2023. A comparative analysis of the performances of macroeconomic indicators during the Global Financial Crisis, COVID-19 Pandemic, and the Russia-Ukraine War: The Ghanaian case. Research in Globalization, 7, 100174. https://doi.org/10.1016/j.resglo.2023.100174

Ullah, S., Ullah, A. and Zaman, M., 2024. Nexus of governance, macroeconomic conditions, and financial stability of banks: a comparison of developed and emerging countries. Financial Innovation, 10, 30. https://doi.org/10.1186/s40854-023-00542-x

APPENDIX

https://github.com/richarddagba/regional-economic-performance-analysis