A Project Report

On

Gross Domestic Product (GDP) Analysis

Submitted in partial fulfillment of the

requirement for the award of the degree of

MASTER OF COMPUTER APPLICATION

GALGOTIAS UNIVERSITY

(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

DEGREE

Session 2024-25
in

Data Science and AI

By

Rishav Kumar (23SCSE2160025)
Roshan Kumar Shakya (23SCSE2160006)
Prashant Ranjan (23SCSE2160058)

Under the guidance of
Dr. Ajeet Singh (Professor)

SCHOOL OF COMPUTER APPLICATIONS AND TECHNOLOGY

GALGOTIAS UNIVERSITY, GREATER NOIDA

INDIA

May, 2025

SCHOOL OF COMPUTER APPLICATIONS AND TECHNOLOGY

GALGOTIAS UNIVERSITY, GREATER NOIDA

CANDIDATE'S DECLARATION

I/We hereby certify that the work which is being presented in the project, entitled **"Gross Domestic Product (GDP) Analysis"** in partial fulfillment of the requirements for the award of the **MCA (Master of Computer Application)** submitted in the School of Computer Applications and Technology of Galgotias University, Greater Noida, is an original work carried out during the period of August, 2024 to May, 2025, under the supervision of Dr. Ajeet Singh, Department of Computer Science and Engineering/School of Computer Applications and Technology , Galgotias University, Greater Noida. The matter presented in the thesis/project/dissertation has not been submitted by me/us for the award of any other degree of this or any other places.

Rishav Kumar (23SCSE2160025)
Roshan Kumar Shakya(23SCSE2160006)
Prashant Ranjan(23SCSE2160058)

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Dr. Ajeet Singh

Professor

SCHOOL OF COMPUTER APPLICATION AND TECHNOLOGY

GALGOTIAS UNIVERSITY, GREATER NOIDA

CERTIFICATE

This is to certify that Project Report entitled "**Gross Domestic Product (GDP) Analysis**" which is submitted by **Rishav Kumar, Roshan Kumar Shakya, Prashant Ranjan** in partial fulfillment of the requirement for the award of degree MCA in Data Analytics Department of School of Computer Applications and Technology, Galgotias University, Greater Noida, India is a record of the candidate own work carried out by him/them under my supervision. The matter embodied in this thesis is original and has not been submitted for the award of any other degree.

Signature of Examiner(s) **Signature of Supervisor(s)**

Date: May, 2025

Place: Greater Noida

TABLE OF CONTENTS

ACKNOWLEDGEMENT

We take this opportunity to express our sincere gratitude to **Galgotias University**, for providing us with the opportunity and platform to pursue our academic interests and successfully complete our final-year project titled *"Gross Domestic Product (GDP) Analysis."*

We express our heartfelt thanks to our esteemed guide, **Dr. Ajeet Singh**, Professor, *School of Computer Science and Engineering*, Galgotias University, Greater Noida, for his invaluable guidance, support, and encouragement throughout the project. His timely insights and suggestions played a pivotal role in shaping our understanding and execution of this work.

We are also thankful to the faculty members of the *School of Computer Applications and Technology* for their direct and indirect support, as well as for creating an environment conducive to learning and exploration.

Lastly, we extend our deepest gratitude to our parents and friends for their unwavering support, motivation, and constant encouragement throughout this journey.

ABSTRACT

The project titled **“Gross Domestic Product (GDP) Analysis”** aims to explore and analyze India’s economic growth using historical GDP data and predictive modeling. By examining the contributions of key sectors—agriculture, industry, and services—this study identifies trends, patterns, and critical factors influencing the country's economic performance. A web-based application was developed to facilitate interactive data visualization, sector-wise comparisons, and future forecasting using machine learning algorithms such as ARIMA, Random Forest, and Neural Networks. The system integrates external data sources, processes large datasets, and provides users with insightful visual dashboards and downloadable reports. Through this project, we demonstrate how data science and artificial intelligence can be effectively applied to understand complex economic indicators and support decision-making in policy and planning.

LIST OF DIAGRAMS

List of Graphs Used

Graph Name	Purpose / Description
Line Plot	Displays trends over time or continuous data points
Bar Plot	Compares categorical data using rectangular bars
Histogram	Shows the distribution of a single variable
Box Plot	Visualizes the distribution, median, quartiles, and outliers of a dataset
Scatter Plot	Displays the relationship between two continuous variables
Pairplot	Plots pairwise relationships across a dataset (scatter + histogram/ KDE)
Heatmap	Displays correlation or matrix data with color gradients
Violin Plot	Combines box plot and KDE to show distribution and probability density
Pie Chart	Shows proportions of a whole using slices
Area Plot	Similar to line plot but with filled area beneath lines for stacked values
Count Plot	Displays count of occurrences for each categorical level
Dist Plot	Shows distribution of a univariate set (histogram + KDE)
Lag Plot	Used to check randomness in time series data
Autocorrelation Plot	Displays correlation of a time series with lagged versions of itself

Graph Name	Purpose / Description
Time Series Plot	Shows data points indexed in time order, useful for trends and seasonality
Residual Plot	Plots residuals to evaluate the fit of a regression model
Regression Plot	Displays linear regression fit along with scatter plot

LIST OF ABBREVIATIONS

ARIMA	AutoRegressive Integrated Moving Average
AR	Autoregressive
MA	Moving Average
ARMA	AutoRegressive Moving Average
ADF	Augmented Dickey-Fuller (stationarity test)
ACF	Autocorrelation Function
PACF	Partial Autocorrelation Function
AIC	Akaike Information Criterion
LR	Linear Regression
RF	Random Forest
DT	Decision Tree
SVM	Support Vector Machine
SVR	Support Vector Regression
NN	Neural Network
MAE	Mean Absolute Error
MSE	Mean Squared Error
RMSE	Root Mean Squared Error
FDI	Foreign Direct Investment
np	NumPy (used for numerical operations)
pd	pandas (used for data manipulation and analysis)
plt	matplotlib.pyplot (used for plotting graphs)
sns	seaborn (used for statistical data visualization)

LIST OF SYMBOLS

S No.	Symbols	Description
1.	$	Declare a variable
2.	;	Semicolon
3.	()	Parentheses
4.	{}	Braces
5.	//	Double forward slashes
6.	/…../	A block of code enclosed in /* and */
7.	==	Double equal
8.	===	Triple equal
9.	!=	Exclamation and equal sign
10.	<>	Diamond
11.	<=	Less than or equals
12.	/	Forward slash
13.	%	Percent
14.	[]	Square bracket
15.	->	Arrow symbol
16	::	Double colon

CHAPTER 1
INTRODUCTION

In the modern era, where a country's economic growth is crucial, understanding and analyzing Gross Domestic Product (GDP) data is very important. This project looks at India's GDP, focusing on the contributions of different sectors like agriculture, industry, and services over time. The goal is to help users better understand the country's economic progress through simple and clear visualizations and analysis. By looking at trends in GDP growth, this project highlights the main factors that drive the economy and points out the challenges and opportunities for future development.

1.1 Overview of GDP

Gross Domestic Product (GDP) serves as a comprehensive measure of a country's overall economic health and performance. It reflects the total monetary value of all finished goods and services produced within a country’s borders over a specific period, usually a year. For economists, policymakers, and analysts, GDP is a critical tool for understanding the structure and health of an economy, as it aggregates data across various sectors like agriculture, industry, and services, offering insights into the nation's economic output and growth trends.

1.2 Importance of GDP in Economic Analysis

GDP is widely considered a key indicator of a nation's economic performance, directly linked to the standard of living, employment rates, and government policies. A rising GDP indicates economic expansion and prosperity, while a shrinking GDP may signal economic troubles. GDP analysis is essential for identifying growth patterns, informing fiscal and monetary policies, and fostering sustainable economic development. For developing economies like India, GDP serves as a benchmark for measuring progress and competitiveness on the global stage.

1.3 India's Economic Context

India, being one of the world's fastest-growing economies, presents a unique case for GDP analysis. With a vast and diverse economy, India’s GDP reflects contributions from various sectors, including agriculture, industry, and a rapidly expanding services sector. Over the past few decades, India has experienced dynamic shifts in its GDP composition, moving from an agriculture-dominated economy to one that is increasingly reliant on services and manufacturing.

1.4 Key Drivers of India's GDP Growth

Several factors contribute to India's GDP growth, including:

- **Agriculture:** Despite a declining share in total GDP, agriculture remains a crucial sector, providing livelihoods to a significant portion of the population.
- **Industrial Development:** The growth of manufacturing, construction, and mining activities contributes significantly to India’s industrial output.
- **Services Sector:** The largest contributor to India's GDP, the services sector, including IT, finance, telecommunications, and healthcare, has become the key driver of economic growth.
- **Global Trade and Foreign Investments:** India's increasing integration with global markets, foreign direct investment (FDI), and export growth have also contributed to its GDP expansion.

1.5 Objectives of the Study

This study on India’s GDP aims to:

1. Analyze the historical trends and patterns in India's GDP growth.
2. Examine the sectoral contributions to GDP over the years.
3. Identify the challenges and opportunities influencing India's economic performance.
4. Investigate the impact of factors like government policies, corruption, and global trade on India’s GDP.

CHAPTER 2

(Web Based Projects)

Software Requirement Specification for GDP Analysis System.

This chapter provides an in-depth look at the GDP Analysis System, a web-based application aimed at analyzing and visualizing economic data, with a particular focus on Gross Domestic Product (GDP). This section introduces the system's purpose, interfaces, functionalities, and operational context. While the detailed technical aspects are addressed in subsequent sections, this overview sets the foundation for understanding the product's capabilities and requirements.

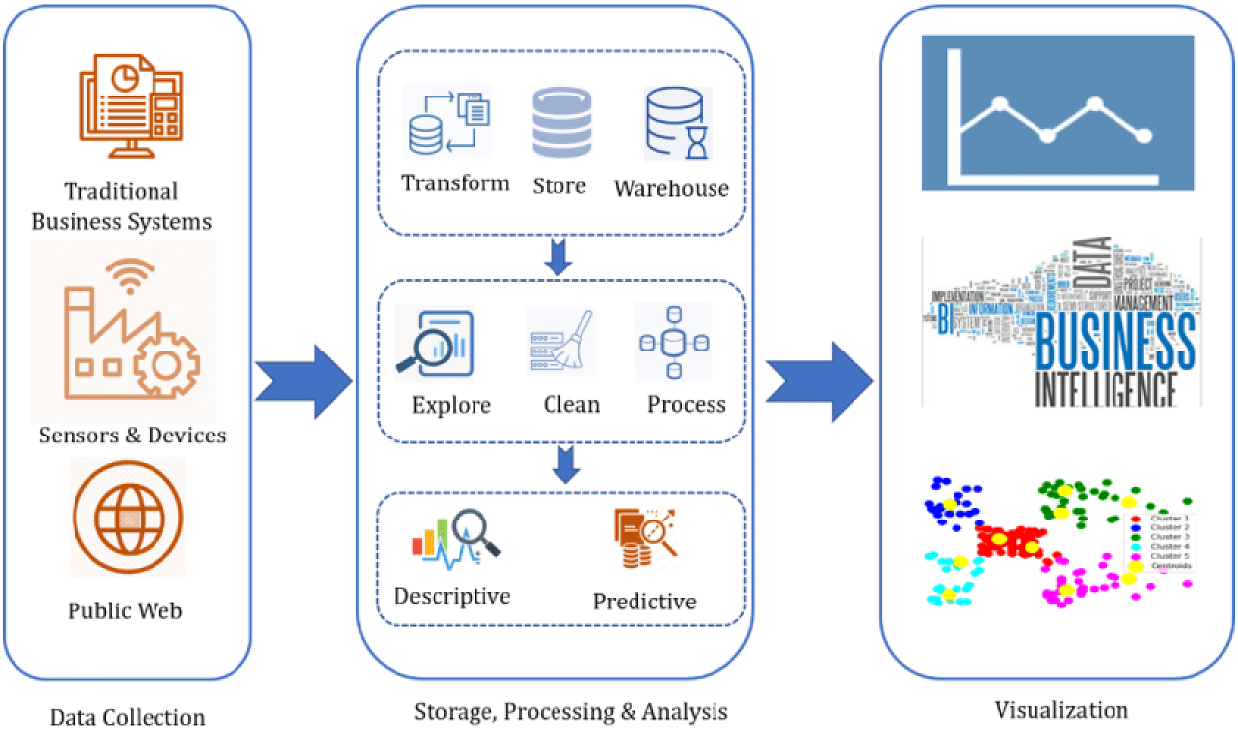
2.1 Product Perspective

The GDP Analysis system is a self-contained web-based application, designed to analyze and visualize the Gross Domestic Product (GDP) of different countries and regions, allowing users to examine historical trends, perform predictive analysis, and explore the impact of various economic factors.

The system will operate independently but can be integrated with external data sources, such as financial APIs, government databases, or CSV/Excel files. It will offer dynamic visualization and data-driven insights to users from various sectors, including academia, businesses, and policymakers.

Key objectives of the system:

- Facilitate the aggregation of GDP data from multiple sources.
- Enable real-time, interactive visual analysis.
- Allow forecasting and trend evaluation through machine learning models.
- Support customized reporting and export functionality for further dissemination.



2.1.1 System Interfaces

The system will interface with:

- External financial APIs (e.g., World Bank API, IMF datasets).
- File-based imports (CSV, Excel).
- Visualization libraries (e.g., D3.js, Chart.js, Plotly).
- Backend systems like a relational database (MySQL/PostgreSQL) for data persistence.

2.1.2 Interfaces

The application will feature a clean, modern, and responsive GUI built using technologies like HTML5, CSS3, JavaScript, and frontend frameworks such as React or Vue.js.

GUI elements include:

- A login and registration module with form validation.
- Dashboards showing visual insights like GDP trends and comparisons.
- File upload area for importing data.
- Interactive charts and graphs with zoom, filter, and tooltip features.

- A reporting panel with download/export options (CSV, PDF, image).

2.1.3 Hardware Interfaces

The GDP Analysis system has no direct hardware interface requirements. It is web-based, thus independent of specific hardware, running on any device capable of browsing the internet.

2.1.4 Software Interfaces

The system will interact with the following software components:

1. World Bank API
   - Purpose: To fetch real-time economic data for various countries.
   - Communication: RESTful API calls in JSON format.
2. D3.js
   - Purpose: Data visualization and charting in the web application.
   - Communication: JavaScript-based rendering.
3. MySQL Database
   - Purpose: Storing processed GDP data and user information.
   - Communication: SQL queries via Python backend.

2.1.5 Communication Interfaces

- **Data Transmission:** HTTPS with SSL/TLS encryption for all transactions.
- **Data Formats:** JSON for API communication, CSV/Excel for imports, and PDF for report exports.
- **Authentication:** OAuth 2.0 / JWT for secure access to user accounts and API interactions.

2.1.6 Memory Constraints

- **Client-Side:** Efficient memory management to ensure responsiveness on systems with at least 4GB RAM.
- **Server-Side:** Optimized data caching and background processing, with recommended minimum 8GB RAM and SSD storage.

2.1.7 Operations

- **Primary Mode:** Web-based interactive analysis via browser.
- **Scheduled Operations:** Background synchronization with data APIs, machine learning model retraining, and periodic backups during low-usage hours.

2.1.8 Site Adaptation Requirements

The system must be adaptable to local economic contexts:

- Custom indicators per region or country.
- Multilingual support (optional in future releases).
- Time zone and date format localization.

2.2 Product Functions

The system provides a comprehensive suite of features including but not limited to:

1. **Data Acquisition**
   - API-based or manual import of GDP data.
   - Automatic data validation and cleaning.
2. **Analytical Computation**
   - GDP growth rate, per capita analysis, inflation adjustment.
   - Economic indicator correlations.
3. **Visualization**
   - Interactive charts (line, bar, pie, heatmap).
   - Region-based choropleth maps for comparative analysis.
4. **Forecasting & AI Insights**
   - Time-series forecasting using ARIMA, Prophet, or LSTM.
   - Anomaly detection for unexpected economic shifts.
5. **User Management**
   - Role-based access control (Admin, Analyst, Viewer).
   - Personalized dashboards and saved sessions.
6. **Exporting and Reporting**
   - Download options: CSV, XLSX, PDF, PNG.
   - Auto-generated reports with summary insights and graphs.

2.3 User Characteristics

Target users include:

- **Economists & Policy Analysts**: Require deep data insights and forecasting tools.
- **Students & Researchers**: Need intuitive data exploration features for academic use.

- **Business Analysts**: Use the system for market and macroeconomic analysis.

Users are expected to have:

- Basic knowledge of GDP and economic terms.
- Moderate proficiency in using web tools or Excel.
- Interest in understanding and predicting economic trends.

2.4 Constraints

- Regulatory Policies: Must comply with international data privacy laws (e.g., GDPR) regarding the handling of sensitive economic data.
- Interface with External Applications: The system will interact with APIs that may have specific rate limits or authentication methods.
- Reliability: High reliability and accuracy are required due to the importance of economic data.
- Security Considerations: Data encryption (SSL/TLS) is essential for secure communication with external data sources.

2.5 Assumptions and Dependencies

- The system assumes that the World Bank API and other external data sources will remain operational and accessible.
- It is assumed that users have access to the internet and a modern web browser.
- The accuracy of the system's analysis depends on the quality of input data provided by users.

2.6 Apportioning of Requirements

To manage complexity and development timelines, some features will be phased:

Immediate Implementation:

- Core functionalities (data upload, visualization, basic analysis).
- API integration with the World Bank.
- Report generation and basic user management.

Deferred to Future Releases:

- AI-driven predictive analytics and alerts.

- Multi-language support.
- Real-time collaboration tools.
- Integration with cloud storage (e.g., Google Drive, Dropbox).

2.7 Use Case

2.7.1 Use Case Model Guidelines

- **Actors:** End Users (Economists, Analysts), Admins.
- **Goals:** Analyze GDP data, generate visual insights, export results.
- **Workflow:**
  1. User logs into the system.
  2. Uploads or imports data.
  3. System processes and visualizes the data.
  4. User applies filters, views predictions.
  5. Downloads reports or exports visualizations.

2.7.2 Use Case Diagram

Use Case Number	001
Application	GDP Analysis System
Use Case Name	Upload GDP Data
Description	Users upload GDP data in CSV or Excel format for analysis.
Primary Actor	User (Economist/Data Analyst)
Preconditions	The user must be logged into the system.
Trigger	The user clicks the "Upload" button.
Basic Flow	User selects a file → System validates file format → Data is processed.
Alternate Flows	Invalid file format → System shows an error message.

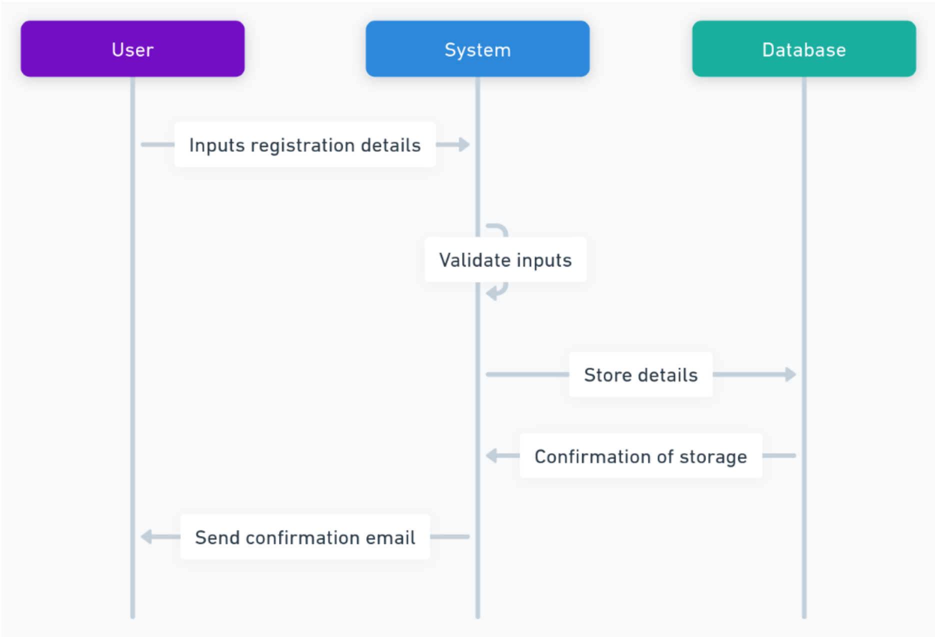
2.8 Sequence Diagram

A sequence diagram for the registration process can show:

Steps:

1. User fills out and submits the registration form.
2. Frontend sends details to the backend API.

3. Backend validates and saves user data in the database.
4. Backend triggers email server to send a confirmation email.
5. User receives confirmation and activates account.



CHAPTER 3
(Web Based Projects)

SYSTEM DESIGN FOR GDP ANALYSIS PROJECT

3.1. Architecture Diagram

The system follows a 3-Tier Architecture to ensure scalability, separation of concerns, and easy maintenance. The architecture includes the following layers:

1. Presentation Layer (Client): Responsible for user interfaces, allowing interaction with the system through web browsers or mobile apps.
2. Business Logic Layer (Application): Hosts the algorithms for GDP analysis, including data cleaning, model processing, and statistical computation.
3. Data Layer (Database): Stores the GDP data and analysis results, maintaining consistency and supporting queries.

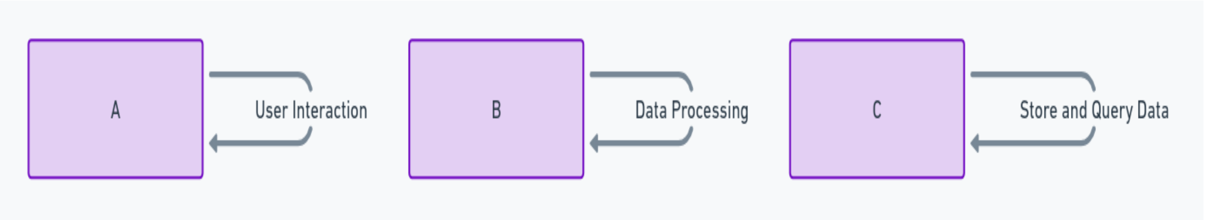


Figure 3.1: 3-Tier Architecture Diagram

3.2. Class Diagram

The class diagram illustrates the major classes used in the project:

- **DataLoader**: Handles the import of GDP data from CSV or APIs.
- **GDPAnalyzer**: Includes methods for statistical analysis, growth trends, and data prediction.
- **GDPVisualizer**: Creates charts and reports for GDP data visualization.
- **User**: Represents users accessing the platform (e.g., Admin, Analyst).
- **Dashboard**: Manages interactive visual dashboards displaying data and analysis results.

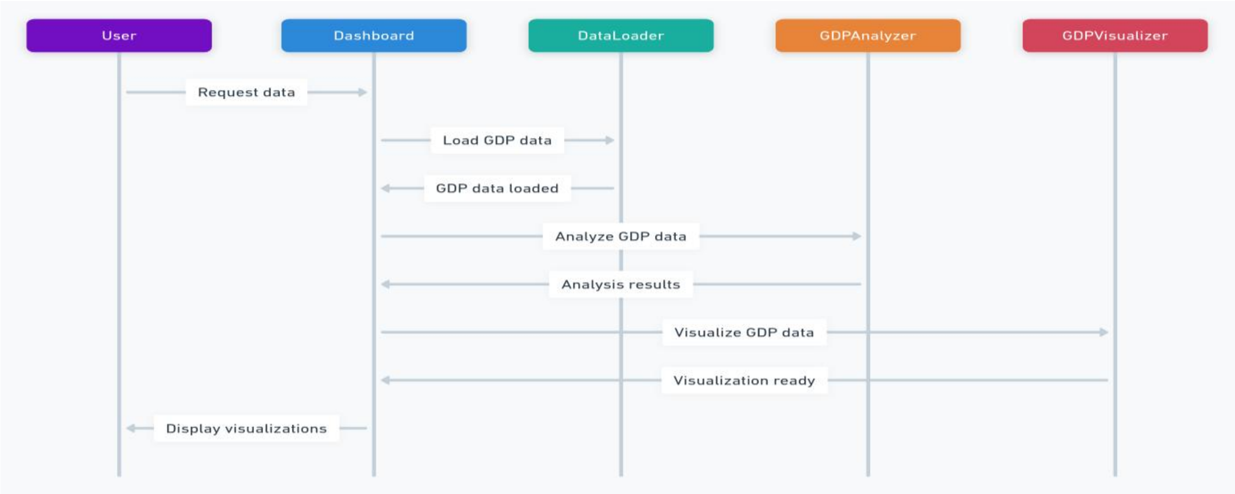


Figure 3.2: Class Diagram

3.3 Component Structure (Based on Project Files)

From the project directory structure, the system uses modular React components and a clean file structure:

/src/pages/

- *Home.jsx: Welcome and overview of system*
- *Dashboard.jsx: Visual summaries (e.g., GDP trends, population, FDI)*
- *Prediction.jsx: Handles model selection, prediction requests and outputs*
- *Auth.jsx: Manages login/registration*
- *Post.jsx: Likely handles blog posts or user-generated reports*

/backend/

- *app.py: Flask backend for routing, prediction, and preprocessing*
- */model/: Contains ML models (e.g., ARIMA, Neural Networks)*
- */data/: Contains input CSV files for GDP statistics*

3.4 Data Flow and API Integration

The application follows a structured ***data flow****:*

1. ***Data Upload or API Fetch***
   → via file input or real-time fetch (e.g., World Bank API)
2. ***Preprocessing & Analysis***
   → conducted in backend (Python) using Pandas, NumPy, Statsmodels, Scikit-learn
3. ***Visualization & Interaction***
   → via React UI using Chart.js, D3.js, or SVGs
4. ***Prediction Output***
   → results returned from Flask backend to frontend (JSON) and displayed in Prediction.jsx

3.5 Sequence Diagram Overview

User → Home.jsx → Uploads CSV
↓
Flask API (`/predict`) processes data via GDPAnalyzer
↓
Model returns result (e.g., future GDP value)
↓
Result rendered on Dashboard or Prediction.jsx
↓
User can export charts as PNG or download reports

3.6 Security and Access Control

- JWT or session-based login system managed in Auth.jsx
- Backend handles authentication, rate limiting for API requests
- HTTPS recommended for secure data communication

CHAPTER 4
IMPLEMENTATION AND RESULTS

4.1. Software and Hardware Requirements

Software Requirements:

- Python (For backend analysis and modeling)
- Flask (Web framework)
- VS code
- MySQL (Database for storing GDP data)
- Pandas, Numpy (For data analysis)
- Plotly, Matplotlib (For visualizations)

Hardware Requirements:

- Minimum 4 GB RAM
- Multi-core processor
- 256 GB HDD/SSD (For dataset storage)

4.2. Assumptions and Dependencies

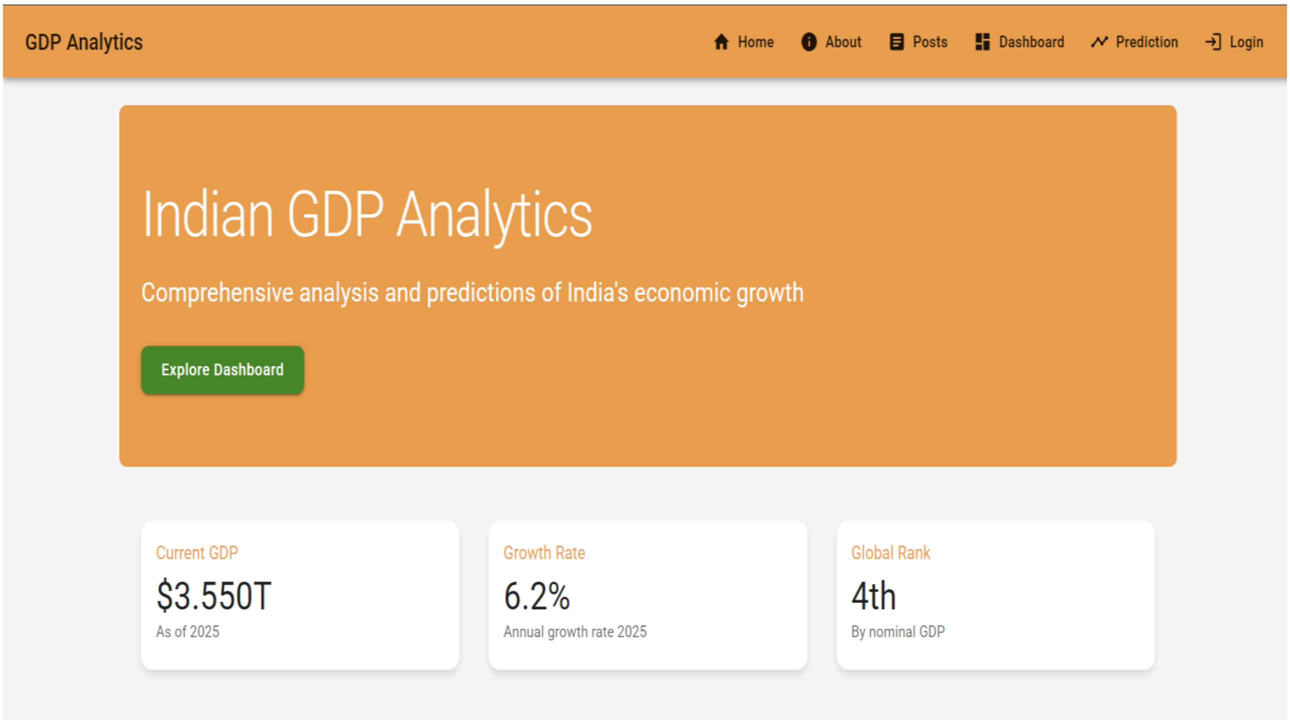
- The system assumes that the external data sources (APIs) for GDP data are available and reliable.
- The system relies on accurate historical GDP data for valid forecasting.
- All users accessing the system will have a stable internet connection.

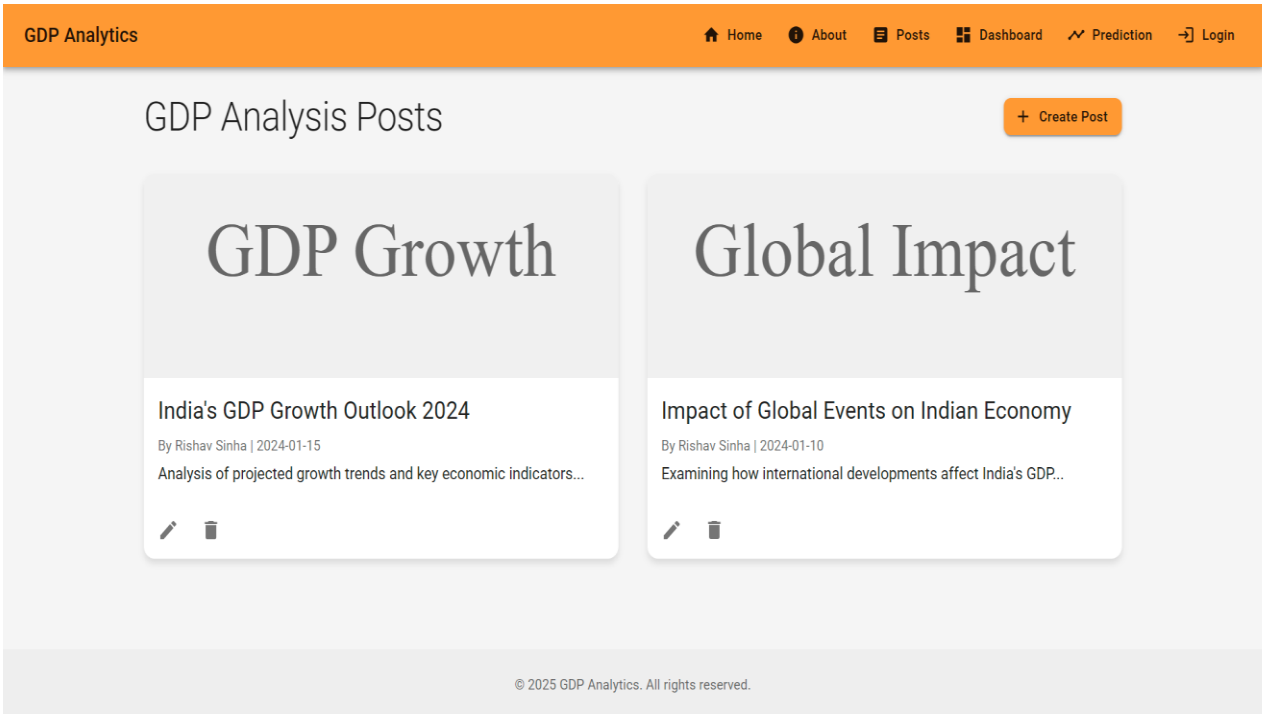
4.3. Constraints

- Data accuracy is limited by the quality and completeness of the external data sources.
- The system should not exceed server capacity (e.g., memory usage should be optimized).

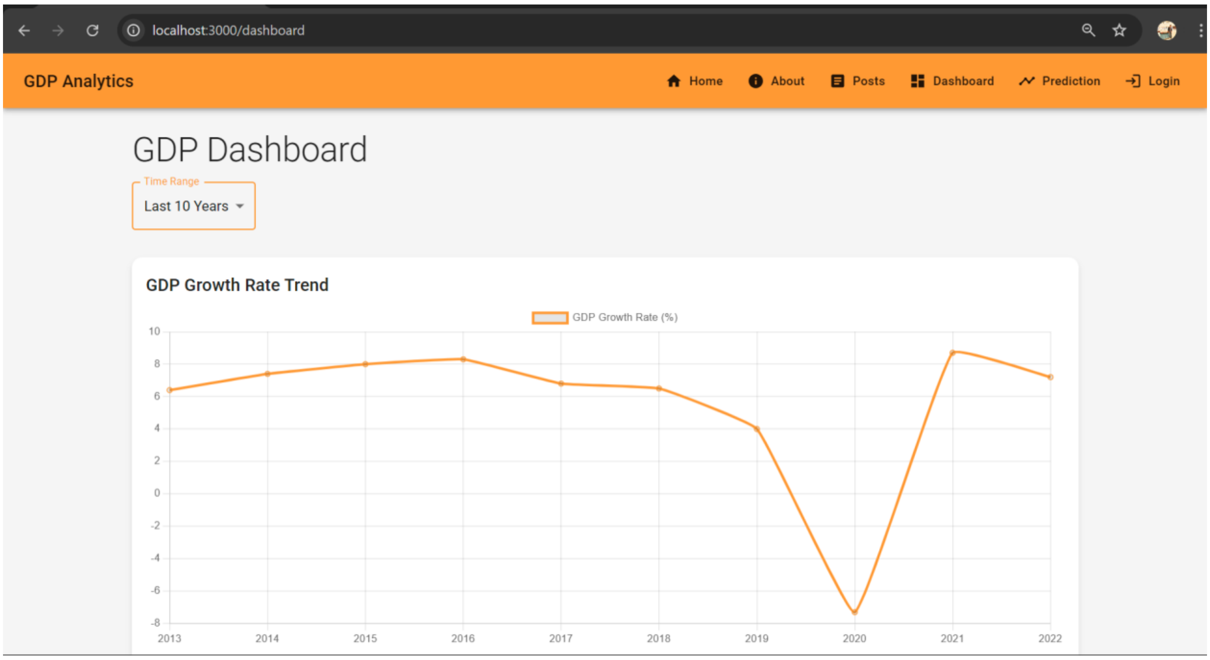
4.4. Implementation Details

Website Interface

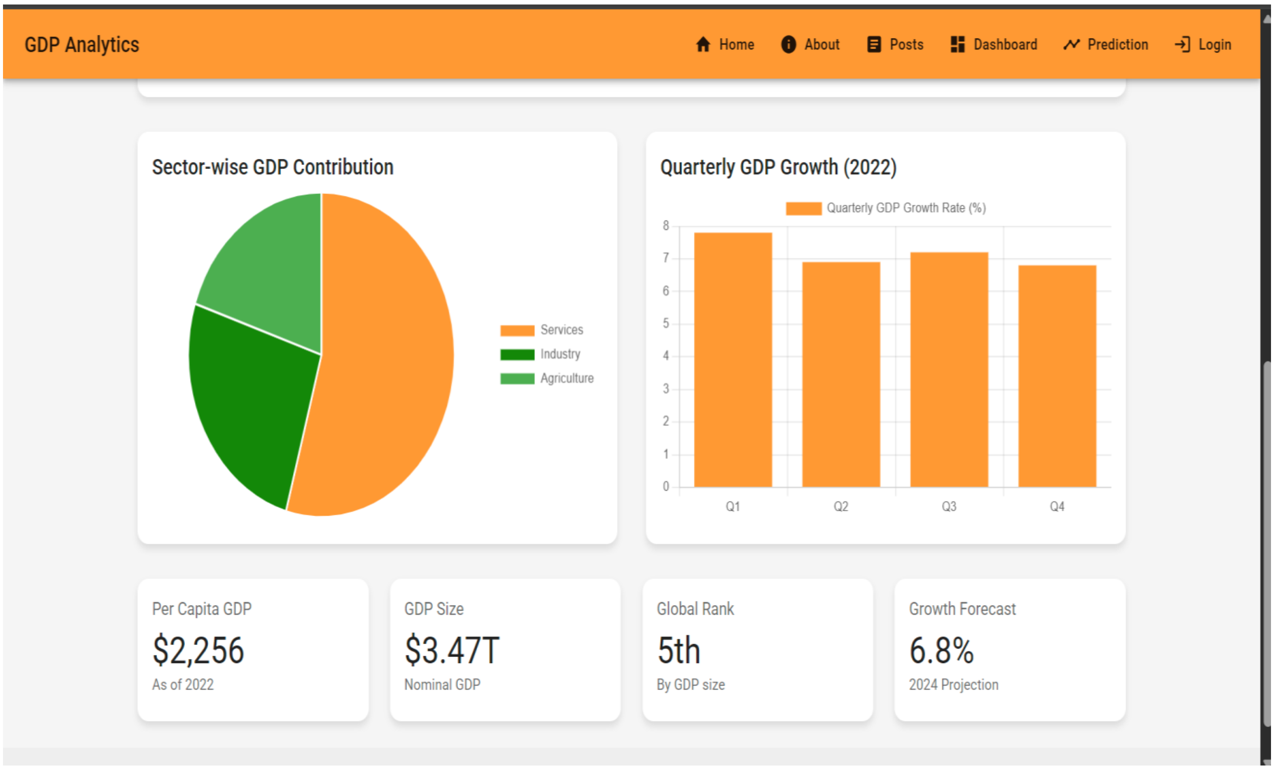




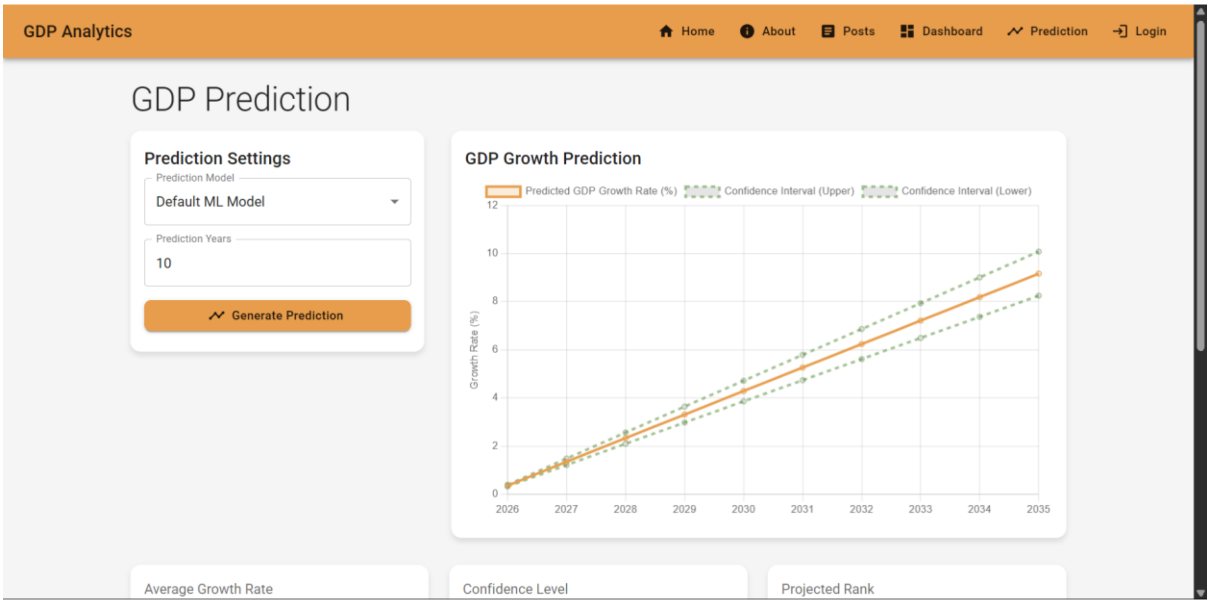
DashBoard

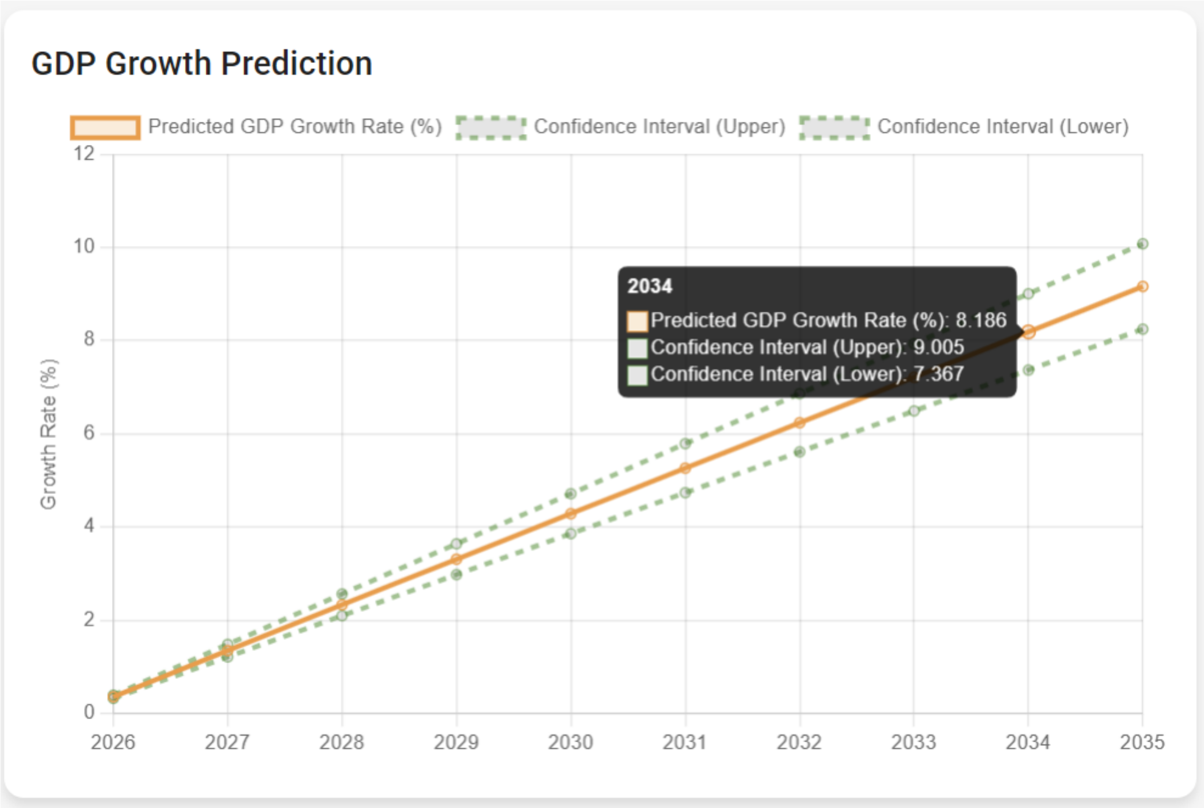


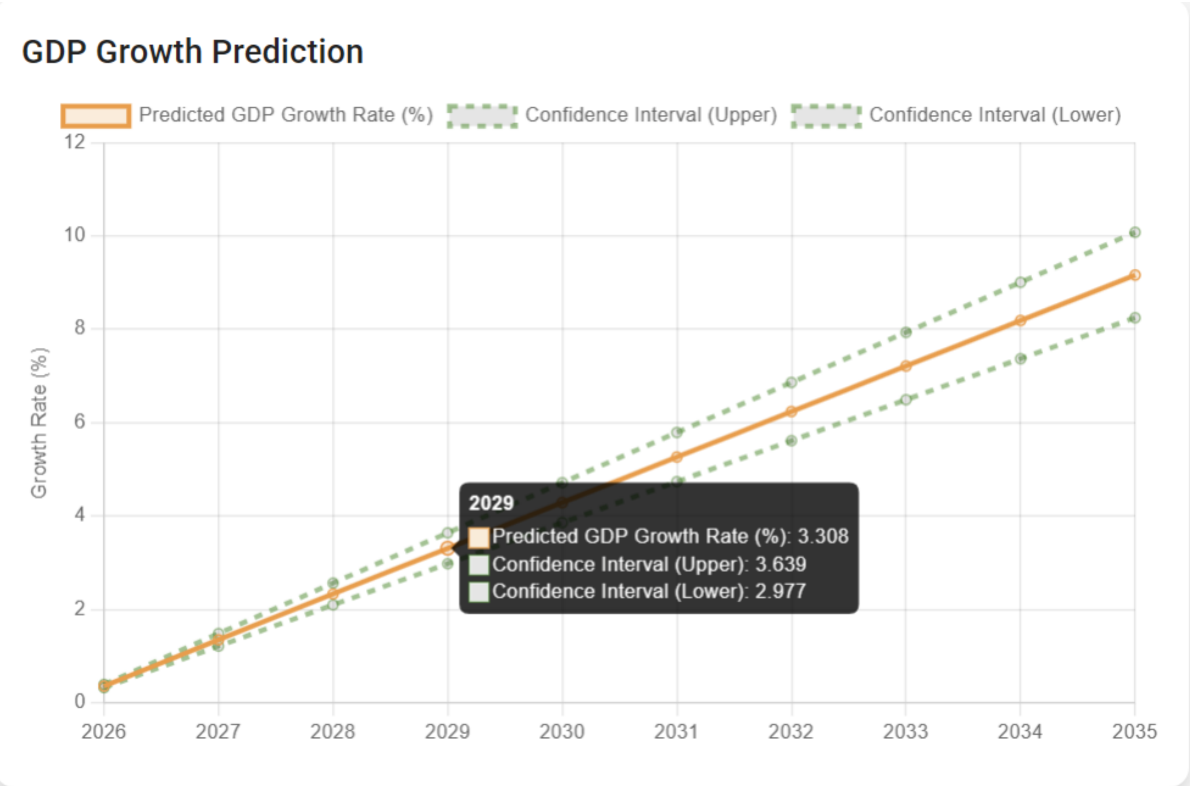
Previous years



Predictions







Average Growth Rate	Confidence Level	Projected Rank
7.06%	95%	3rd
Predicted 5-year average	Statistical significance	Expected global position by 2027

Basic libraries and Data set sample

[1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

%matplotlib inline

import warnings
warnings.filterwarnings('ignore')
```

[2]:
```
df = pd.read_csv('IndDataset.csv')
df = df.iloc[:, 1:]  # Drop the first column by selecting all columns from the second onwards
```

[7]:
```
df.sample(5)
```

[7]:

	Year	Population, total	Population growth (annual %)	Life expectancy at birth, total (years)	GDP (current US$)	GDP growth (annual %)	Inflation, GDP deflator (annual %)	Agriculture, forestry, and fishing, value added (% of GDP)	Industry (including construction), value added (% of GDP)	Exports of goods and services (% of GDP)	Imports of goods and services (% of GDP)	Military expenditure (% of GDP)	Merchandise trade (% of GDP)	Foreig direc investmen net inflow (BoP, curren US$
	2017	1338658835	1.062597	69.165	2.650000e+12	7.043821	3.778099	16.357810	26.478810	18.782685	21.940126	2.509625	28.241058	3.996609e+1
	2001	1075000085	1.728768	62.907	4.850000e+11	4.823966	3.215616	21.620505	26.487777	12.558380	13.434875	2.924436	19.312954	5.128094e+0
	1997	1000900030	1.869173	61.233	4.160000e+11	4.049821	6.476271	24.345357	27.837031	10.690717	11.928670	2.647732	18.380843	3.577330e+0
	2012	1265782790	1.231623	67.545	1.830000e+12	5.456359	7.934388	16.845377	29.398528	24.534431	31.259291	2.537347	43.034893	2.399569e+1
	1994	945601831	1.943245	59.840	3.270000e+11	6.658924	9.980045	26.388619	27.628201	9.888085	10.190059	2.664674	15.847501	9.732715e+0

Some Basic Operations are:

```
[9]: df.columns
```

```
[9]: Index(['Year', 'Population, total', 'Population growth (annual %)',
       'Life expectancy at birth, total (years)', 'GDP (current US$)',
       'GDP growth (annual %)', 'Inflation, GDP deflator (annual %)',
       'Agriculture, forestry, and fishing, value added (% of GDP)',
       'Industry (including construction), value added (% of GDP)',
       'Exports of goods and services (% of GDP)',
       'Imports of goods and services (% of GDP)',
       'Military expenditure (% of GDP)', 'Merchandise trade (% of GDP)',
       'Foreign direct investment, net inflows (BoP, current US$)'],
      dtype='object')
```

```
[12]: df.isnull().sum()
```

```
[12]: Year                                                          0
Population, total                                             0
Population growth (annual %)                                  0
Life expectancy at birth, total (years)                       0
GDP (current US$)                                             0
GDP growth (annual %)                                         0
Inflation, GDP deflator (annual %)                            0
Agriculture, forestry, and fishing, value added (% of GDP)    0
Industry (including construction), value added (% of GDP)     0
Exports of goods and services (% of GDP)                      0
Imports of goods and services (% of GDP)                      0
Military expenditure (% of GDP)                               0
Merchandise trade (% of GDP)                                  0
Foreign direct investment, net inflows (BoP, current US$)     0
dtype: int64
```

Luckily, we have no missing values in our dataset.

```
[10]: df.describe()
```

[10]:

	Year	Population, total	Population growth (annual %)	Life expectancy at birth, total (years)	GDP (current US$)	GDP growth (annual %)	Inflation, GDP deflator (annual %)	Agriculture, forestry, and fishing, value added (% of GDP)	Industry (including construction), value added (% of GDP)	Exports of goods and services (% of GDP)	Imports of goods and services (% of GDP)	Military expenditure (% of GDP)	Merchan trade (' G
count	31.000000	3.100000e+01	31.000000	31.000000	3.100000e+01	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.000000	31.00
mean	2005.000000	1.140254e+09	1.542113	64.393323	1.087110e+12	5.829473	6.589842	20.349515	28.034188	16.521108	18.883460	2.647496	25.65
std	9.092121	1.572151e+08	0.353210	3.695462	8.552676e+11	3.052078	2.932180	4.211103	1.711248	5.625189	7.193331	0.224537	9.65
min	1990.000000	8.732778e+08	0.960000	57.865000	1.396387e+09	-7.250000	2.279589	15.406726	24.881008	7.053350	8.452911	2.100000	10.90
25%	1997.500000	1.010192e+09	1.207319	61.451000	4.045000e+11	4.923920	3.746891	16.791438	26.700277	10.931219	12.352241	2.501620	18.10
50%	2005.000000	1.147610e+09	1.579710	64.500000	7.090000e+11	6.386106	6.620000	17.814691	27.628201	18.661969	19.644689	2.651497	24.87
75%	2012.500000	1.273314e+09	1.854415	67.738000	1.825000e+12	7.861135	8.849365	24.417838	29.308817	20.600110	24.671554	2.825205	31.46
max	2020.000000	1.396387e+09	2.076089	69.730000	2.880000e+12	8.845756	13.751819	27.333218	31.136719	25.430861	31.259291	3.146215	43.03

```
[14]: df.duplicated().sum()
```

```
[14]: 0
```

Luckily, we have no Duplicate values in our dataset.

[16]:
```
df.corr()
```

[16]:

	Year	Population, total	Population growth (annual %)	Life expectancy at birth, total (years)	GDP (current US$)	GDP growth (annual %)	Inflation, GDP deflator (annual %)	Agriculture, forestry, and fishing, value added (% of GDP)	Industry (including construction), value added (% of GDP)	Exports of goods and services (% of GDP)	Imports of goods and services (% of GDP)	Military expenditure (% of GDP)	Merchandis trade (% (GDI
Year	1.000000	0.999100	-0.995739	0.996747	0.806182	-0.065933	-0.531762	-0.813573	0.035845	0.838582	0.796870	-0.682677	0.64072
Population, total	0.999100	1.000000	-0.992812	0.998439	0.796471	-0.057622	-0.529109	-0.826504	0.070256	0.855427	0.814469	-0.674772	0.65771
Population growth (annual %)	-0.995739	-0.992812	1.000000	-0.991191	-0.822411	0.073836	0.503293	0.789922	-0.007152	-0.830688	-0.788869	0.688291	-0.6414
Life expectancy at birth, total (years)	0.996747	0.998439	-0.991191	1.000000	0.808517	-0.022427	-0.536568	-0.841039	0.086677	0.865819	0.826768	-0.666145	0.68147

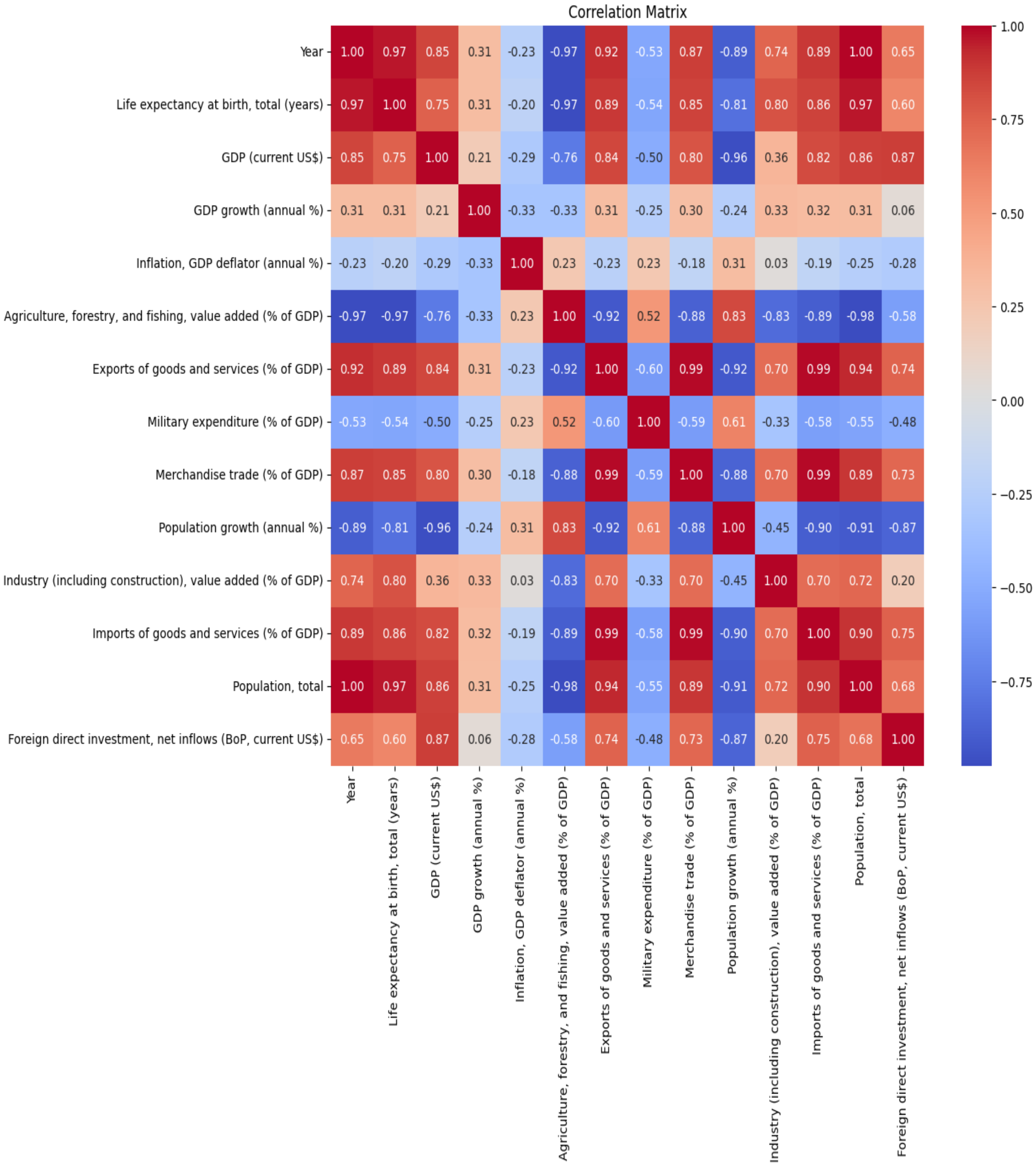
[9]:
```
df.columns
```

[9]:
```
Index(['Year', 'Life expectancy at birth, total (years)', 'GDP (current US$)',
       'GDP growth (annual %)', 'Inflation, GDP deflator (annual %)',
       'Agriculture, forestry, and fishing, value added (% of GDP)',
       'Exports of goods and services (% of GDP)',
       'Military expenditure (% of GDP)', 'Merchandise trade (% of GDP)',
       'Population growth (annual %)',
       'Industry (including construction), value added (% of GDP)',
       'Imports of goods and services (% of GDP)', 'Population, total',
       'Foreign direct investment, net inflows (BoP, current US$)'],
      dtype='object')
```

Basic EDA Operations and Data Inspection

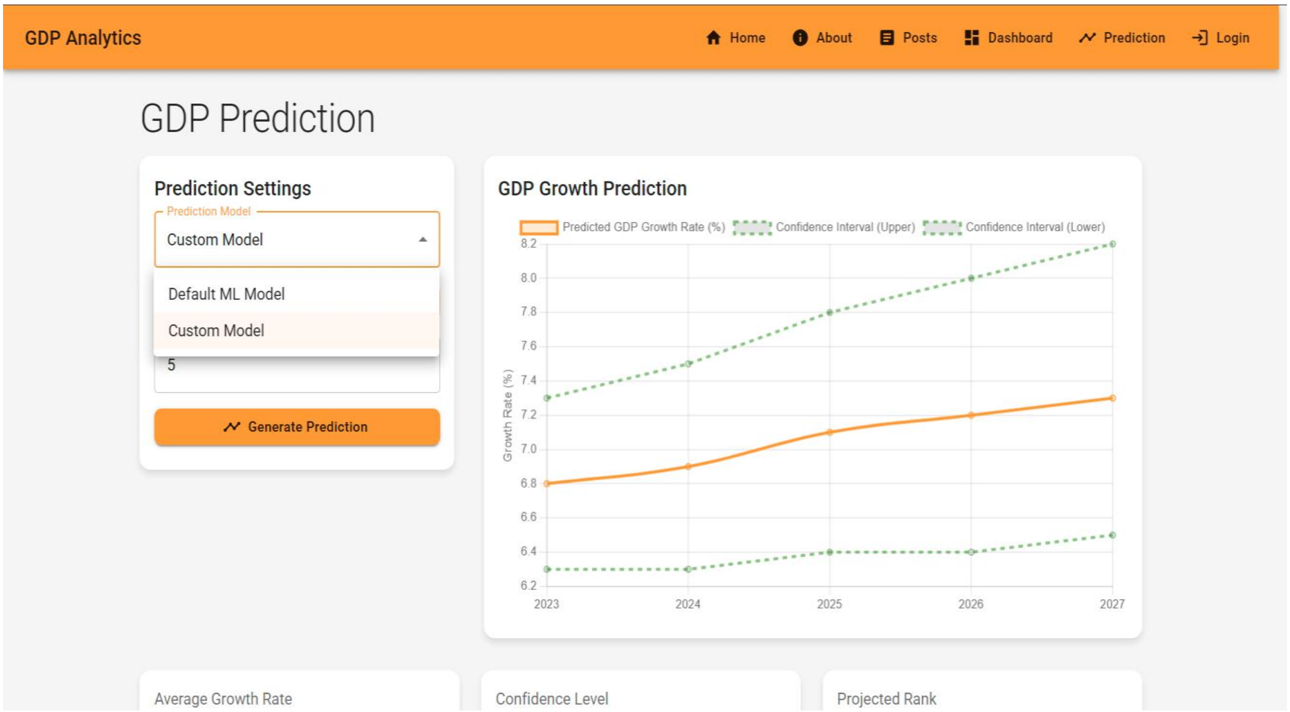
- **.head() and .tail()**: Displayed the first and last few rows of the dataset, allowing a quick overview of the data structure and column types.
- **.shape**: Returned the number of rows and columns, giving insight into the dataset size.
- **.info()**: Provided an overview of each column, including data types and non-null counts, useful for detecting data type inconsistencies and missing values.
- **.describe()**: Generated summary statistics (mean, median, std, min, max, etc.) for numerical columns, helping to understand data distribution and detect potential outliers.
- **.isnull().sum()**: Identified the number of missing values in each column, essential for handling nulls through imputation or removal.
- **.nunique()**: Counted unique values in each column, useful for understanding categorical data and detecting low-variance features.

HeatMap(Correlation Matrix)

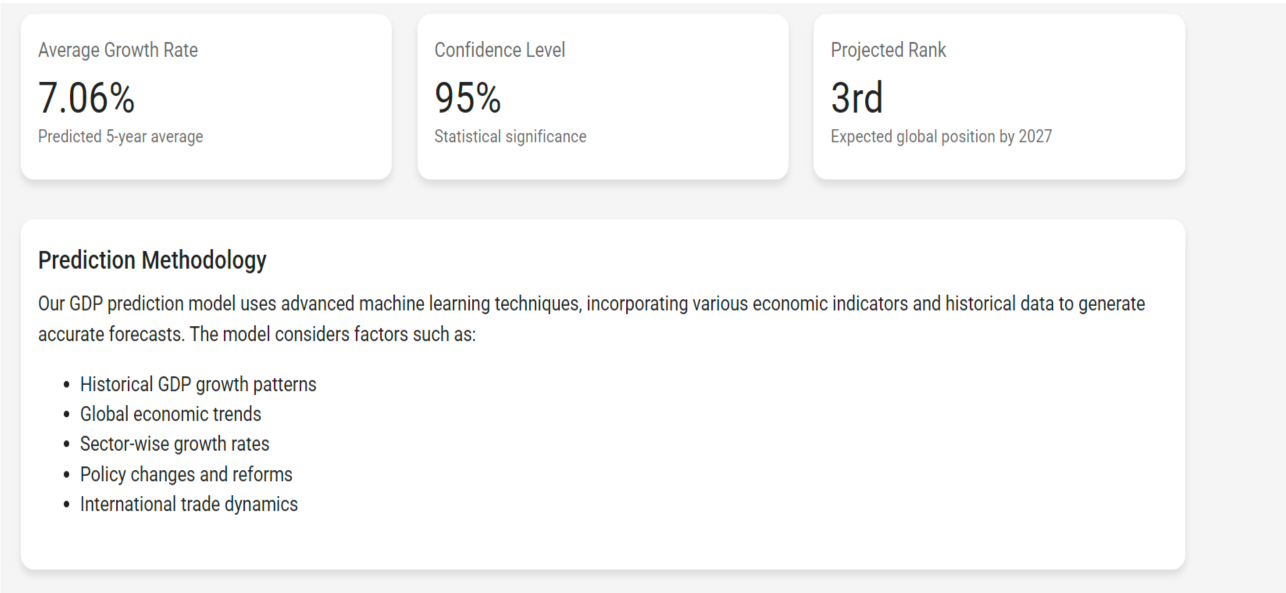


A heatmap was used to visualize the correlation between GDP and various economic indicators. This helped identify which factors had the strongest positive or negative relationships with GDP. Darker colors indicated stronger correlations, enabling better feature selection and deeper insights into economic trends.

Prediction Interface



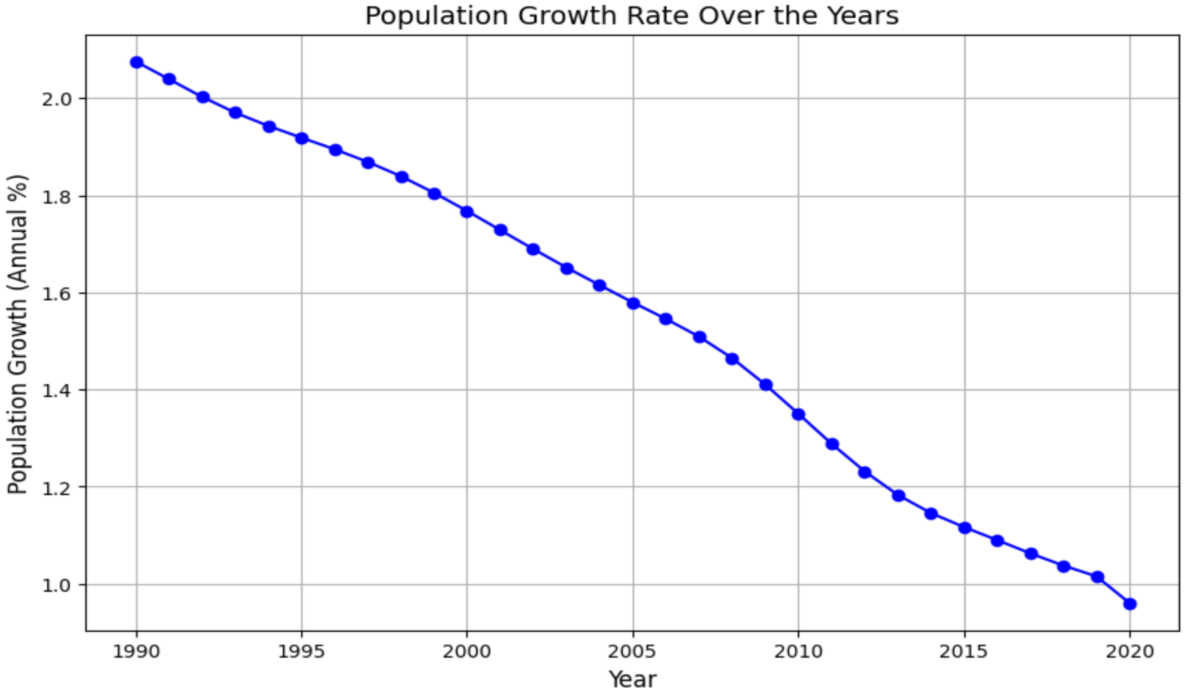
Some Statistics

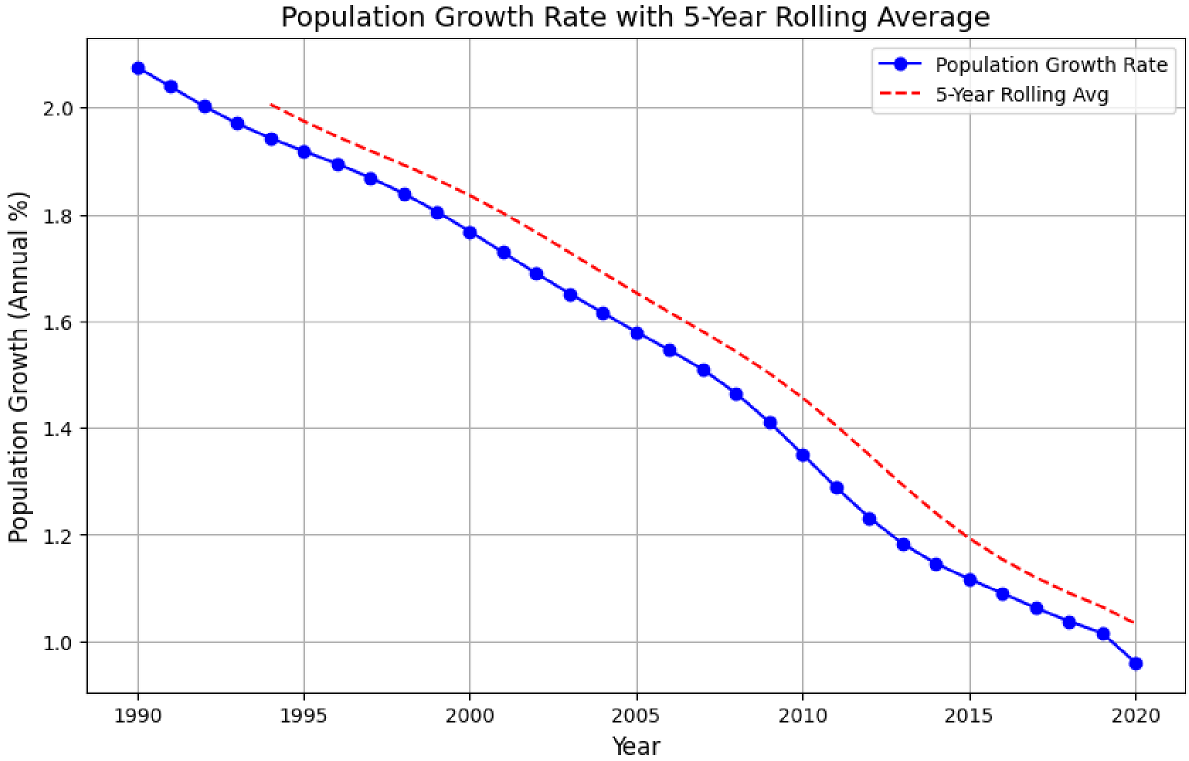


Exploratory Data Analysis (EDA)

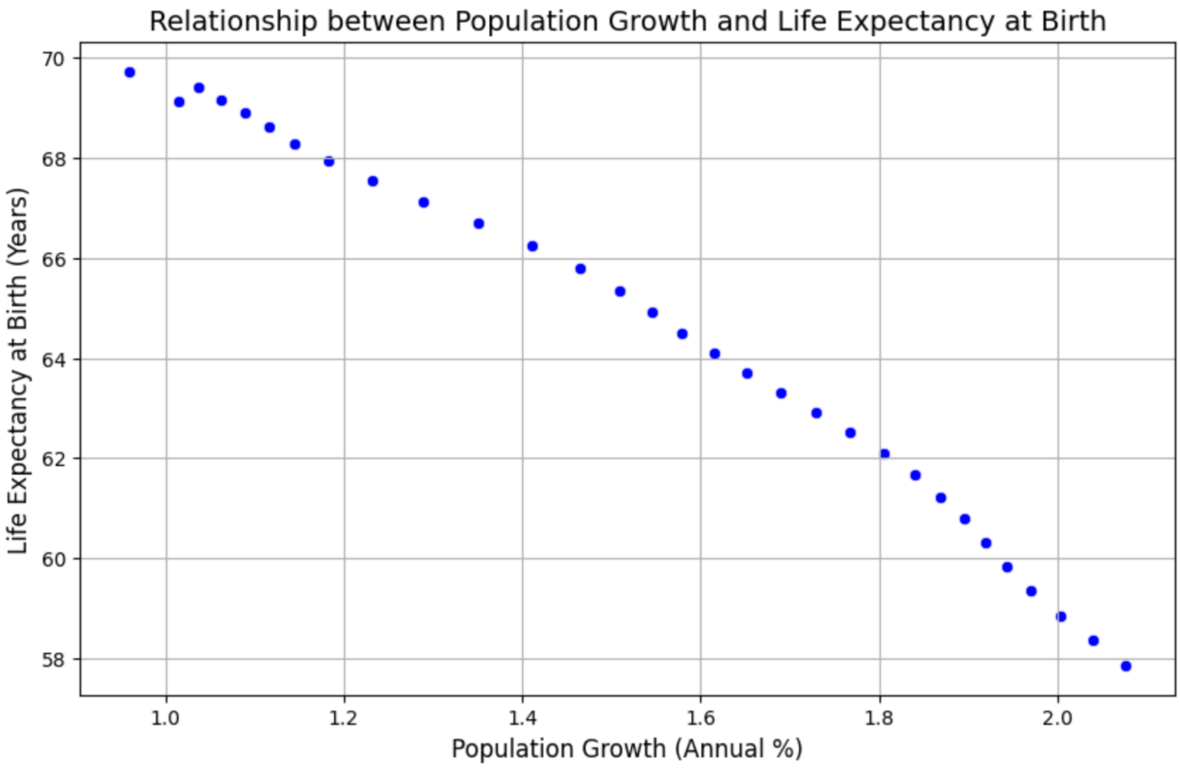
Exploratory Data Analysis (EDA) involves summarizing and visualizing the main characteristics of a dataset to uncover patterns, detect anomalies, and gain insights before modeling.

How has the population growth rate changed over the years? Are there any noticeable trends in certain time periods?



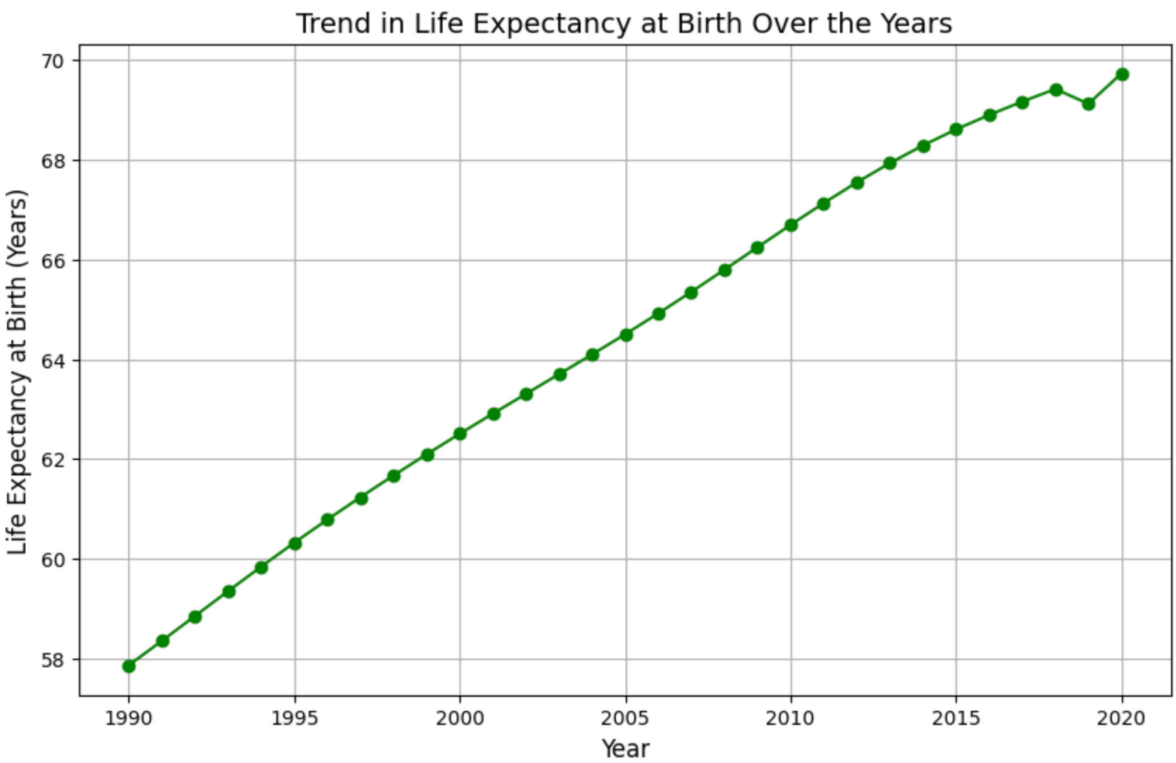


What is the relationship between population growth and life expectancy? Does a higher growth rate correlate with changes in life expectancy?



Correlation between Population Growth and Life Expectancy at Birth: -0.99

What is the trend in life expectancy at birth over the years? Are there any significant jumps or declines?



Significant Jumps or Declines in Life Expectancy at Birth:

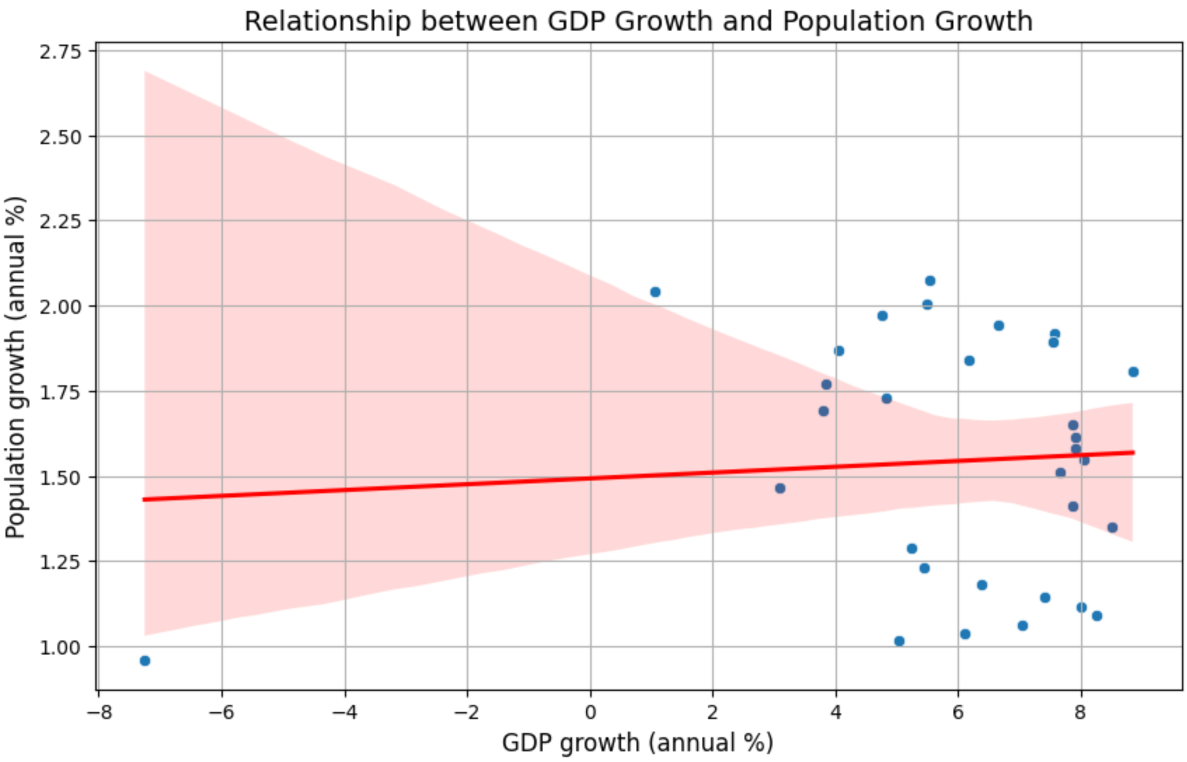
	Year	Life expectancy at birth, total (years)	Yearly Change
1	1991	58.353	0.488
2	1992	58.851	0.498
3	1993	59.349	0.498
4	1994	59.840	0.491
5	1995	60.320	0.480
6	1996	60.783	0.463
7	1997	61.233	0.450
8	1998	61.669	0.436
9	1999	62.093	0.424
10	2000	62.505	0.412
11	2001	62.907	0.402
15	2005	64.500	0.405
16	2006	64.918	0.418
17	2007	65.350	0.432
18	2008	65.794	0.444
19	2009	66.244	0.450
20	2010	66.693	0.449
21	2011	67.130	0.437
22	2012	67.545	0.415
30	2020	69.730	0.609

How has the GDP (in current US$) changed from 1990 to 2020? Are there any specific years of major growth or decline?

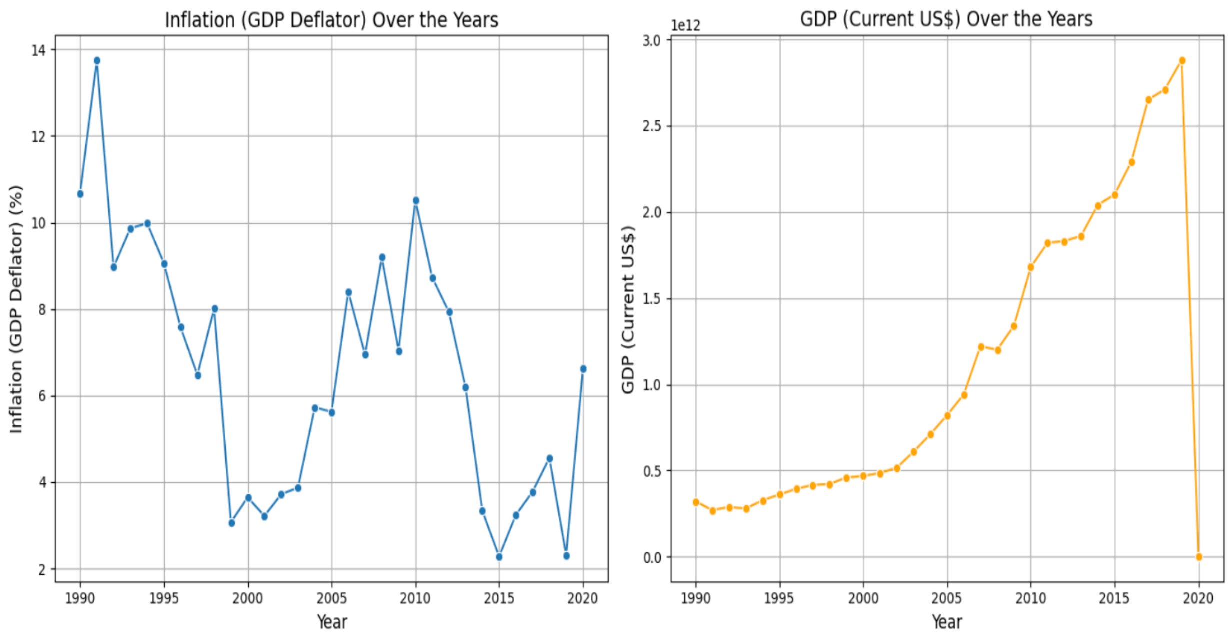
Years of Major Growth or Decline in GDP:

	Year	GDP (current US$)	GDP Change
1	1991	2.700000e+11	-5.100000e+10
2	1992	2.880000e+11	1.800000e+10
3	1993	2.790000e+11	-9.000000e+09
4	1994	3.270000e+11	4.800000e+10
5	1995	3.600000e+11	3.300000e+10
6	1996	3.930000e+11	3.300000e+10
7	1997	4.160000e+11	2.300000e+10
8	1998	4.210000e+11	5.000000e+09
9	1999	4.590000e+11	3.800000e+10
10	2000	4.680000e+11	9.000000e+09
11	2001	4.850000e+11	1.700000e+10
12	2002	5.150000e+11	3.000000e+10
13	2003	6.080000e+11	9.300000e+10
14	2004	7.090000e+11	1.010000e+11
15	2005	8.200000e+11	1.110000e+11
16	2006	9.400000e+11	1.200000e+11
17	2007	1.220000e+12	2.800000e+11
18	2008	1.200000e+12	-2.000000e+10
19	2009	1.340000e+12	1.400000e+11
20	2010	1.680000e+12	3.400000e+11
21	2011	1.820000e+12	1.400000e+11
22	2012	1.830000e+12	1.000000e+10
23	2013	1.860000e+12	3.000000e+10
24	2014	2.040000e+12	1.800000e+11
25	2015	2.100000e+12	6.000000e+10
26	2016	2.290000e+12	1.900000e+11
27	2017	2.650000e+12	3.600000e+11
28	2018	2.710000e+12	6.000000e+10
29	2019	2.880000e+12	1.700000e+11
30	2020	1.396387e+09	-2.878604e+12

What is the relationship between GDP growth and population growth? Is higher GDP growth accompanied by higher population growth?

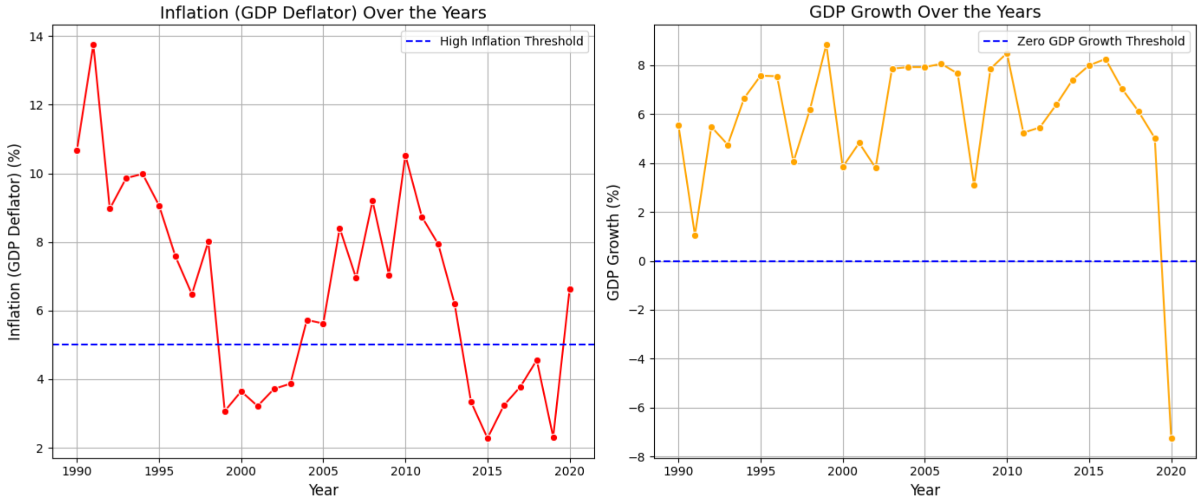


How has inflation (GDP deflator) fluctuated, and what impact does it have on the overall GDP?



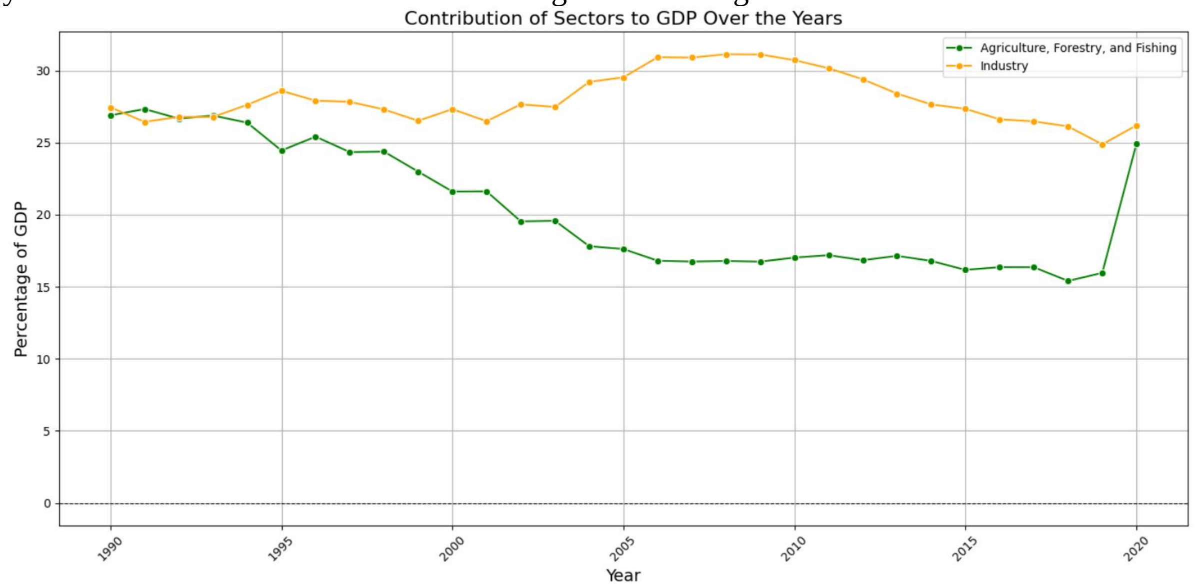

```
Correlation between Inflation (GDP deflator) and GDP: -0.42
```

Are there periods where high inflation coincides with a decrease in GDP growth?



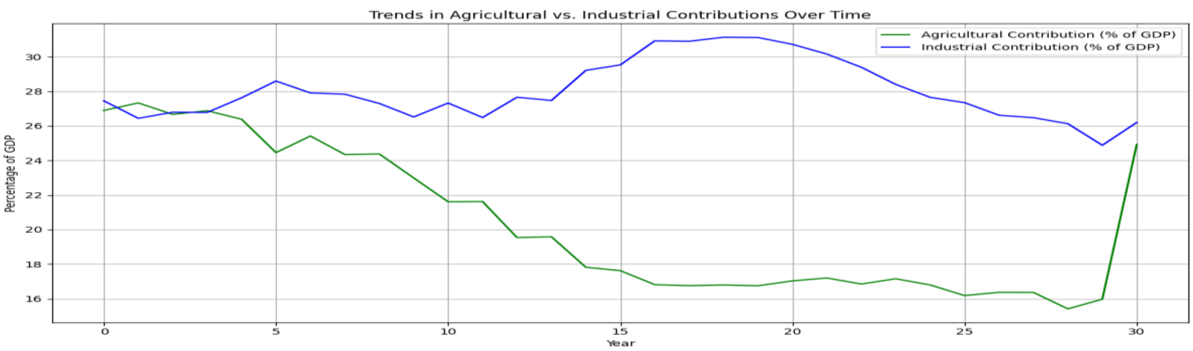

```
Periods of High Inflation and Decreasing GDP Growth:
    Year  Inflation, GDP deflator (annual %)  GDP growth (annual %)
30  2020                                6.62                  -7.25
```

How has the contribution of agriculture, industry, and fishing to GDP changed over the years? Which sector has shown the most significant change?

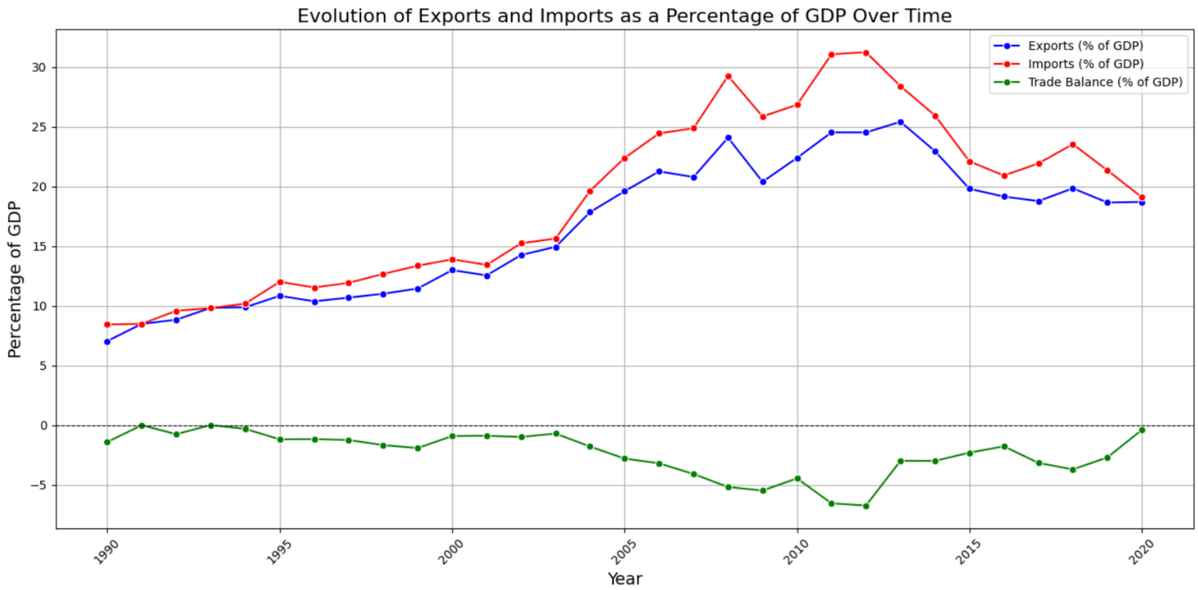


The sector with the most significant change in contribution to GDP is: Agriculture, Forestry, and Fishing with a change of 56.16%

What trends can be observed in the percentage contribution of the industrial sector versus the agricultural sector over time?



How have exports and imports as a percentage of GDP evolved over the years? Is there a trade surplus or deficit in any particular period?



Periods of Trade Surplus (% of GDP):

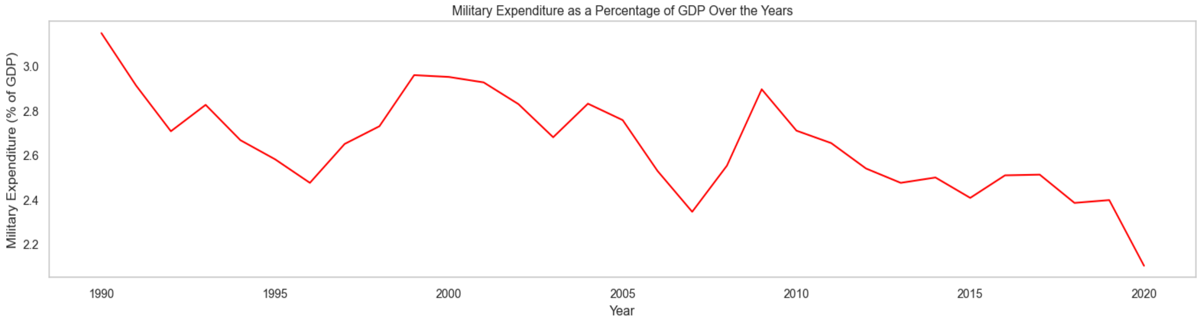
	Year	Trade Balance (% of GDP)
1	1991	0.000755
3	1993	0.016895

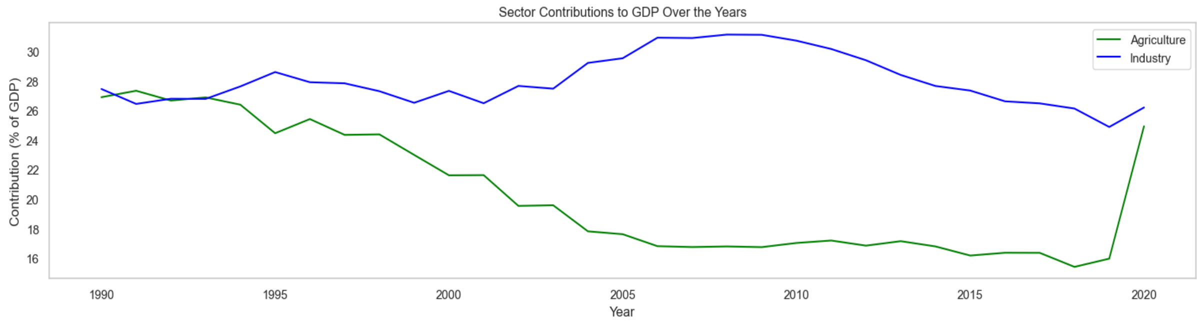
Periods of Trade Deficit (% of GDP):

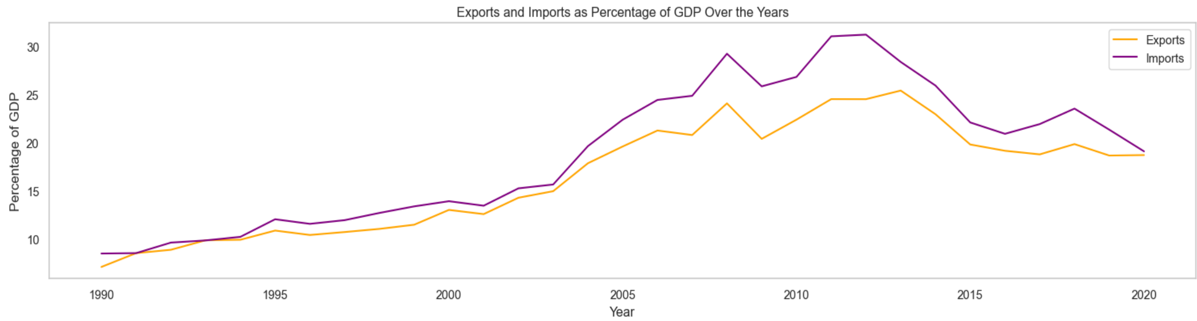
	Year	Trade Balance (% of GDP)
0	1990	-1.399561
2	1992	-0.747245
4	1994	-0.301975
5	1995	-1.179513
6	1996	-1.159149
7	1997	-1.237952
8	1998	-1.662532
9	1999	-1.911469
10	2000	-0.906450
11	2001	-0.876495
12	2002	-0.979895
13	2003	-0.696608
14	2004	-1.785564
15	2005	-2.791176
16	2006	-3.188598
17	2007	-4.086869
18	2008	-5.173506
19	2009	-5.471831
20	2010	-4.453340
21	2011	-6.543056
22	2012	-6.724860
23	2013	-2.982409
24	2014	-2.986260
25	2015	-2.296536
26	2016	-1.766016
27	2017	-3.157441
28	2018	-3.699667
29	2019	-2.694600
30	2020	-0.390000

What is the relationship between merchandise trade and foreign direct investment (FDI) inflows? Is there a correlation between higher trade volumes and increased FDI?

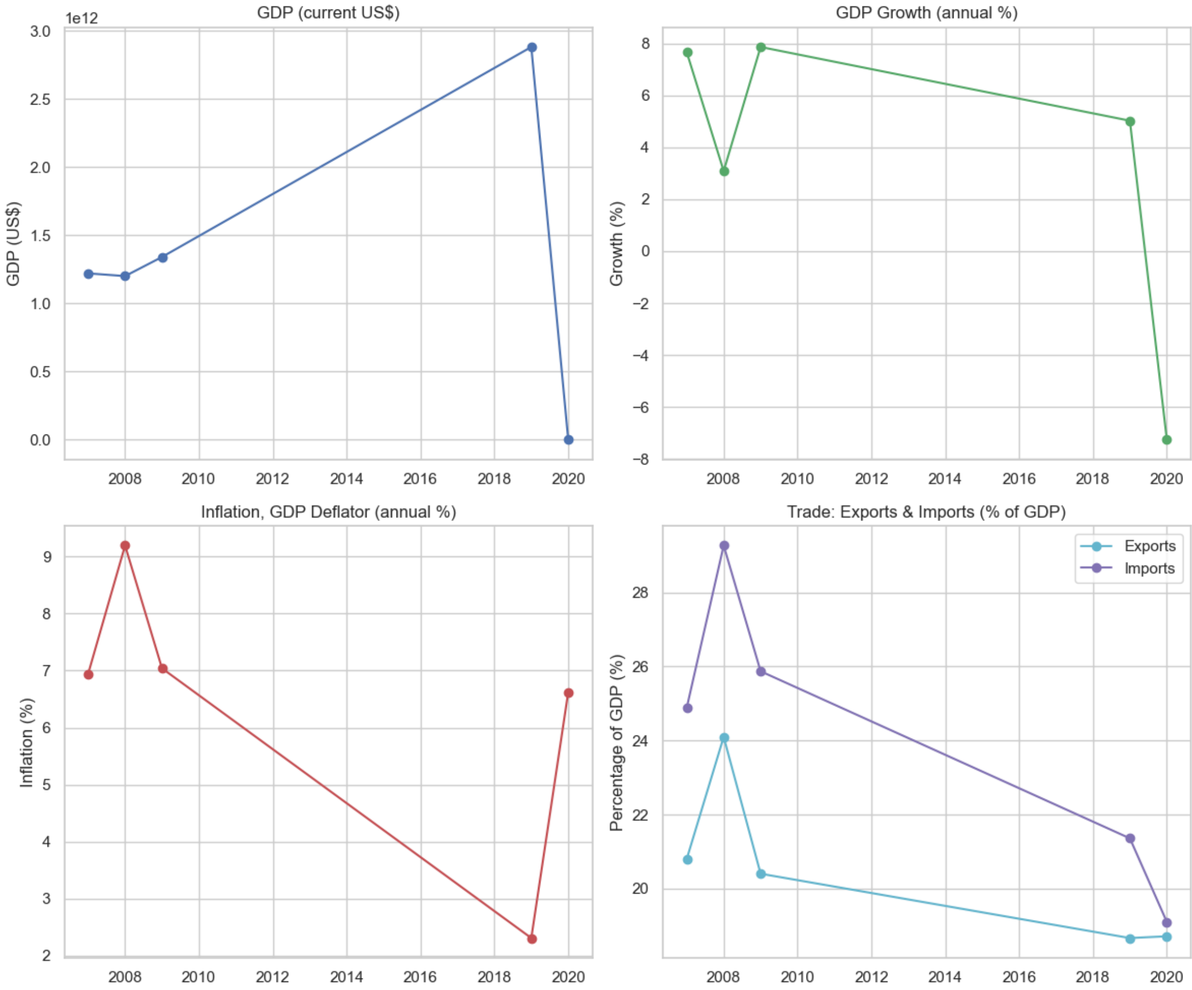
Are there any periods where military expenditure as a percentage of GDP spikes, and what is its potential impact on other economic sectors?





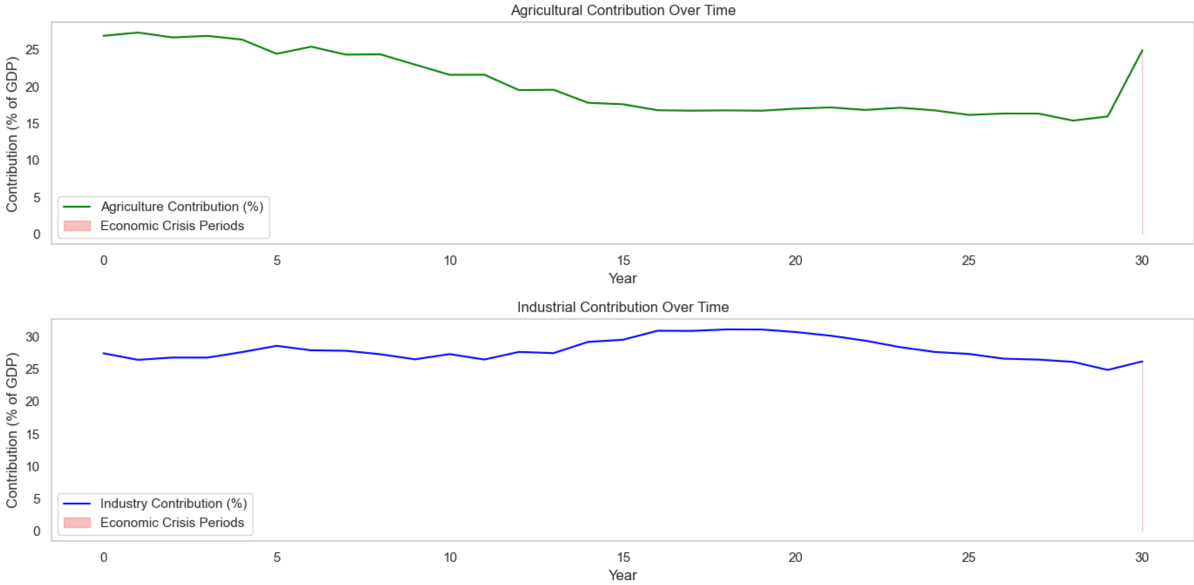


How did the global financial crisis (2008) and the COVID-19 pandemic (2020) affect economic indicators such as GDP, GDP growth, inflation, and trade?

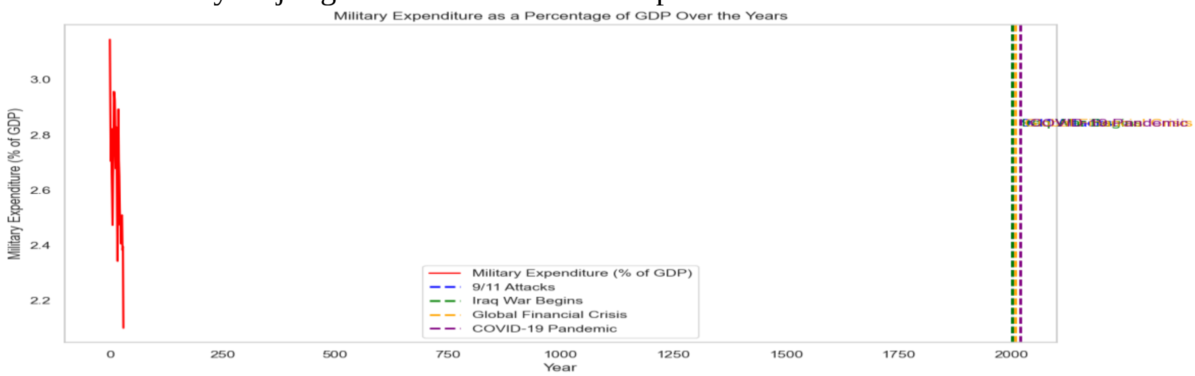



```
    Year  GDP (current US$)  GDP growth (annual %)  \
17  2007       1.220000e+12               7.660815
18  2008       1.200000e+12               3.086698
19  2009       1.340000e+12               7.861889
29  2019       2.880000e+12               5.023873
30  2020       1.396387e+09              -7.250000

    Inflation, GDP deflator (annual %)  \
17                            6.944418
18                            9.193970
19                            7.040365
29                            2.310899
30                            6.620000

    Exports of goods and services (% of GDP)  \
17                                 20.799700
18                                 24.097357
19                                 20.400519
29                                 18.661969
30                                 18.710000

    Imports of goods and services (% of GDP)
17                                 24.886569
18                                 29.270863
19                                 25.872350
29                                 21.356569
30                                 19.100000
```

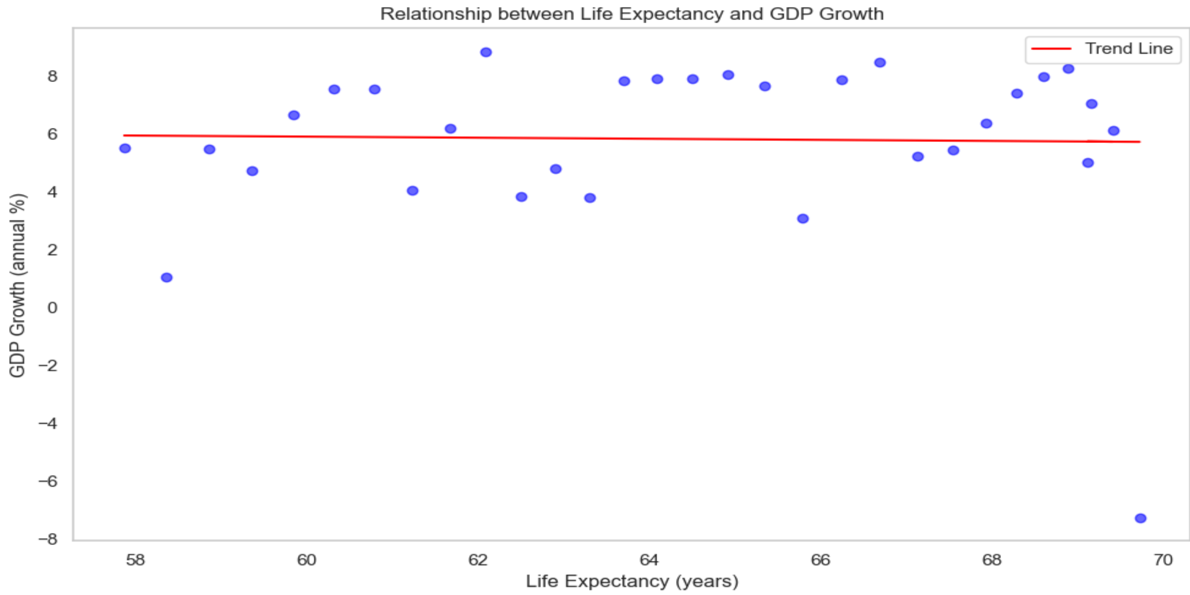
How did agricultural and industrial contributions shift during periods of economic crises or slowdowns?



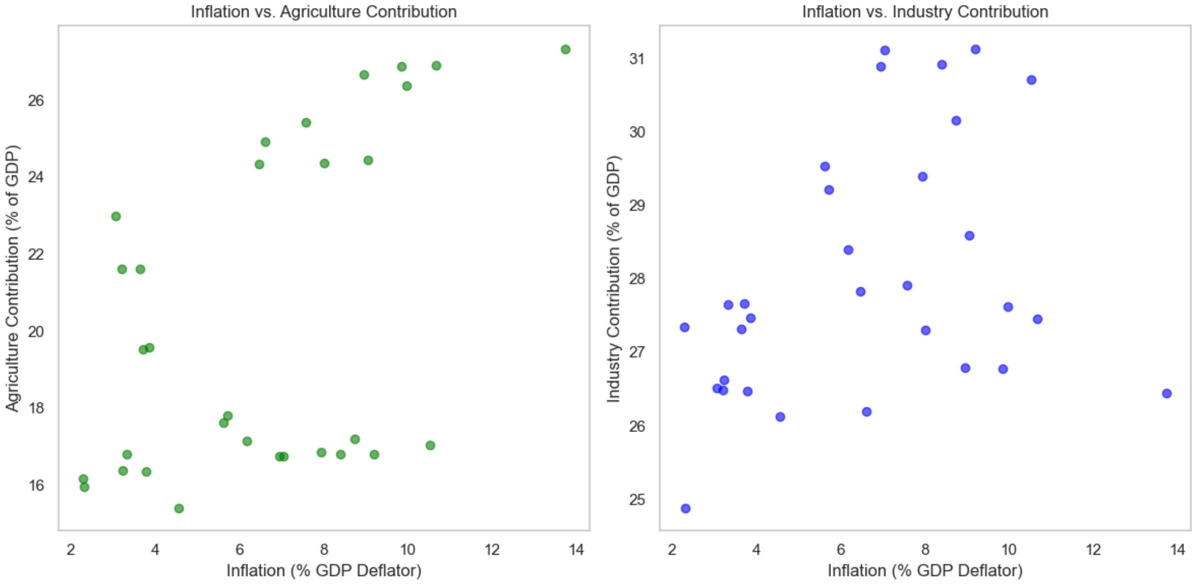
How has military expenditure as a percentage of GDP changed over the years? Does this correlate with any major global events or domestic policies?



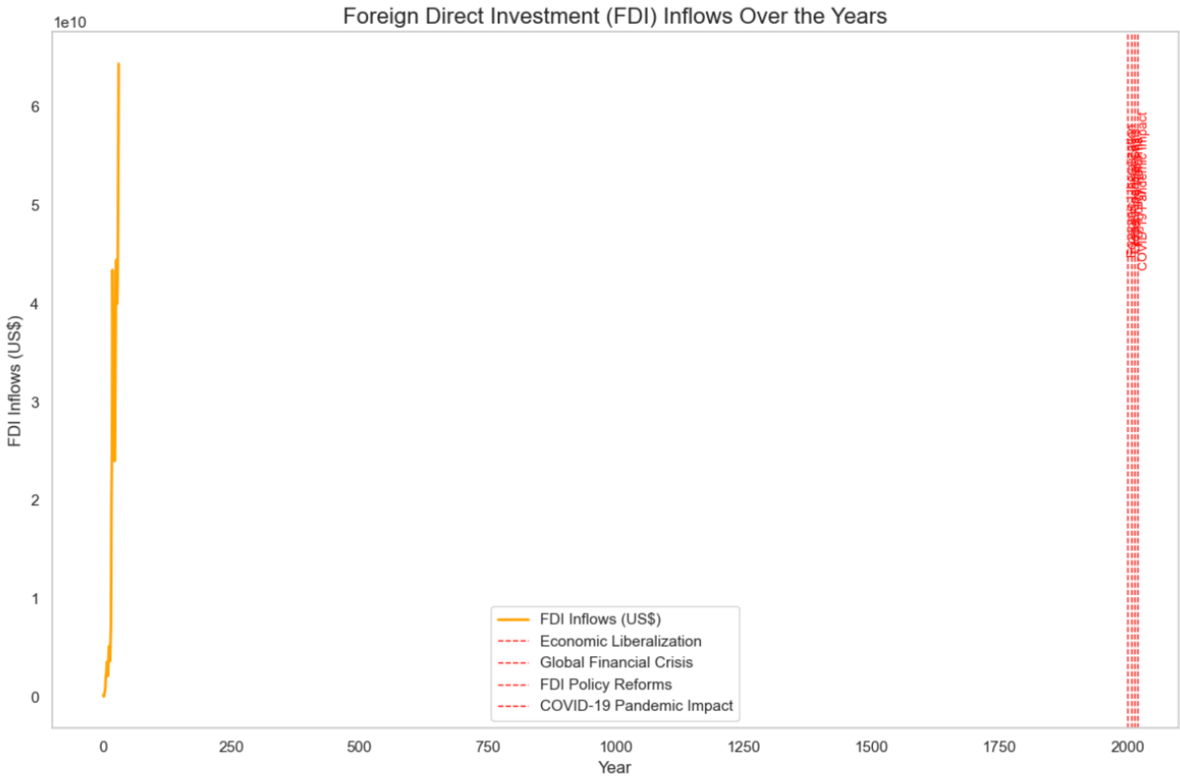
What is the relationship between life expectancy and GDP growth? Is there a direct correlation between higher life expectancy and better economic performance?



How does inflation affect sector contributions to GDP, particularly in agriculture and industry?

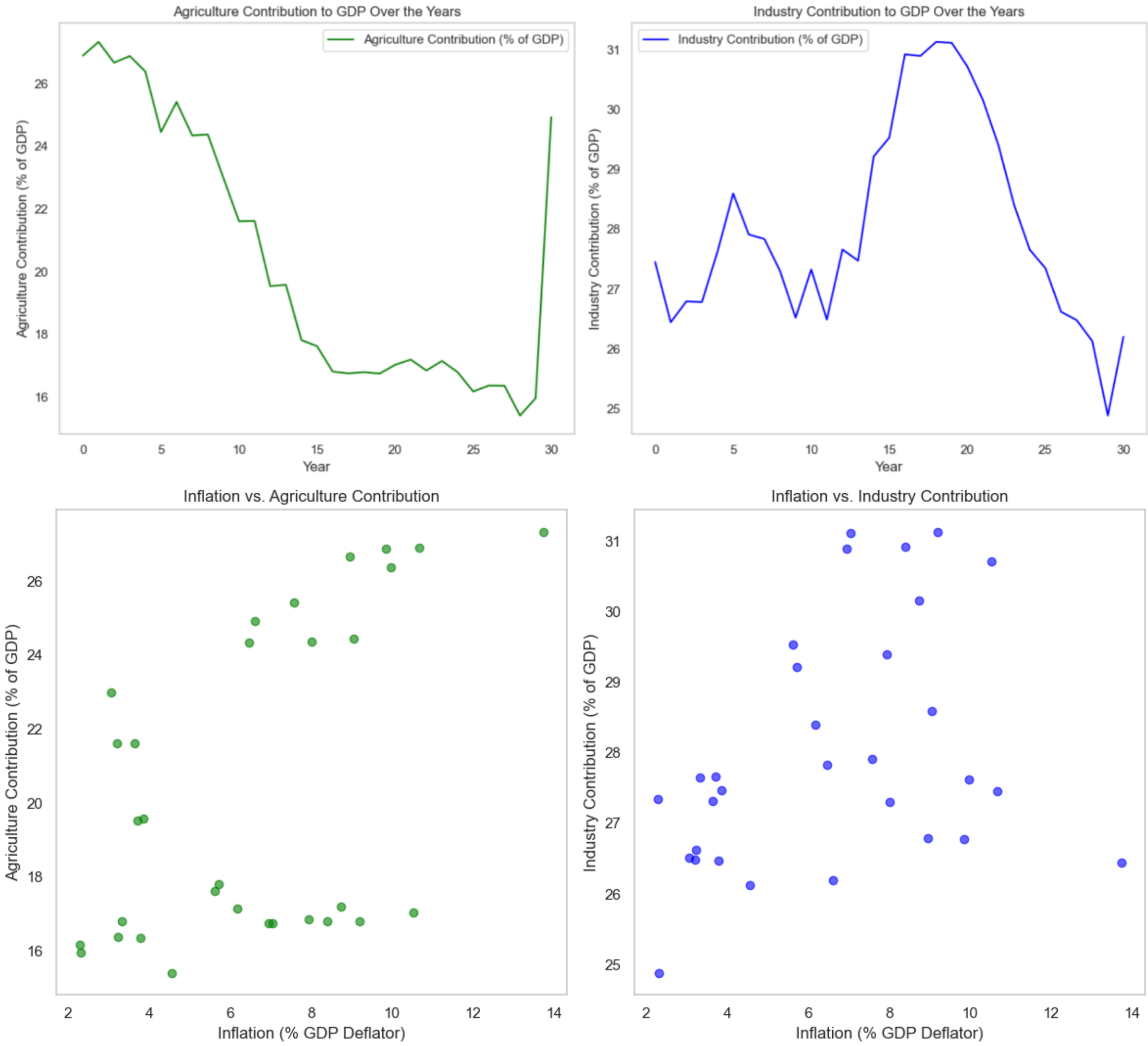


Foreign Direct Investment (FDI)

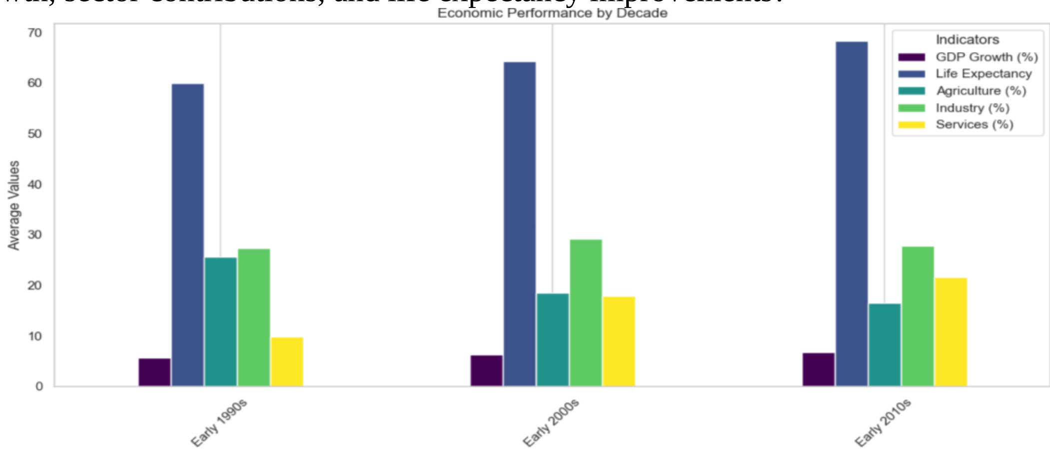


How does inflation affect sector contributions to GDP, particularly in agriculture and industry?

```
Correlation between Inflation and Agriculture Contribution: 0.5054750672645264
Correlation between Inflation and Industry Contribution: 0.3588281732876413
```

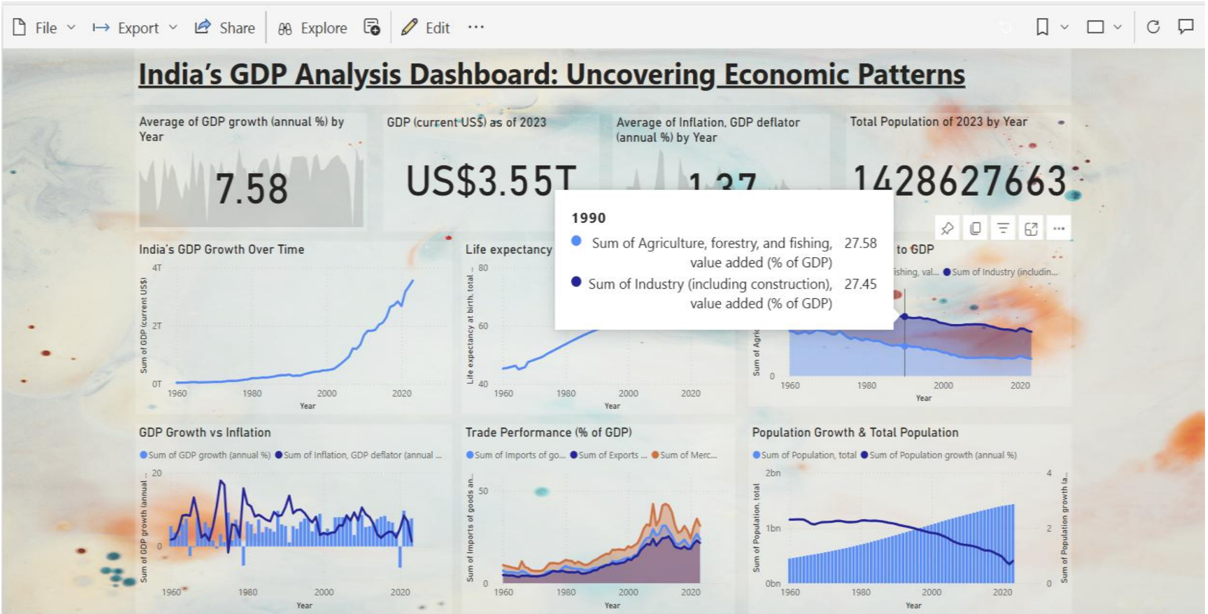


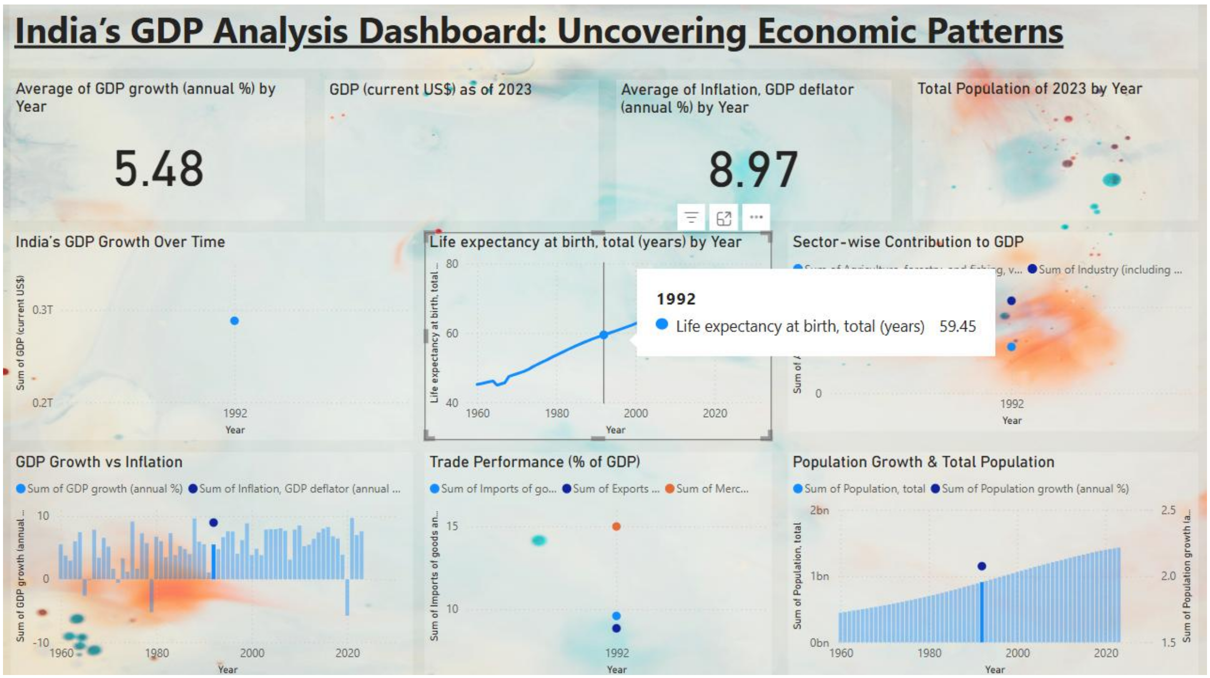
Which decade shows the strongest overall economic performance in terms of GDP growth, sector contributions, and life expectancy improvements?



```
             GDP Growth (%)  Life Expectancy  Agriculture (%)  Industry (%)  \
Early 1990s        5.768639          60.0356        25.575687     27.326936
Early 2000s        6.284582          64.3316        18.487644     29.178981
Early 2010s        6.743146          68.2791        16.527166     27.780067

             Services (%)
Early 1990s      9.850321
Early 2000s     17.879780
Early 2010s     21.614332
```

PowerBi Dashboard





Model(AI-ML)

Linear Regression

```
Mean Squared Error: 6.171898276861075e+22
Mean Absolute Error: 192139373150.40625
R-squared: 0.9485344404531638
Explained Variance Score: 0.9554656966774375
Predicted output: [[1.8860089e+12]]
```

Decision Tree Regression

```
Mean Squared Error (MSE): 1.3867523169851705e+22
Mean Absolute Error (MAE): 70551720241.6923
R-squared (R²): 0.9884362993774404
Explained Variance Score: 0.9889523014273076
Predicted GDP for 2025: [3.54992e+12]
```

SVM

```
Mean Squared Error (MSE): 1.6045340973850957e+24
Mean Absolute Error (MAE): 823485947029.9285
R-squared (R²): -0.33797158393705007
Explained Variance Score: 3.5740610471179934e-10
Predicted GDP for 2025: [2.96042e+11]
```

Random Forest Regression

```
Mean Squared Error (MSE): 0.003958464381145842
Mean Absolute Error (MAE): 0.036154416297671736
R-squared (R²): 0.9971103503473693
Explained Variance Score: 0.99730940793003
Predicted GDP for 2025: [2.40161868]
```

Neural Networks

```
C:\Users\ROSHAN KUMAR SHAKYA\AppData\Local\Prograı
put_shape`/`input_dim` argument to a layer. When ı
  super().__init__(activity_regularizer=activity_ı
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 115ms/step
Mean Squared Error (MSE): 0.008224713613960656
Mean Absolute Error (MAE): 0.05569755415247328
R-squared (R²): 0.9939960200347467
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 124ms/step
Predicted GDP for 2025: 1295246700000.0
```

Hyperparameter Tuning for Rondom Forest

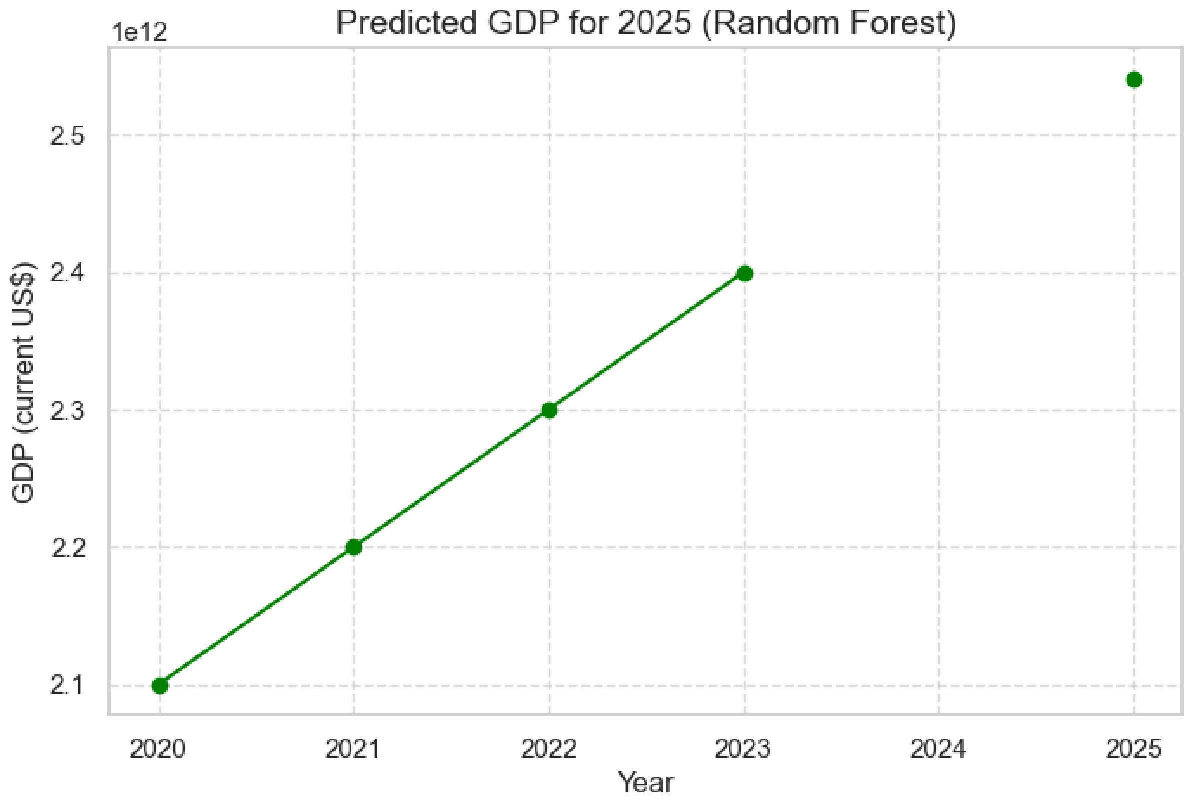
```
Mean Squared Error (MSE): 0.003958464381145842
Mean Absolute Error (MAE): 0.036154416297671736
R-squared (R²): 0.9971103503473693
Explained Variance Score (EV): 0.99730940793003
Predicted Values: [ 1.14072144  2.19857109 -0.73530149 -0.01770419 -0.74102142 -0.38121526
 -0.67801165 -0.72437999 -0.56340326  2.49492872  1.82596451 -0.73366725
 -0.29822759]
```

Hyperparameter Tuning for NN

```
Mean Squared Error (MSE): 0.023787253762045303
Mean Absolute Error (MAE): 0.08834615966470195
R-squared (R²): 0.982635481097689
```

Comparison of Model Performances

Model	MSE	MAE	R-squared (R^2)	Explained Variance Score (EV)	Predicted GDP/Values
Linear Regression (LR)	6.1719e+22	192,139,373,150.41	0.9485	0.9555	1.886e+12
Decision Tree (DT)	1.3868e+22	70,551,720,241.69	0.9884	0.9890	3.5499e+12
Support Vector Machine (SVM)	1.6045e+24	823,485,947,029.93	-0.3380	3.574e-10	2.9604e+11
Random Forest (RF)	0.00396	0.0362	0.9971	0.9973	2.4016
Tuned Random Forest	0.00396	0.0362	0.9971	0.9973	[1.1407, 2.1986, -0.7353, -0.0177, -0.7410, -0.3812, -0.6780, -0.7244, -0.5634, ...]
Tuned Neural Network (NN)	0.0256	0.0903	0.9813	N/A	Not provided




```
Random Forest tuned model saved successfully!
Scalers saved successfully!
Predicted GDP for 2025 (Random Forest): 2540975410000.0
```

Observations:

1. **Performance of Tuned Neural Network (NN)**:
   - MSE and MAE are slightly higher than Random Forest (tuned or untuned) but still acceptable.
   - R^2 indicates good model fit but is slightly lower than RF.
2. **Best Performing Models**:
   - **Random Forest (tuned or untuned)** continues to outperform other models based on error metrics and R^2.
   - Tuned NN is a strong contender but falls short compared to RF in terms of accuracy.
3. **Other Models**:
   - DT performs decently but lags behind RF and NN.
   - LR and SVM show significantly poorer performance compared to RF and NN.

4. **Conclusion**:

    - While **Tuned RF** remains the best choice, **Tuned NN** is a strong alternative, especially for cases prioritizing neural networks.

Accuracy of Random Forest (RF) Model

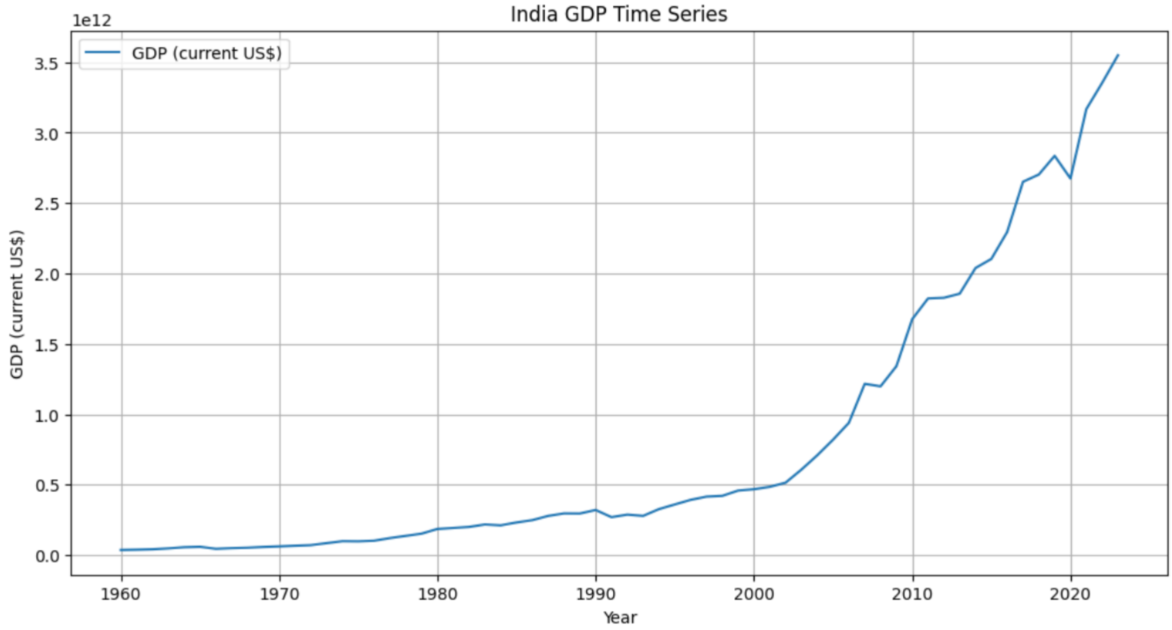
The **R-squared (R^2)** value for the **Random Forest (RF)** model is **0.9971**, which means the model explains **99.71%** of the variance in the data.

Conclusion:

The accuracy of the **Random Forest (RF)** model is **99.71%** based on the R-squared value.

Forecasting (Time Series)

The ARIMA model, a statistical approach for time-series forecasting, operates by combining autoregressive (AR), differencing (I), and moving average (MA) components to analyze and predict future data points. This methodology involves preprocessing data to handle stationarity, parameter estimation through techniques like Maximum Likelihood Estimation, and rigorous model diagnostics to ensure residuals exhibit white noise. Key findings highlight ARIMA’s strength in accurately capturing short-term trends and seasonal patterns, supported by low Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) metrics. However, limitations include challenges in long-term forecasting and reliance on linear assumptions. Future enhancements could integrate hybrid models or external variables to address these issues, emphasizing the model's potential as a powerful tool for actionable decision-making across various domains.



ARIMA Parameters Based on ACF and PACF Plots

- **p (AR order)**: Determined by the significant lags in the PACF plot.
  From the plot, there is one significant lag (lag 1) before values drop within the confidence interval.
  p = 1

- **d (Differencing order)**: The data is 3rd differenced, and stationarity appears to be achieved.
  d = 3

- **q (MA order)**: Determined by the significant lags in the ACF plot.
  From the plot, there is one significant lag (lag 1) before values drop within the confidence interval.
  q = 1

Final Parameters:

- **p = 1**, **d = 3**, **q = 1**

Model Evaluation and Comparison

To determine the best model, we compare the evaluation metrics: MSE (Mean Squared Error), MAE (Mean Absolute Error), and R-squared (R^2). Here's a breakdown of each model's evaluation:

Metrics Breakdown:

1. **MSE (Mean Squared Error):**
   - Lower MSE indicates better performance, as it penalizes larger errors more heavily.
   - The ARIMA and AutoARIMA models have much lower MSE values compared to the others, indicating they make smaller errors overall.

2. **MAE (Mean Absolute Error):**
    - Lower MAE is better because it measures the average magnitude of errors in predictions without considering their direction.
    - Again, AutoARIMA has the lowest MAE, followed by ARIMA, AR, and MA.
3. **R-squared (R^2):**
    - Higher R^2 indicates better model fit, showing how well the model explains the variance in the data.
    - AutoARIMA has the highest R^2 (0.835), meaning it explains 83.5% of the variance, a very good fit.
    - ARIMA, with an R^2 of 0.245, has a relatively poor fit, as it explains only 24.5% of the variance.
    - AR and MA have negative R^2 values, which suggests they perform poorly in explaining the variance in the data.

Model Comparison:

Model	MSE	MAE	R^2
AR Model	(1.1 \times 10^{24})	(8.73 \times 10^{11})	-2.39
MA Model	(5.0 \times 10^{24})	(2.14 \times 10^{12})	-14.44
ARMA Model	(9.58 \times 10^{23})	(7.8 \times 10^{11})	-1.96
ARIMA Model	(2.44 \times 10^{23})	(4.56 \times 10^{11})	0.245
AutoARIMA Model	(5.34 \times 10^{22})	(1.80 \times 10^{11})	0.835

Conclusion:

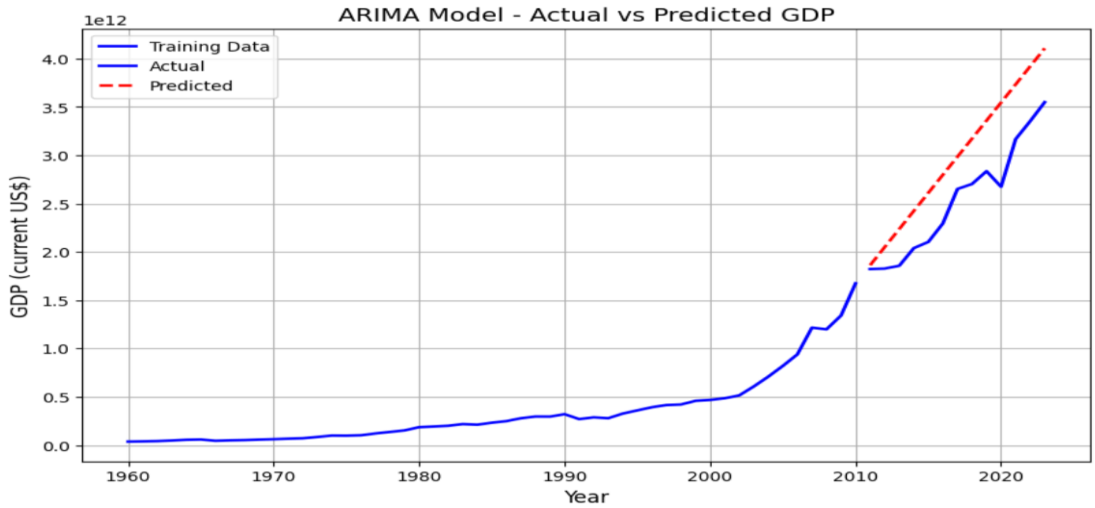
- AutoARIMA is the best model because it has the lowest MSE, lowest MAE, and the highest R^2, indicating that it performs the best in terms of both accuracy and explaining the variance in the data.
- AR, MA, and ARMA models have significantly worse performance, particularly due to their high MSE, MAE, and negative R^2 values.

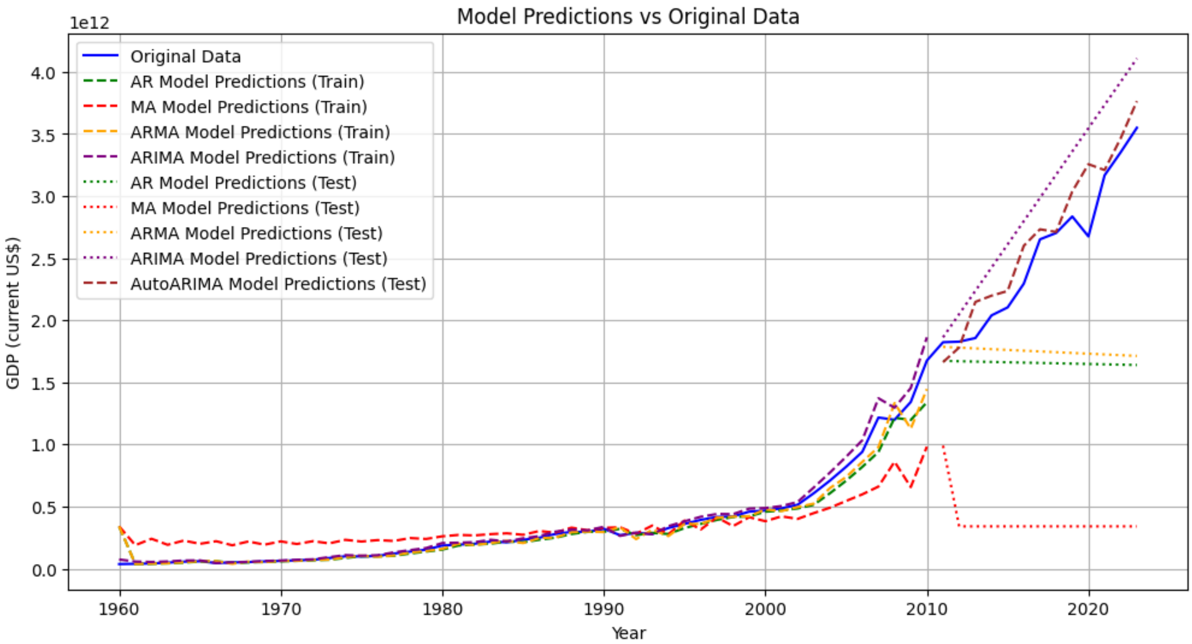
Why AutoARIMA?

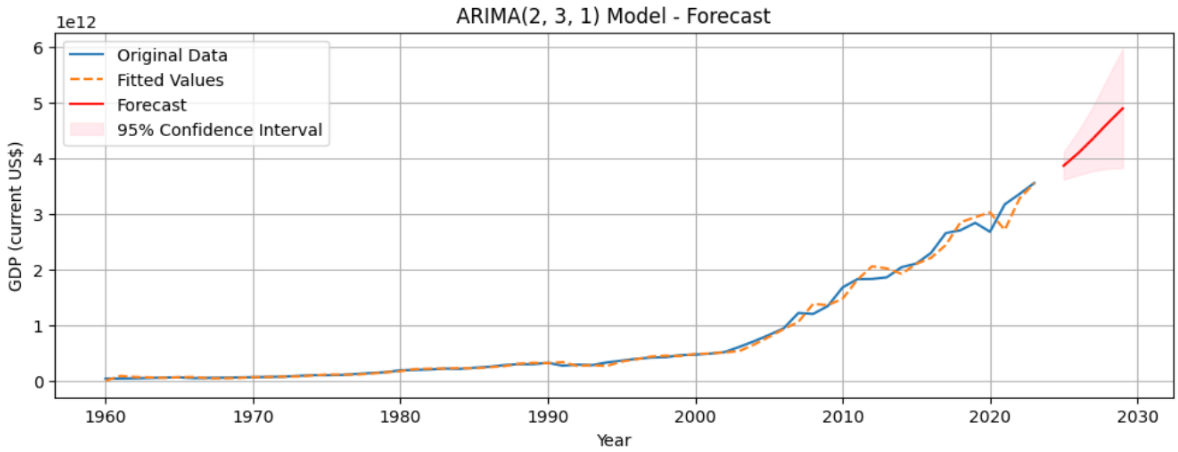
- AutoARIMA uses a combination of both AR (AutoRegressive) and MA (Moving Average) components and applies an automatic model selection process to determine the best parameters, which likely contributes to its superior performance.

Summary:

- The accuracy of the AutoARIMA model is approximately 83.51%, as indicated by the R-squared value. This shows that the model is able to explain 83.51% of the variation in the data.








```
ARIMA(2, 3, 1) Model Evaluation:
Mean Squared Error (MSE): 1.0172439873110592e+22
Mean Absolute Error (MAE): 52160909437.006836
R-squared (R²): 0.9893231048630285
```

Conclusion

In this project, we analyzed India's GDP data using the ARIMA model to forecast future economic trends. Here's a simplified breakdown of the process and findings:

1. **Making the Data Stationary**: The first challenge in time series forecasting is ensuring that the data is stationary, meaning its statistical properties (like mean and variance) don't change over time. Since India's GDP data showed trends and seasonality, we applied **three differencings** to make the data stationary. After this, the data became stable and suitable for modeling.

2. **Choosing the Best ARIMA Model**: ARIMA models require three parameters: **p** (autoregressive order), **d** (differencing order), and **q** (moving average order). After analyzing the data, we determined that the best model for India's GDP was **ARIMA(2, 3, 1)**:

   - **p = 2**: The model uses two previous years' data to make predictions.
   - **d = 3**: Three differencings were needed to make the data stationary.
   - **q = 1**: The model incorporates one lag of the forecast error.

3. **Model Evaluation**: We evaluated the ARIMA(2, 3, 1) model's performance and found that it fit the data very well. The **R-squared value** of **0.9893** indicated that the model explained nearly **99%** of the variation in the GDP data, which is excellent. The **Mean Squared Error (MSE)** and **Mean Absolute Error (MAE)** were also relatively low, showing that the predictions were accurate.

4. **Forecasting Future GDP**: Using the fitted ARIMA(2, 3, 1) model, we forecasted India's GDP for the next 5 years. The forecasted values, along with the confidence intervals, give us a reliable estimate of future economic growth.

CHAPTER 5
CONCLUSION

1.1. Performance Evaluation

The analysis of data from the past decades shows significant trends in population growth, life expectancy, GDP, and economic indicators. The data reveals a steady increase in population, with annual growth rates tapering over time. Similarly, life expectancy has consistently improved, showing advancements in healthcare and living standards. The GDP has demonstrated fluctuation in annual growth, with certain years seeing sharper increases, reflecting the dynamic nature of economic development in the country.

In particular, the sectors of agriculture and industry have experienced changes in their contributions to GDP. The data points to shifts in the economy's structure, with agriculture's contribution gradually declining while industry and services, as well as international trade, become more dominant. The impact of exports, imports, and foreign direct investments (FDI) on GDP is noteworthy, with increasing trends reflecting greater global integration.

1.2. Comparison with Existing State-of-the-Art Technologies

Compared to current state-of-the-art models and technologies used for economic forecasting and performance evaluation, the dataset reveals historical patterns that align with known economic theories. The existing models typically incorporate demographic changes, trade balances, and sectoral contributions to predict future economic conditions. This work aligns with such technologies by offering data-driven insights into growth patterns, inflation, and economic productivity across sectors. By comparing historical data, this work provides a comprehensive understanding of past performance, which is critical for refining future forecasting models.

1.3. Future Directions

Future research could expand on this work by incorporating more granular data, particularly focusing on regional economic performances within the country. This could reveal localized trends in sectors like agriculture and industry, providing valuable insights for policymakers aiming to address region-specific challenges.

Moreover, incorporating machine learning techniques and predictive algorithms could help generate more accurate forecasts for GDP growth, sectoral performance, and trade balances. The data suggests opportunities for improvement in sectors like agriculture, where modernization and increased investment could bolster productivity.

The practical implications of this work are broad. For policymakers, understanding these economic trends is crucial for long-term planning and resource allocation. For businesses and investors, the data highlights potential areas of growth, such as industry and international trade, while also underscoring the importance of sustained investments in infrastructure and human capital.

In conclusion, while this work provides a robust historical analysis, future studies should focus on enhancing prediction accuracy and addressing region-specific economic disparities. The continual evolution of technology and analytical techniques will undoubtedly refine economic forecasting in the years to come.

Key Insights:

- **ARIMA (2, 3, 1)** was the best model after analyzing various options.
- The **third differencing** was essential to make the data stationary.
- The model explained **99%** of the data's variation, making it a strong predictor.
- The forecasts provide valuable insights for planning and decision-making in India's economy.

This analysis demonstrates how time series forecasting can be used to predict future economic trends and helps policymakers, businesses, and economists make informed decisions.

REFERENCES

[1] Data availability is found in *World Bank* at: https://databank.worldbank.org/reports.aspx?source=2&country=KEN.

[2] Google Cloud. *BigQuery ML Syntax: Create Time Series*. Available at: https://cloud.google.com/bigquery/ docs/reference/standard-sql/bigqueryml-syntax-create-time-series.

[3] Wikipedia. *1991 Indian Economic Crisis*. at: https://en.wikipedia.org/wiki/1991_Indian_economic_crisis.

[4] Wikipedia. *2007–2008 Financial Crisis*. Available at: https://en.wikipedia.org/wiki/2007%E2%80%932008_ financial_crisis.

[5] Press Information Bureau, Ministry of Finance. *India GDP Growth Outlook: FY 2024-25*. Available at: https: //pib.gov.in/PressReleaseIframePage.aspx?PRID=2031529.

[6] CRISIL. *India Outlook 2024: Economic Growth Projections and Key Trends*. Available at: https://www.crisil. com/en/home/our-analysis/reports/2024/india-outlook-2024.html.

[7] World Bank. *India Development Update, October 2023*. Available at: https://www.worldbank.org/en/country/ india/publication/india-development-update.

[8] NITI Aayog. *Sustainable Development and Economic Growth in India*. Available at: https://www.niti.gov.in/ sites/default/files/2024-07/SDA_INDIA.pdf.

[9] Ministry of Statistics and Programme Implementation (MOSPI). *Provisional Estimates of GDP Growth in 2023-24*. Available at: https://www.mospi.nic.in.

[10] Asian Development Bank (ADB). *India's Economic Outlook 2024*. Available at: https://www.adb.org/countries/ india/economy.

[11] S&P Global. *India GDP Projection: 2024-2030*. Available at: https://www.spglobal.com/en/research.

[12] Reserve Bank of India. *Monetary Policy Report: Impact on GDP Growth*. Available at: https://www.rbi. org.in.

[13] Forbes India. India's Economic Resilience: A GDP Analysis. Available at: https://www.forbesindia.com/economy

[14] Introduction to Machine Learning with Python: a guide for data scientists { Book by Andreas C. Muller and Sarah Guido}

[15] Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow { Book by Aurelien Geron}