Project Report

On

RECESSION ANALYSIS

Submitted

In partial fulfilment

For the award of the Degree of

PG-Diploma in Big Data Analytics

(PG-DBDA)

C-DAC, ACTS (Pune)

Guided By:

Mr. Saurabh Panwar

Submitted By:

Bisure Prateek Pratapsinh (230940125010)

Kodak Rishikesh Pradeep (230940125027)

Munfan Aditya Ramrao (230940125029)

Mali Omkar Nandkumar (230940125032)

Tambolkar Shubham Dilip (230940125051)

Centre for Development of Advanced Computing(C-DAC), ACTS

(Pune- 411008)

CERTIFICATE

TO WHOMSOEVER IT MAY CONCERN

This is to certify that

Bisure Prateek Pratapsinh	(230940125010)
Kodak Rishikesh Pradeep	(230940125027)
Munfan Aditya Ramrao	(230940125029)
Mali Omkar Nandkumar	(230940125032)
Tambolkar Shubham Dilip	(230940125051)

Have successfully completed their project on

RECESSION ANALYSIS

Under the guidance of Mr. Saurabh Panwar Sir

Acknowledgement

This is to acknowledge our indebtedness to our Project Guide, **Mr. Saurabh Panwar** for his constant guidance and helpful suggestion for preparing this project **Recession Analysis.** We express our deep gratitude towards him for inspiration, personal involvement, constructive criticism that she provided us along with technical guidance during the course of this project.

We take this opportunity to thank Head of the department **Mr. Gaur Sunder** for providing us such a great infrastructure and environment for our overall development.

We express sincere thanks to **Mrs. Namrata Ailawar**, Process Owner, for their kind cooperation and extendible support towards the completion of our project.

It is our great pleasure in expressing sincere and deep gratitude towards **Mrs. Risha P R** (Program Head) and **Mrs. Srujana Bhamidi** (Course Coordinator, PG-DBDA) for their valuable guidance and constant support throughout this work and help to pursue additional studies.

Also, our warm thanks to **C-DAC ACTS Pune**, which provided us this opportunity to carry out, this prestigious Project and enhance our learning in various technical fields.

Bisure Prateek Pratapsinh (230940125010)

Kodak Rishikesh Pradeep (230940125027)

Munfan Aditya Ramrao (230940125029)

Mali Omkar Nandkumar (230940125032)

Tambolkar Shubham Dilip (230940125051)

ABSTRACT

A recession is a significant, widespread and prolonged downturn in economic activity. It is usually characterized by negative GDP growth, rising unemployment, falling business profits and declining stock prices. Recessions are a normal part of the business cycle, but they can be very painful. By understanding the causes and consequences of recessions, we can better prepare for them and mitigate their effects.

Analyzing a recession involves a comprehensive examination of various economic indicators and trends. The non-profit organization known as National Bureau of Economic Research (NBER) officially determines recessions based on a broader set of indicators including GDP, employment, income and consumer spending.

Two consecutive quarters of negative GDP growth is the most widely used indicator for recession analysis. For the task of Recession analysis, we found datasets of the yearly GDP growth of top 5 countries in the world by GDP. With these datasets we analyzed the Recession and GDP growth over the period. Using Python and Machine Learning we tried to calculate GDP growth over time.

Table of Contents

1. INTRODUCTION

Gross Domestic Product (GDP) and Recessions are key components in assessing the economic health of a country.

1.1 Gross Domestic Product (GDP):

GDP is the total value of all goods and services produced within a country's borders over a specific period, usually measured annually or quarterly.[1]

Components:

Consumption (C): Expenditures by households on goods and services.

Investment (I): Spending on capital goods, such as machinery and buildings, and changes in business inventories.

Government Spending (G): Expenditures by the government on goods and services.

Net Exports (Exports - Imports): The value of a country's exports minus its imports.

1.2 Recession:

A recession is a significant decline in economic activity across the economy, lasting for an extended period, typically marked by a contraction in GDP, rising unemployment, and decreased consumer and business spending.[2]

Indicators of a Recession:

Negative GDP Growth: If the GDP contracts for two consecutive quarters or more, it is often considered a technical recession.

Unemployment Rate: A rise in unemployment indicates a slowdown in economic activity.

Consumer and Business Confidence: A decline in confidence can lead to reduced spending and investment.

Stock Market Performance: A bear market, with declining stock prices, can signal economic downturns.

Recession Causes and Impacts:

Causes: Recessions can be triggered by various factors such as a financial crisis, high levels of debt, external shocks, or a combination of these.

Impacts: Recessions lead to reduced income, increased unemployment, lower consumer spending, decreased business investment, and often a decline in international trade.

Policy Responses:

Monetary Policy: Central banks may lower interest rates to stimulate borrowing and spending.

Fiscal Policy: Governments may implement stimulus packages, tax cuts, or increased public spending to boost economic activity.

Long-Term Economic Health:

GDP Growth Rate: Positive and sustained GDP growth is essential for long-term economic health.

Structural Reforms: Governments may implement reforms to improve productivity, enhance competitiveness, and address underlying economic issues.

In summary, GDP is a crucial metric for assessing a country's economic performance, and recession analysis involves examining various indicators to identify and understand periods of economic downturn. Effective policy responses are essential for mitigating the impacts of recessions and fostering long-term economic stability and growth.

Dataset & Features

Data used in this project is collected in the different csv and excel files and then combined to form a raw dataset which is used for the further processing and evaluation.[3]

Raw Dataset contains 5 columns:

- Country
- Date
- GDP (US Billion $)
- GDP per Capita (US $)
- Annual % Change

There were total 1525 records in the raw dataset.

2. LITERATURE REVIEW

Shelley G. L. and Wallace F. H (2004) [4] studied the relation between M1 money, real GDP and inflation in Mexico for the period 1944 to 1991. Co-integration relationships existed between the inflation and money hence study suggested that reduction in money growth may have been resulted in reduction of inflation in Mexico. The variations in inflation were divided into two components namely predictable and unpredictable components. In differenced inflation, predictable increases resulted in having a negative effect on GDP growth. Unpredictable increases resulted in having a positive effect on real GDP growth. This paper considered only money as an economical factor and neglected other important indicators of economic health.

The paper titled “GDP Prediction by Support Vector Machine Trained with Genetic Algorithm” was published in 2010 by Gang Long [5]. In this study, a support vector machine trained with a genetic algorithm is applied in GDP forecasting. Author concluded that the genetic algorithm can get optimal solutions in a short time, which is an excellent method in parameters selection of support vector machine. Then, a genetic algorithm is introduced to simultaneously optimize the SVM parameters. The GDP data from 1989 to 2002 used for training and 2003 to 2007 for testing. But the limitation of this project is that other Machine Learning algorithms can achieve better performance than SVM.

Another paper titled “Predicting Gross Domestic Product Using Autoregressive Models” published in 2017 by J. Roush, K. Siopes and G. Hu [6]. They used autoregressive models and then constructed a vector autoregressive model to predict GDP. The predicted result matches historical GDP data and predicts consistent future growth. Restriction of this approach is that it fails to overcome historic economic recession. This paper also didn’t account for the other parameters such as trade, economic, geographical to predict the GDP growth.

Martin Schneider, Martin Spitzer [7] developed a framework for short-term forecasting of real GDP for Austria using the generalized dynamic factor model.

3. SYSTEM REQUIREMENTS

Hardware Requirements :

- Platform – Windows
- RAM – 8 GB
- Peripheral Devices – Mouse, Keyboard, Monitor
- A Network Connection

Software Requirements :

- Anaconda
- Google Colab
- Python3
- Tableau

4. FUNCTIONAL REQUIREMENTS

Anaconda :

- Anaconda Navigator is a desktop graphical user interface (GUI) included in Anaconda® Distribution that allows you to launch applications and manage conda packages, environments, and channels without using command line interface (CLI) commands.

- Navigator can search for packages on Anaconda.org or in a local Anaconda Repository. It is available for Windows, macOS, and Linux.

- Anaconda supports multiple data science/machine learning libraries such as Pandas, NumPy, Scikit-learn, Matplotlib, etc. to perform various tasks such as statistical analysis, data mining, data exploration, feature engineering, and modelling.

Python3 :

- Python is a general purpose and high-level programming language.[8]

- It is used for developing desktop GUI applications, websites and web apps.
- Python allows to focus on core functionality of the application by taking care of common programming tasks.
- Python is derived from many other languages, including ABC, Modula-3, C, C++, Algol-68, Small Talk, and Unix shell and other scripting languages.

Tableau :

- Data visualization is the graphical representation of information and data.

- It helps create interactive elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.[9]

- Tableau is widely used for Business Intelligence but is not limited to it.

- It helps create interactive graphs and charts in the form of dashboards and worksheets to gain business insights.
- All of this is made possible with gestures as simple as drag and drop.

5. DATA PREPROCESSING

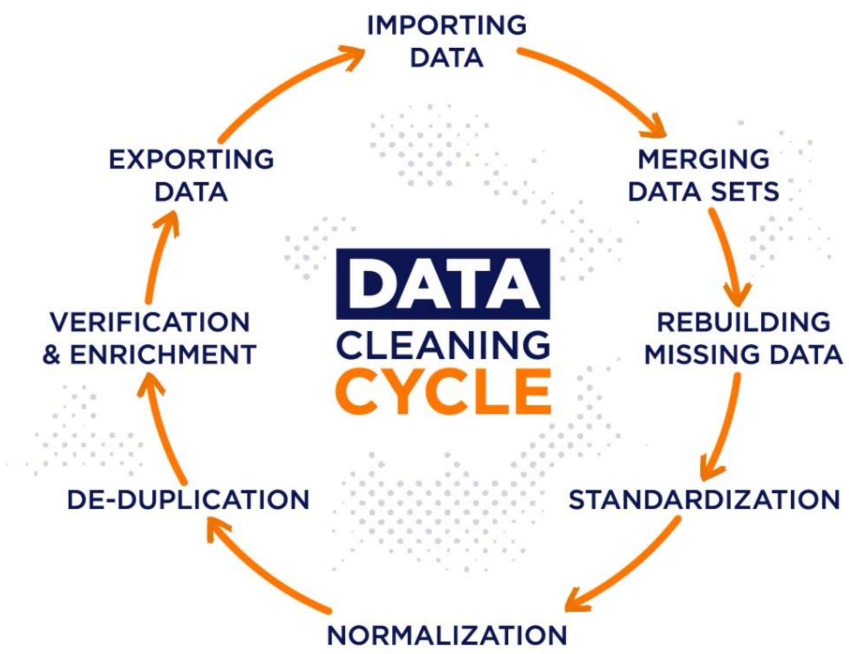


Fig : Data Cleaning Process

Data cleansing or data cleaning is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table or database and refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the data and then replacing, modifying or deleting the dirty or coarse data. Data cleansing may be performed interactively with data wrangling tools or as batch processing through scripting.

After cleansing, a data set should be consistent with other similar data sets in the system. The inconsistencies detected or removed may have been originally caused by user entry errors, by corruption in transmission or storage, or by different data dictionary definitions of similar entities in different stores. Data cleaning differs from data validation in that validation almost invariably means data is rejected from the system at entry and is performed at the time of entry, rather than on batches of data.

Here in this project after collecting the data we pre-processed the data using Python. We dropped the null/irrelevant values present in the dataset and so that the final dataset created which is used for further operations.

6. SYSTEM ARCHITECTURE

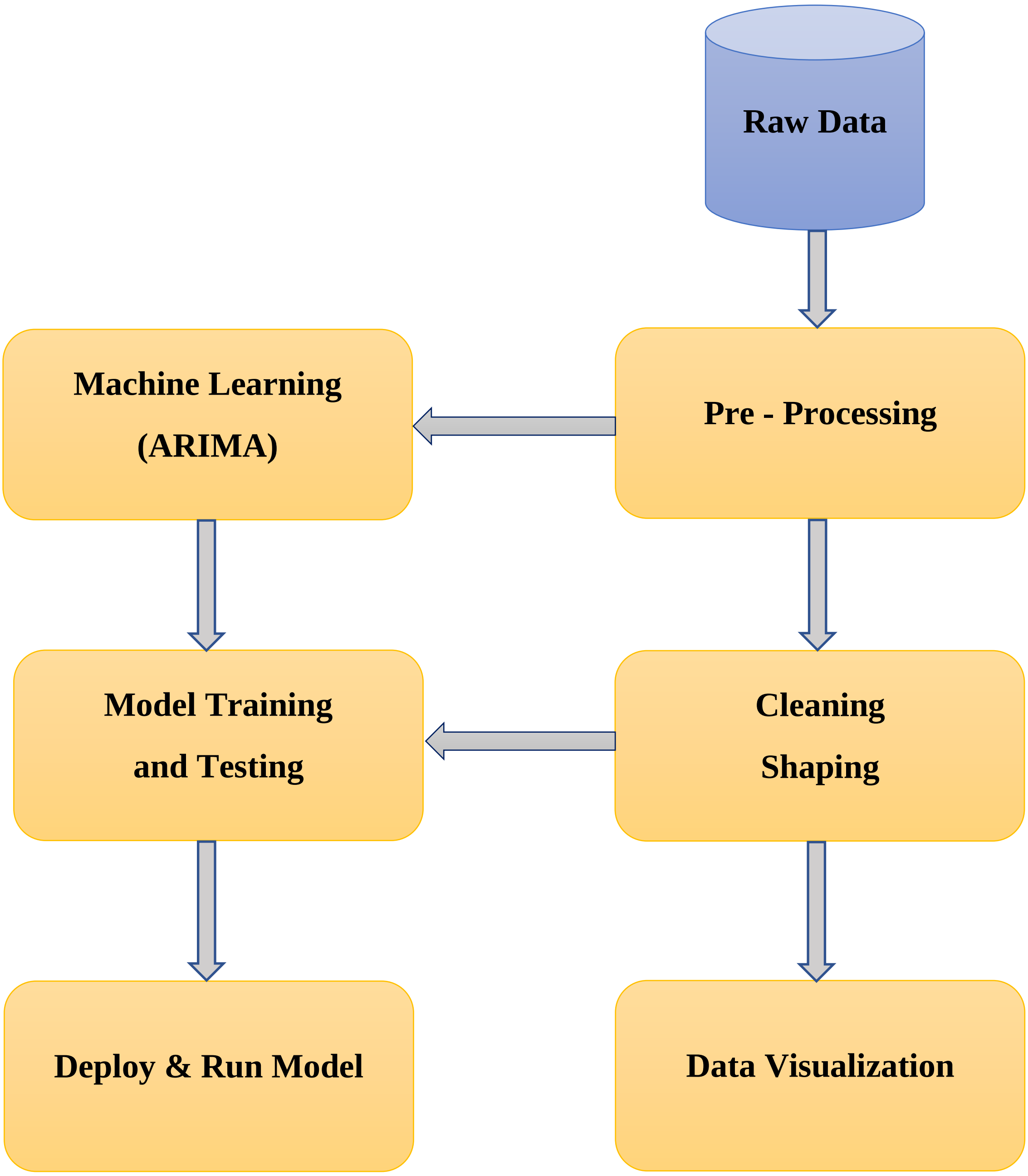


Fig : System Architecture of Recession Analysis Project

7. METHODOLOGY

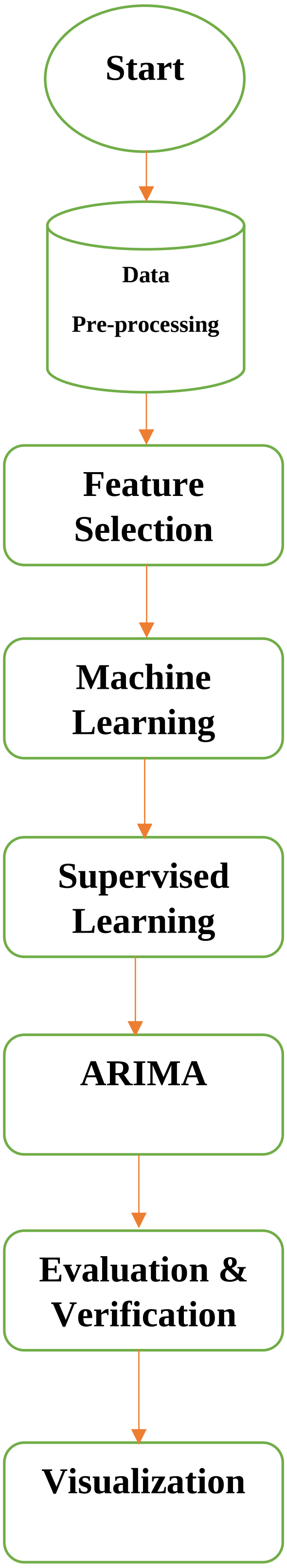


Fig : Methodology of Recession Analysis

8. DATA VISUALIZATION

Creating a dynamic dashboard in Tableau involved several steps.

Step 1: Loading the Data

We loaded the dataset with the necessary information, including year wise GDP, GDP growth rate, and other relevant data for the top 5 countries.

Step 2: Connected Tableau to Dataset

1. Opened Tableau Desktop.

2. Connected the dataset by selecting the appropriate data source.

Step 3: Created Sheets for Year wise GDP and Growth Rate

For Year wise GDP:

1. Dragged the "Year" field to Columns shelf.

2. Dragged the "GDP" & "Country" field to Rows shelf.

3. Selected the appropriate visualization type (line chart).

For Year wise GDP Growth Rate:

1. Dragged the "Year" field to Columns shelf.

3. Dragged "Annual % Change" field to Rows shelf.

4. Selected the appropriate visualization type (line chart).

Step 4: Created a Box Plot for Annual % Change in GDP

1. Drag the "Country" field to Columns shelf.

2. Dragged "Annual % Change" field to Rows shelf.

3. Changed the chart type to "Box Plot" from the "Show Me" menu.

Step 5: Built a Dashboard

1. Clicked on the "New Dashboard" tab.
2. Dragged the World Map Sheet, Yearwise GDP sheet, Yearwise GDP Growth Rate sheet, and Box Plot sheet onto the dashboard.
3. Arranged them as per your preference.

Step 7: Added Interactivity to the Dashboard

1. Used the filter to make the dashboard interactive.
2. Added tooltips to display additional information when users hover over data points.

Step 8: Fine-tune Formatting

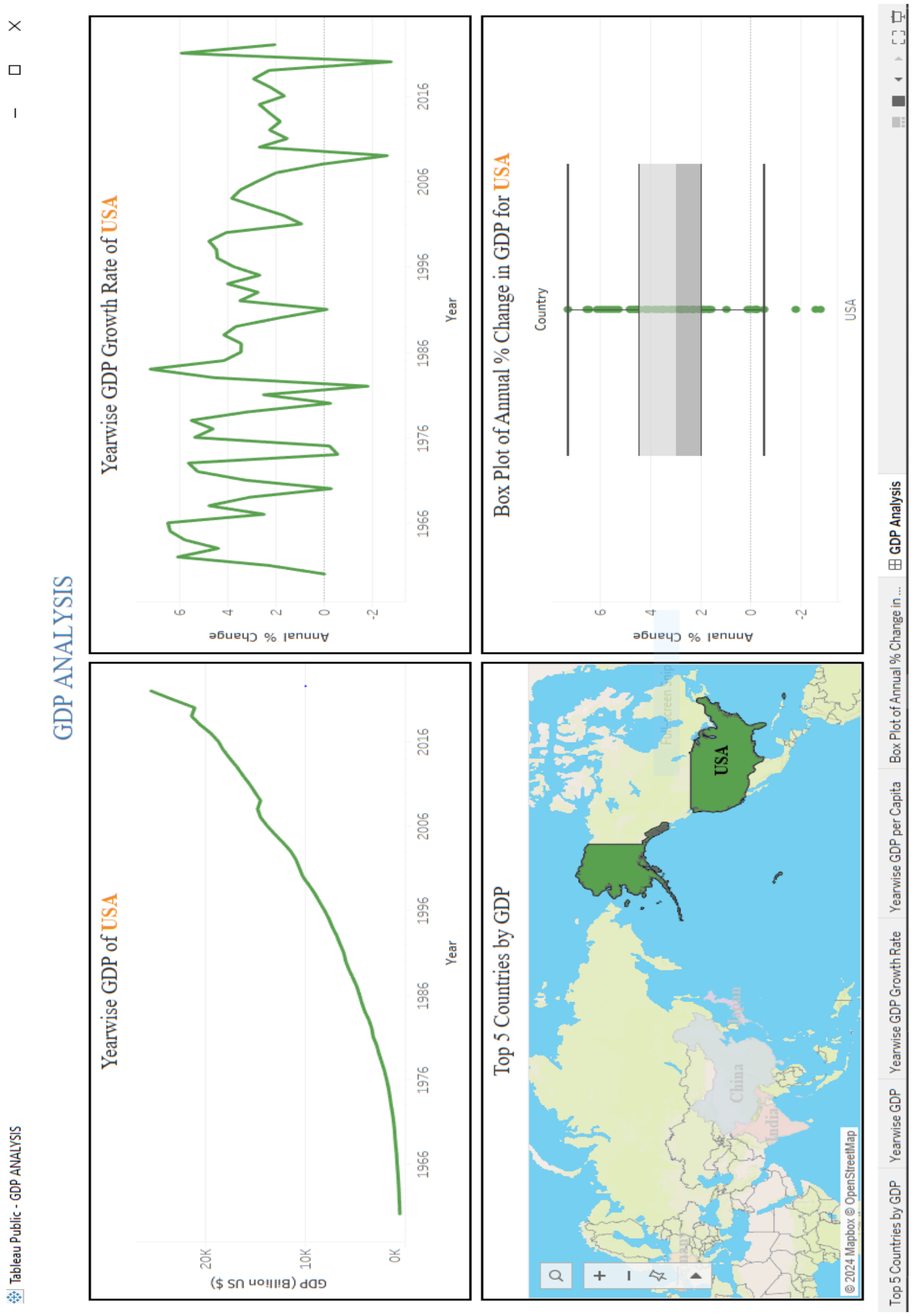
1. Adjusted the axis labels, titles, and legends for clarity.
2. Customized colours and styles to enhance the visual appeal.

Step 9: Tested and Refined

1. Interacted with the dashboard to ensure that it provides the desired insights.
2. Made necessary adjustments based on own observations.

8.1 Dynamic Dashboard Snapshots –

1. USA



2. China

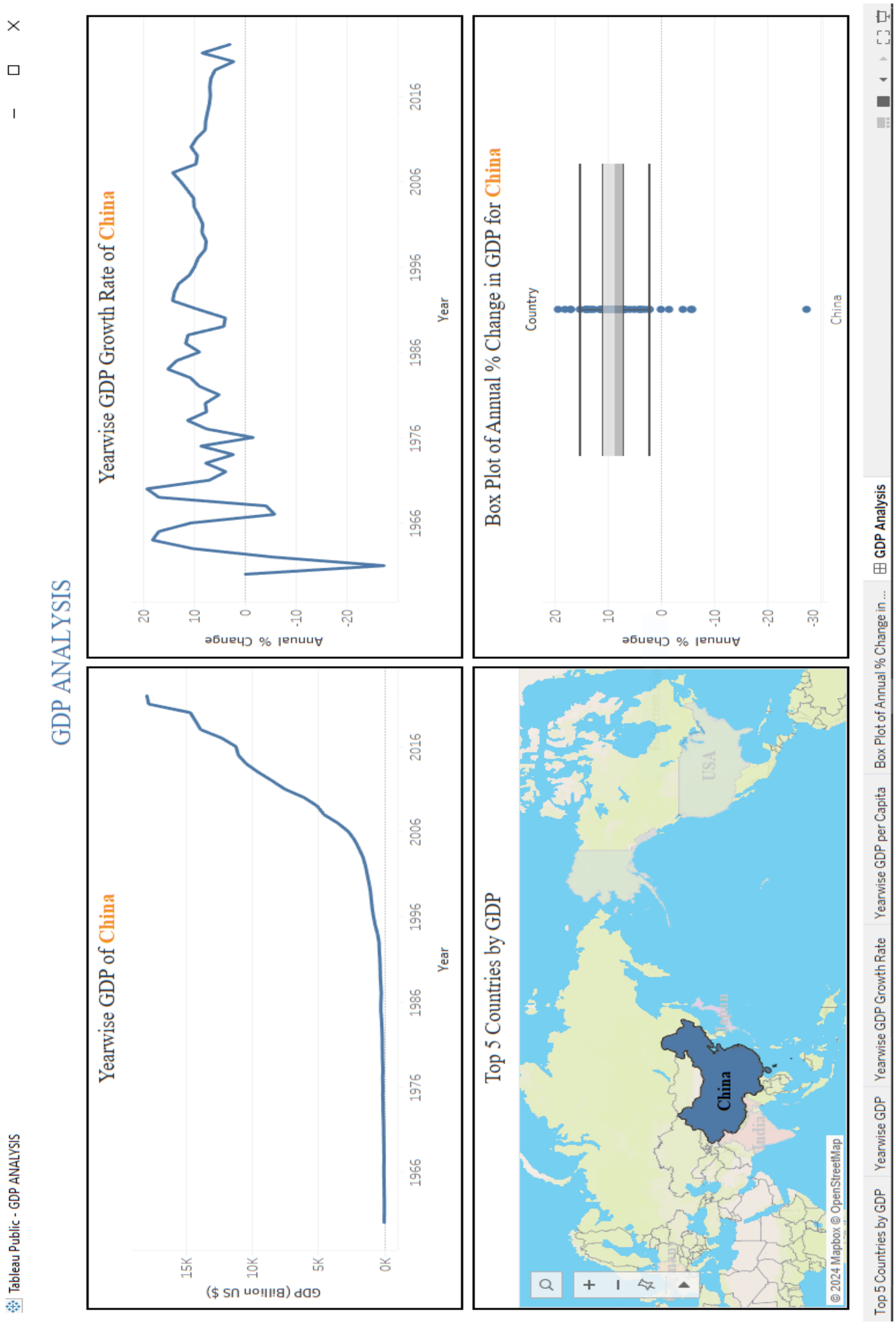


Tableau Public - GDP ANALYSIS
GDP ANALYSIS
Yearwise GDP of China
GDP (Billion US $)
Year
Yearwise GDP Growth Rate of China
Annual % Change
Year
Top 5 Countries by GDP
China
USA
© 2024 Mapbox © OpenStreetMap
Box Plot of Annual % Change in GDP for China
Country
China
Annual % Change
Top 5 Countries by GDP
Yearwise GDP
Yearwise GDP Growth Rate
Yearwise GDP per Capita
Box Plot of Annual % Change in...
GDP Analysis

3. Japan

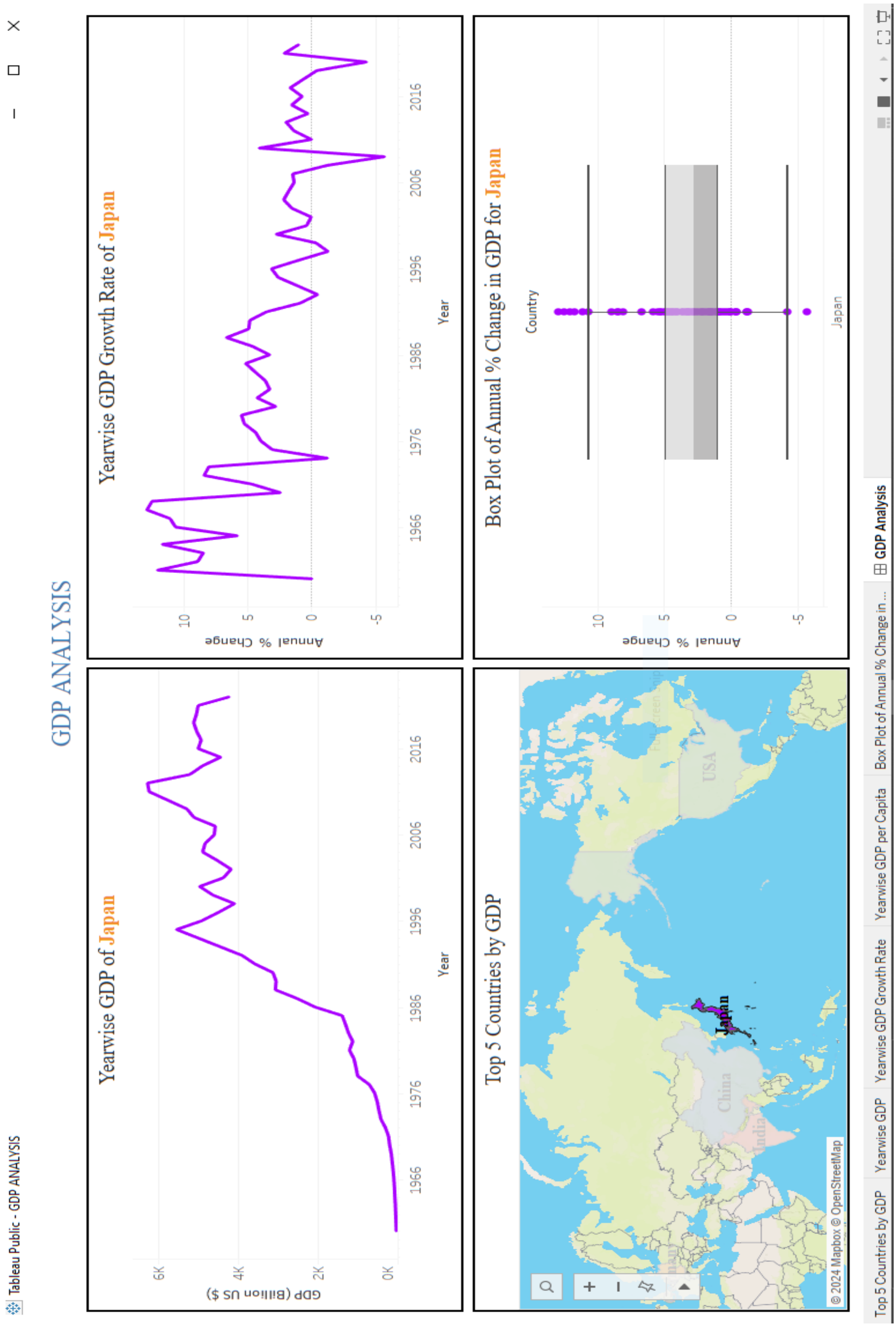


Tableau Public - GDP ANALYSIS
GDP ANALYSIS
Yearwise GDP of Japan
GDP (Billion US $)
Year
Yearwise GDP Growth Rate of Japan
Annual % Change
Top 5 Countries by GDP
Box Plot of Annual % Change in GDP for Japan
Country
Japan
© 2024 Mapbox © OpenStreetMap
Top 5 Countries by GDP
Yearwise GDP
Yearwise GDP Growth Rate
Yearwise GDP per Capita
Box Plot of Annual % Change in ...
GDP Analysis

4. Germany

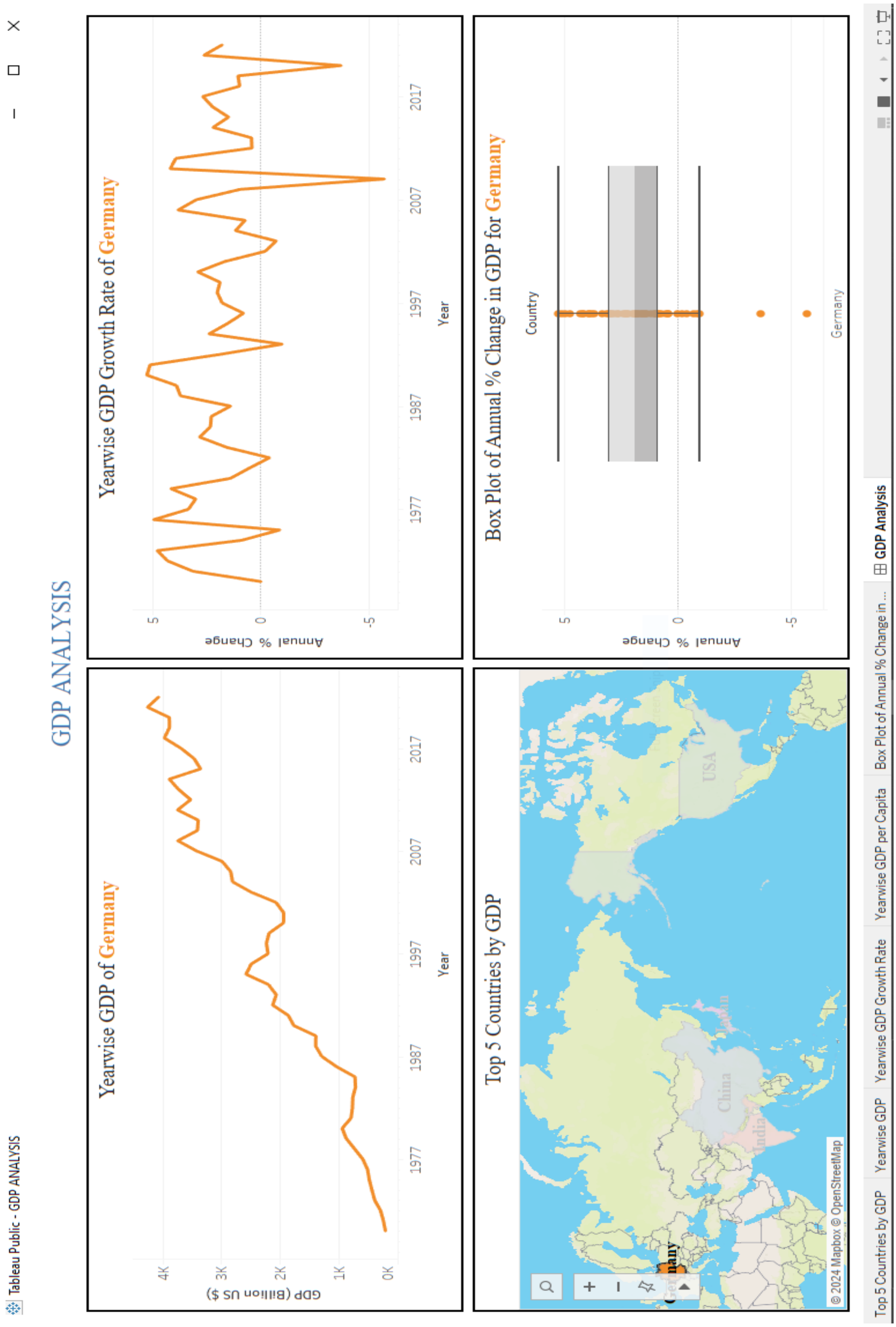


Tableau Public - GDP ANALYSIS
GDP ANALYSIS
Yearwise GDP of Germany
GDP (Billion US $)
Year
Yearwise GDP Growth Rate of Germany
Annual % Change
Top 5 Countries by GDP
Box Plot of Annual % Change in GDP for Germany
Country
Germany
USA
China
India
© 2024 Mapbox © OpenStreetMap
Top 5 Countries by GDP
Yearwise GDP
Yearwise GDP Growth Rate
Yearwise GDP per Capita
Box Plot of Annual % Change in ...
GDP Analysis

5. India

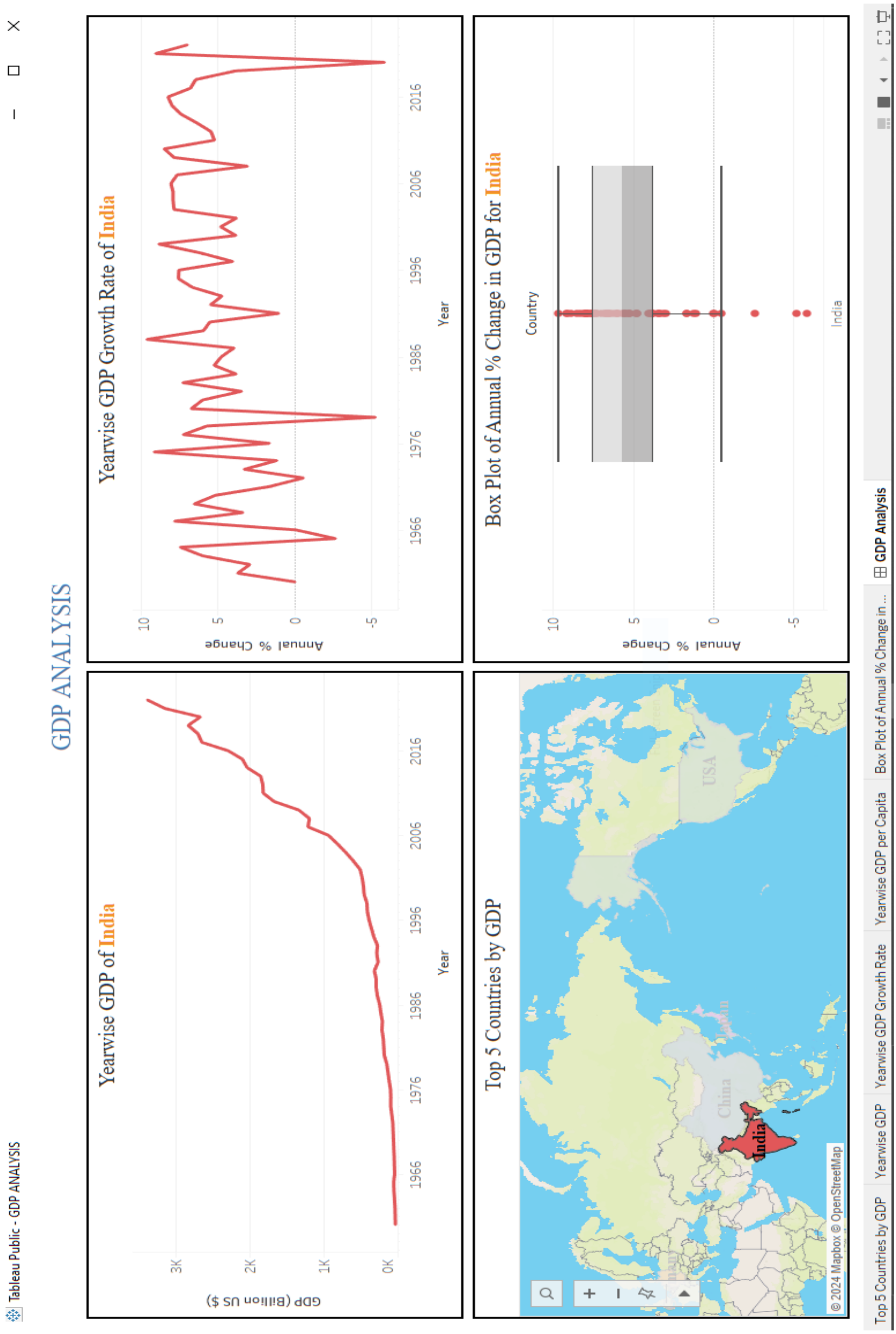


Tableau Public - GDP ANALYSIS
GDP ANALYSIS
Yearwise GDP of India
GDP (Billion US $)
3K
2K
1K
0K
1966
1976
1986
1996
2006
2016
Year
Yearwise GDP Growth Rate of India
Annual % Change
10
5
0
-5
Top 5 Countries by GDP
USA
China
India
© 2024 Mapbox © OpenStreetMap
Box Plot of Annual % Change in GDP for India
Country
India
Top 5 Countries by GDP
Yearwise GDP
Yearwise GDP Growth Rate
Yearwise GDP per Capita
Box Plot of Annual % Change in ...
GDP Analysis

9. MACHINE LEARNING ALGORITHMS

In this project we have used a machine learning model called as Auto - Regressive Integrated Moving Average (ARIMA) and various tools for data pre-processing.

9.1 ARIMA – AutoRegressive Integrated Moving Average

Time series analysis involves modelling and forecasting data points collected over time. ARIMA, which stands for AutoRegressive Integrated Moving Average, is a popular statistical model widely used for time series forecasting. It combines autoregression, differencing, and moving averages to capture different aspects of time series data.[10]

Components of ARIMA:

1. <u>AutoRegressive (AR) Component (p)</u> :

- The autoregressive component represents the relationship between the current observation and its past observations.
- The parameter 'p' denotes the order of the autoregressive component, indicating how many past observations are considered in the model.

2. <u>Integrated (I) Component (d)</u> :

- The integrated component involves differencing the time series data to make it stationary.
- The parameter 'd' represents the number of differencing required to achieve stationarity.

3. <u>Moving Average (MA) Component (q)</u> :

- The moving average component captures the relationship between the current observation and a residual error from a moving average of past observations.
- The parameter 'q' denotes the order of the moving average component, indicating how many past residuals are considered in the model.

Working of ARIMA Model:

1. Data Pre-processing :

Time series data often has trends and seasonality. ARIMA assumes the data to be stationary, so pre-processing steps like differencing may be needed.

2. Model Specification :

The order of the ARIMA model is specified by (p, d, q) parameters. Choosing appropriate values for these parameters is crucial and can be done using methods like grid search or analysing autocorrelation and partial autocorrelation plots.

3. Model Training :

The ARIMA model is trained using historical time series data. The model estimates coefficients for the autoregressive, differencing, and moving average components.

4. Forecasting :

Once trained, the ARIMA model can be used to make future predictions. Forecasting involves generating predictions based on the learned patterns and relationships in the historical data.

Advantages of ARIMA:

1. Simplicity :

ARIMA is a relatively simple model that doesn't require extensive feature engineering.

2. Effective for Stationary Data :

ARIMA performs well when the time series data is stationary, and trends and seasonality are well-captured.

3. Interpretability :

The model's parameters have clear interpretations, making it easy to understand the relationships within the time series data.

Limitations of ARIMA:

1. Assumption of Stationarity :

ARIMA assumes that the data is stationary, and transforming non-stationary data might be challenging.

2. May Not Capture Complex Patterns :

ARIMA might struggle with capturing complex patterns or sudden changes in the time series data.

3. Parameter Selection Challenges :

Selecting the appropriate values for (p, d, q) can be challenging and may require domain knowledge or additional analysis.

Thus, the ARIMA model is a valuable tool for time series forecasting, providing a balance between simplicity and effectiveness. Understanding the components and working principles of ARIMA is crucial for making informed decisions when applying this model to real-world time series data.

9.2 Python Code for Forecasting –

```
import pandas as pd
from statsmodels.tsa.arima.model import ARIMA
import matplotlib.pyplot as plt

# Read data from excel file
data = pd.read_excel("/content/sample_data/India.xlsx")

# Selecting the data for ARIMA model
india_gdp = data[['Year', 'GDP']]
india_gdp.set_index('Year', inplace=True)

# Fit the ARIMA model
model = ARIMA(india_gdp, order=(2,1,2))
model_fit = model.fit()

# Forecasting for the next 5 years
forecast = model_fit.forecast(steps=5)

# Plotting the forecast
plt.figure(figsize=(10, 6))
plt.plot(india_gdp.index, india_gdp['GDP'], label='Historical GDP')
plt.plot(range(2023, 2028), forecast, color='red', linestyle='--', label='Forecasted
GDP')
plt.title('India GDP Forecast 2023-2027')
plt.xlabel('Year')
plt.ylabel('GDP in Billion USD')
plt.legend()
plt.show()

# Output the forecast values
print('Forecasted GDP values for 2023 to 2027 in Billion USD :\n',forecast)
```

10. RESULTS

1. USA

YEAR	Forecasted GDP (Billion USD)
2023	25792.30
2024	26502.32
2025	28516.36
2026	29955.06
2027	30502.64

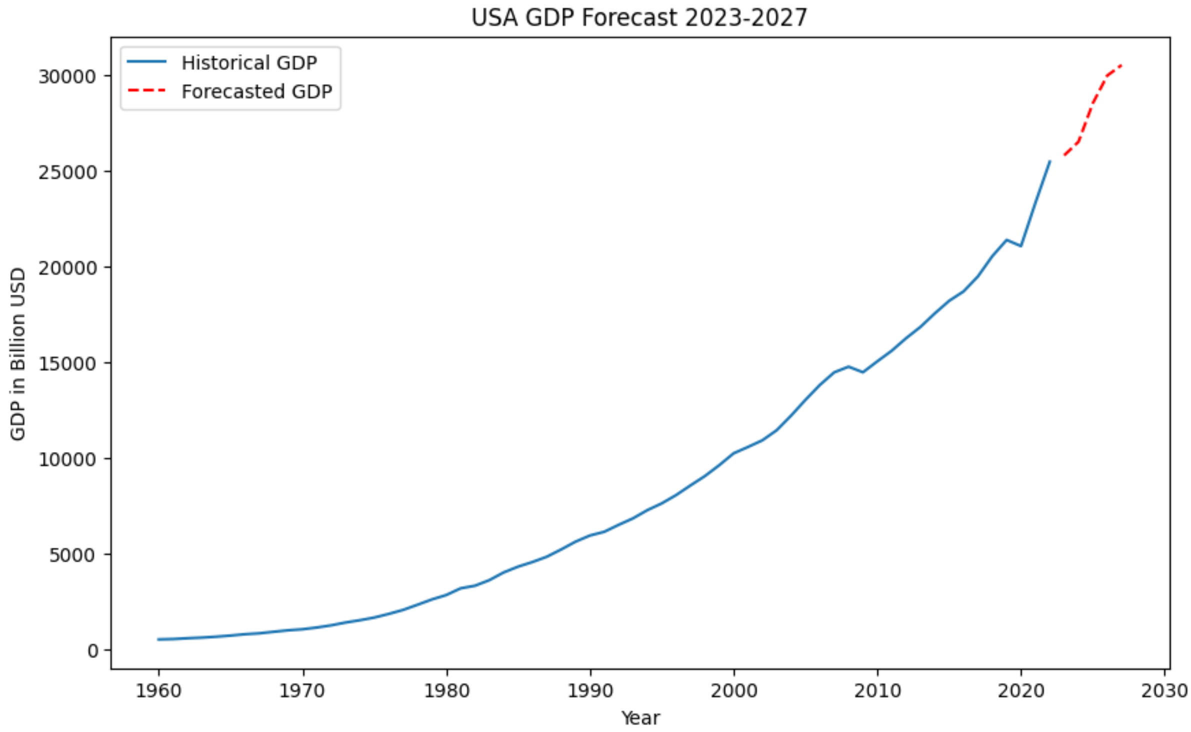


Fig : GDP Forecast for USA

2. China

YEAR	Forecasted GDP (Billion USD)
2023	18068.86
2024	20748.39
2025	21735.59
2026	21363.47
2027	23569.07

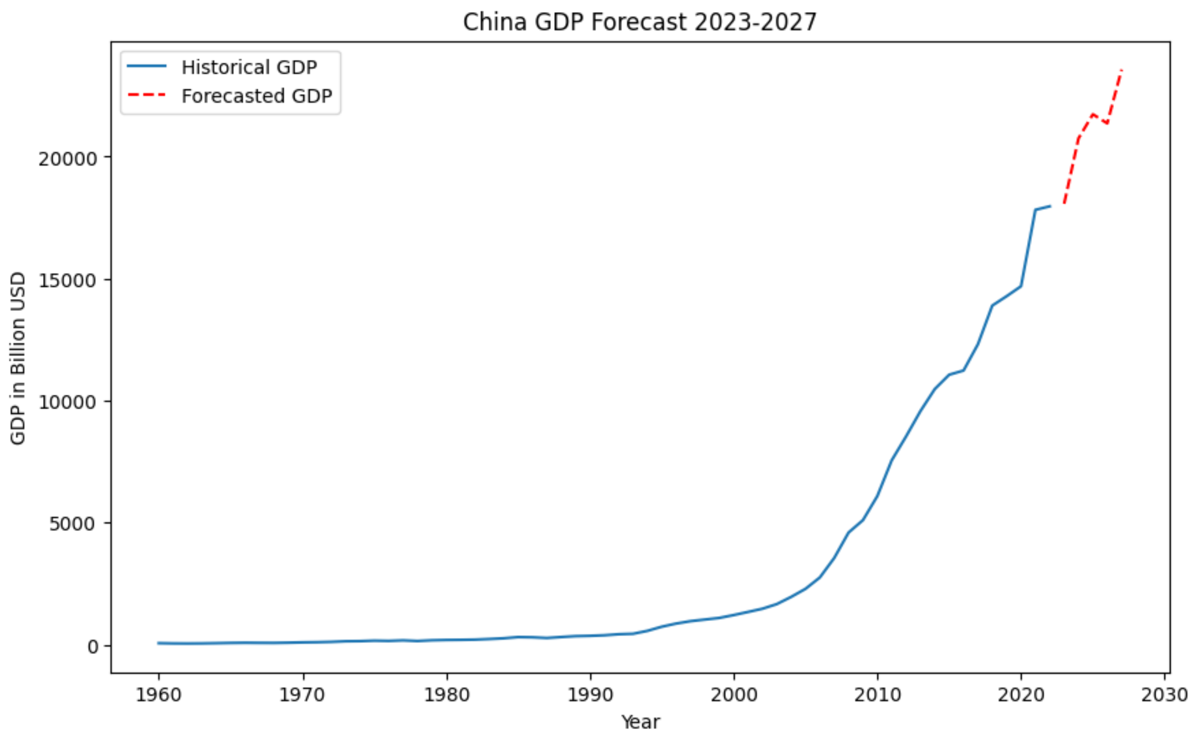


Fig : GDP Forecast for China

3. Japan

YEAR	Forecasted GDP (Billion USD)
2023	3972.09
2024	3892.37
2025	4067.42
2026	4041.73
2027	3974.29

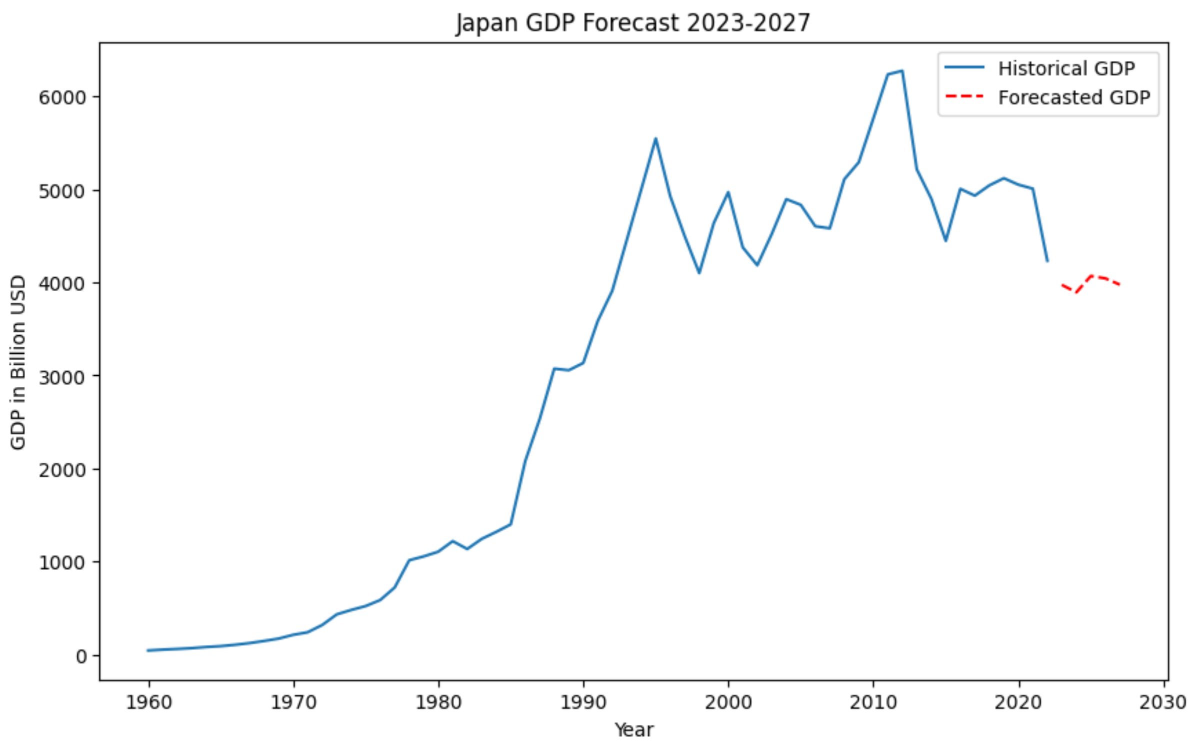


Fig : GDP Forecast for Japan

4. Germany

YEAR	Forecasted GDP (Billion USD)
2023	3978.66
2024	4123.87
2025	4117.93
2026	4074.89
2027	4101.08

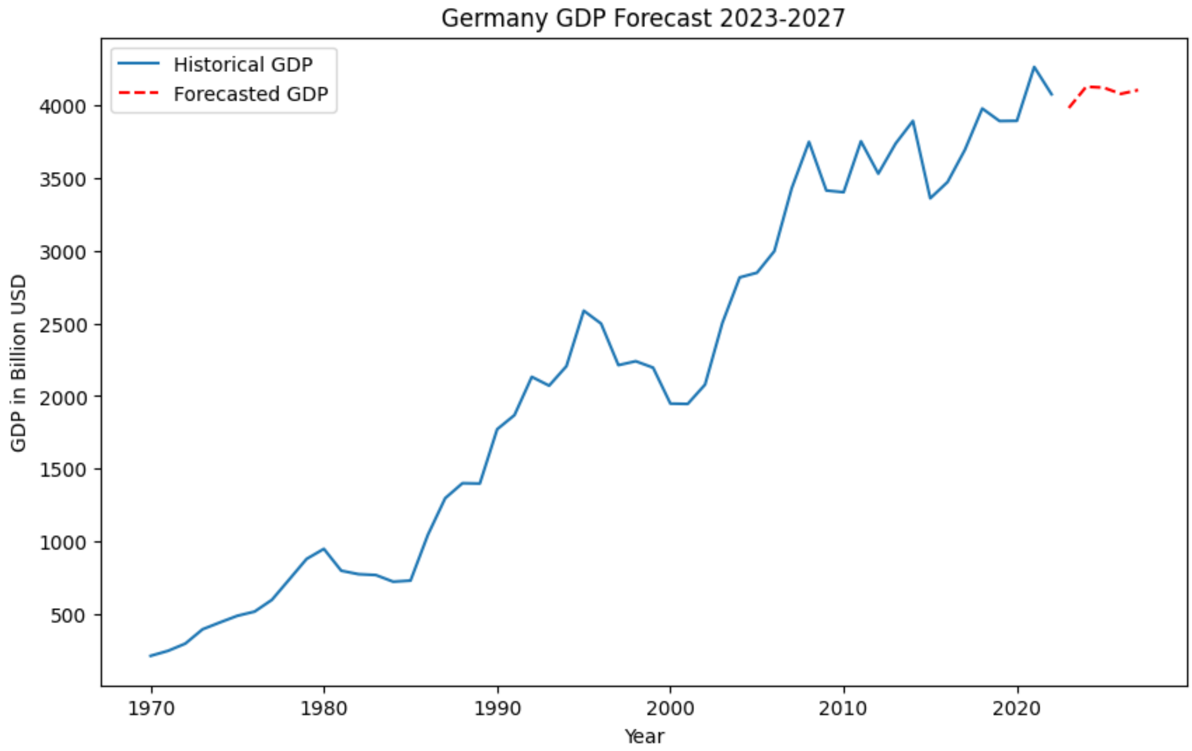


Fig : GDP Forecast for Germany

5. India

YEAR	Forecasted GDP (Billion USD)
2023	3486.08
2024	3619.01
2025	3767.39
2026	3922.63
2027	4050.28

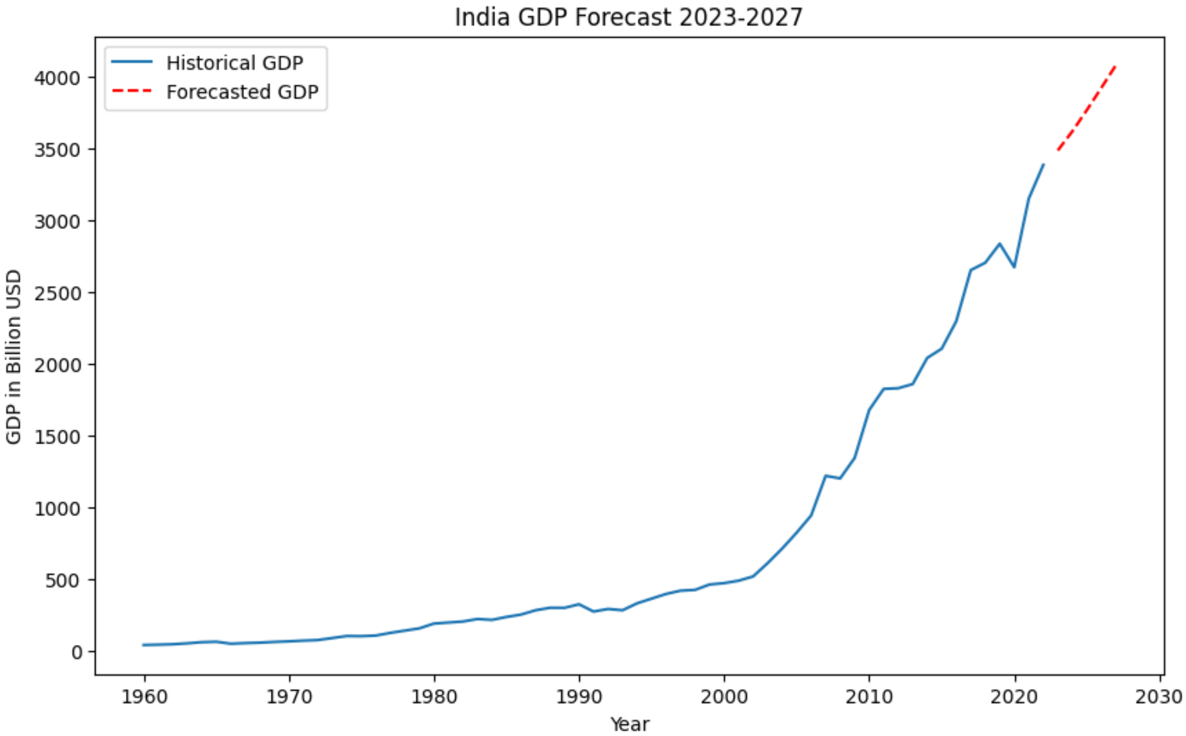


Fig : GDP Forecast for India

11. CONCLUSION

In conclusion, this ambitious project delved into the intricate analysis of six decades of Gross Domestic Product (GDP) data from five countries. The endeavour seamlessly integrated data visualization, dynamic dashboard creation, and the application of forecasting models to offer a comprehensive understanding of the economic trajectories of these nations.

Through meticulous data pre-processing and transformation, the project ensured the quality and integrity of the extensive historical GDP datasets. The development of insightful visualizations not only facilitated the identification of long-term patterns but also enabled a nuanced exploration of the economic trends that shaped each country's trajectory.

The creation of a dynamic dashboard served as a user-friendly interface, allowing stakeholders to interact with the wealth of information dynamically. This not only enhanced accessibility but also provided a powerful tool for decision-makers to gain insights into the economic dynamics of the selected countries.

The implementation of advanced forecasting models for projecting GDP values over the next five years highlighted a forward-looking perspective, incorporating cutting-edge Machine Learning algorithms. This forecasting capability adds a predictive dimension to the project, offering valuable insights for strategic planning and policy formulation.

By visualizing and forecasting GDP data on a multi-decade scale, this project contributes significantly to the understanding of economic evolution, paving the way for informed decision-making and policy development. The amalgamation of historical analysis, dynamic visualization, and future projections positions this project as a valuable asset for researchers, policymakers, and anyone seeking a nuanced comprehension of the economic trajectories of the selected nations. Overall, the project stands as a testament to the power of data-driven insights in unravelling the complexities of economic landscapes over time.

12. REFERENCES

[1] https://www.imf.org/en/Publications/fandd/issues/Series/Back-to-Basics/gross-domestic-product-GDP

[2] https://www.investopedia.com/terms/r/recession.asp

[3] https://www.macrotrends.net/countries/ranking/gdp-gross-domestic-product

[4] Gary L Shelley and Frederick H Wallace. Inflation, money, and real GDP in Mexico: a causality analysis. Applied Economics Letters, 11(4):223–225, 2004.

[5] Gang Long. GDP prediction by support vector machine trained with genetic algorithm. In 2010 2nd International Conference on Signal Processing Systems, volume 3, pages V3–1. IEEE, 2010.

[6] J. Roush, K. Siopes and G. Hu, ”Predicting gross domestic product using autoregressive models,” 2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA), 2017, pp. 317-322.

[7] Martin Schneider, Martin Spitzer, et al. Forecasting Austrian GDP using the generalized dynamic factor model. Technical report, 2004.

[8] https://docs.python.org/3/

[9] https://www.tableau.com/blog/beginners-guide-tableau-public

[10] https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/