**Macroeconomic Indicators and GDP: An Exploratory Modelling and Scenario Analysis for India (2004–2024)**

**Author: Shubh Bhaskar**

This project investigates the complex relationships between key macroeconomic indicators and India's quarterly GDP performance over 2016–2024, this study aims to:

* Understand how variables like CPI, M3, Exchange Rate, Repo Rate, GFCF, etc. influence GDP.
* Perform extensive Exploratory Data Analysis (EDA).  
  Engineer meaningful features (lags, rolling stats, changes) to enhance economic interpretation.
* Fit interpretable models (Linear Regression) to quantify relationships.
* Simulate hypothetical economic scenarios (e.g., CPI shocks) to assess model-driven insights.

**Data Overview:**

| **Type** | **Details** |
| --- | --- |
| Source | IMF, OECD Development Release |
| Period | 2004 Q2 – 2024 Q4 |
| Frequency | Quarterly |
| Total Quarters | 83 |
| Target Variable | GDP (Real Gross Domestic Product) |

**Feature Understanding:**

The dataset used for this project contains a range of quarterly macroeconomic indicators that are theoretically known to influence a country’s economic performance, especially its GDP. Each variable was carefully selected based on its economic relevance to India’s growth dynamics:

* **CPI (Consumer Price Index)**: Represents inflation in the economy. A rise in CPI indicates increasing price levels for goods and services. Inflation affects purchasing power, consumption, and investment decisions, thereby potentially influencing GDP.
* **GDP (Gross Domestic Product)**: The target variable for this project. It represents the total monetary value of all goods and services produced within India during a quarter — a direct measure of economic health.
* **M3 (Broad Money Supply)**: Reflects the total money available in the economy. An increase in money supply (M3) can stimulate demand and investment but may also cause inflation if unchecked.
* **XI RATIO (Export-Import Ratio):** Indicates the trade balance. A higher ratio suggests stronger exports relative to imports, which can contribute positively to GDP via net exports.
* **ER (Exchange Rate - INR/USD):** Affects the competitiveness of exports and the cost of imports. A weaker INR (higher ER) can boost exports but raise import costs.
* **PROD (Manufacturing Production)**: A direct measure of industrial output, contributing to GDP growth through production and employment generation.
* **INT (Interbank Interest Rate)**: Reflects liquidity in the market and the cost of short-term borrowing among banks, affecting credit availability to businesses and consumers.
* **SP (Share Prices)**: Indicates stock market performance, investor confidence, and indirectly reflects economic growth expectations.
* **WUI (World Uncertainty Index):** Measures global economic and political uncertainty. Higher uncertainty can reduce trade, investment, and economic activity.
* **GFCF (Gross Fixed Capital Formation**): Represents investment in infrastructure, machinery, and equipment. A critical factor in long-term productive capacity and GDP growth.
* **EP (Electricity Production)**: A proxy for industrial activity. Higher electricity generation often signals economic expansion.
* **WP (Wholesale Price Index):** Measures wholesale-level inflation. Like CPI, this affects production costs and competitiveness.
* **CAB (Capital Account Balance):** Indicates net foreign exchange inflows/outflows. Persistent deficits or surpluses impact GDP through capital flows and exchange rate adjustments.
* **REIR (Real Effective Exchange Rate):** Measures currency strength after adjusting for inflation relative to trading partners — impacts competitiveness.
* **WGDP (World GDP Growth):** Global economic conditions that indirectly influence Indian exports and investment flows.
* **REPO (Repo Rate):** The policy interest rate set by RBI. Lower repo rates typically stimulate borrowing and spending, while higher rates restrain inflation.
* **UR (Unemployment Rate):** The percentage of the labor force that is unemployed. High unemployment usually corresponds to lower GDP due to reduced productivity and consumption.

**Data Preprocessing:**

Before modelling, several data cleaning and preparation steps were performed:

1. **Missing Value Handling:**

* For quarterly data converted from annual sources (like GFCF, FDI), forward filling or mean imputation was used to maintain data consistency.
* Rolling, lagged, and percentage change features naturally produced missing values at the start of the dataset — these rows were dropped as they could not be meaningfully filled.

1. **Final Data Shape**:  
   After cleaning:  
   83 observations with 31 usable features remained, ready for feature engineering and modelling.
2. **Exploratory Data Analysis (EDA):**

Trend Analysis:

* **GDP:** Demonstrated a steady upward trend with a clear dip during 2020 (COVID-19 impact) and recovery afterward.
* **CPI and M3:** Both displayed consistent growth, reflecting inflation and expansionary monetary policy.
* **Exchange Rate:** Showed gradual depreciation of the INR against USD.
* **REPO Rate**: Fluctuated according to RBI monetary policy — cuts during crisis periods, hikes during inflationary phases.
* **WUI**: Spiked during global uncertainty events such as COVID-19, indicating rising market fears.

**Correlation Analysis**:

* M3, GFCF, and manufacturing production were positively correlated with GDP, as expected — more money, investment, and output drive economic growth.
* Surprisingly, CPI and GDP correlation was weak, possibly due to India's inflation-tolerant economy and government controls.
* Repo Rate and GDP showed mild inverse correlation — consistent with monetary policy theory (tightening policy restricts growth).

**Feature Engineering**:

To capture time-dependent and interaction effects, the following were created:

1. **Lag Features:** Previous quarter values (1-period lag) for CPI, M3, ER, SP, REPO, etc.
2. **Percentage Change Features:** Quarter-on-quarter growth/decline rates.
3. **Rolling Means (4 quarters):** To smooth seasonal and short-term noise (for CPI, M3, ER, SP).
4. **Interaction Terms:** CPI x M3 (explores how inflation and money supply combined influence GDP).
5. **Shock Features:** Captured sudden changes (CPI\_Shock) to model economic surprises.

**Model Building and Evaluation:**

* **Linear Regression** was chosen as the final model — interpretable, simple, and appropriate for small datasets.
* Top 25 features were selected using SelectKBest (F-regression) to avoid overfitting and focus on economically meaningful predictors.

**Performance Metrics on Test Set**:

| Metric | Value | Interpretation |
| --- | --- | --- |
| R² Score | 0.774 | Model explains ~77% of GDP variance |
| RMSE | 1.78 million | Error is only ~6.5% of mean GDP — low prediction uncertainty. |
| MAE | 1.23 million | Average absolute error is ~4.5% of mean GDP |

**Scenario Simulation — CPI Shock (+10%):**

To demonstrate model interpretability:

* Simulated a +10% CPI increase (hypothetical inflation shock).
* Model predicted GDP would rise by ~249,484 units (~0.6%) — possibly reflecting India’s historical resilience to inflation shocks.

**Conclusion and Insights:**

* The project effectively demonstrated the relationship between macroeconomic variables and GDP using transparent, interpretable methods.
* The model performed reliably with low error, considering economic data complexity and size limitations.
* The shock simulation revealed how the model responds to hypothetical scenarios, adding value beyond pure prediction.

**Limitations:**

* Small dataset (83 quarters) limits complexity.
* Some annual data had to be estimated quarterly — introducing potential smoothing artifacts.
* The model is not designed for real-time forecasting but for relationship exploration and scenario analysis.