



**SYMBIOSIS STATISTICAL INSTITUTE**

[Constituent of SYMBIOSIS INTERNATIONAL (DEEMED UNIVERSITY)]

Established under section 3 of the UGC Act, 1956 Re-accredited by NAAC with A++ grade [Awarded Category - I by UGC]

## Internal component-4-Report

Vivek Sharma – 23060641081

Shruti Shukla – 23060641083

Course: Linear Models

Under the Guidance of **Dr. Priya Deshpande**

### **Forecasting 2023 GDP: A Comparative Analysis of Predicted vs. Actual Economic Performance**

#### ➤ **Introduction:**

1. GDP prediction for 2023 is essential for understanding how economies are expected to perform.
2. We're examining if these predictions match up with what actually happened.
3. This project endeavors to assess the accuracy of GDP predictions for the year 2023 across various countries, employing a comprehensive set of economic indicators. By scrutinizing factors such as GDP, GDP per capita, Real GDP growth, government expenditure, export and import figures, tax total rate, unemployment rate, minimum wages, and population size, we aim to discern the extent to which predictive models align with real-world economic performance.
4. We're doing this to see how accurate predictions are and what factors influence economic performance.
5. The significance of this analysis lies in its implications for policy formulation, investment strategies, and economic forecasting

methodologies. Understanding the disparities between predicted and actual GDP offers valuable insights into the complexity of economic systems and the challenges of forecasting future trends accurately.

6. Why It Matters:

Accurate GDP predictions help plan for the future.

Understanding the difference between predictions and reality helps us improve forecasting methods.

It shows us where our economy is strong or needs improvement.

This knowledge can guide decisions on spending, taxes, and other policies.

It helps us understand the complex world of economics better.

7. What We're Doing:

Gathering data on predicted and actual GDP for 2023.

Looking at factors like GDP per person, economic growth rate, government spending, etc.

Comparing predictions with what actually happened.

Analyzing why predictions might have been right or wrong.

Drawing conclusions about the accuracy of GDP predictions and their implications.

8. What We Hope to Find:

Whether predictions matched reality across different countries.

Which factors most strongly influence GDP predictions.

How accurate predictions are overall.

Ways to improve future GDP predictions.

How this information can help decision-makers and investors.

9. Comparing the actual values of the dependent variable (GDP in this case) with the values predicted by the model (fitted values) helps us evaluate how well the model performs in making predictions.

10. Here's why we compare the actual vs. fitted values:

#### A. Assessing Model Accuracy

Comparing actual and predicted values allows us to assess how accurately the model predicts the dependent variable.

It helps us understand how well the model generalizes to new, unseen data.

#### B. Identifying Model Errors

Differences between actual and predicted values (residuals) highlight where the model makes errors.

By examining these errors, we can identify patterns or trends that the model might have missed.

#### C. Model Validation

Comparing actual and predicted values is a crucial step in validating the predictive performance of the model.

If the predicted values closely match the actual values, it suggests that the model is reliable and can be used for making predictions on new data.

#### D. Model Improvement

Discrepancies between actual and predicted values provide insights into areas where the model can be improved.

Understanding where the model performs well and where it struggles can guide adjustments to the model structure or data preprocessing.

#### ➤ **Problem Statement:**

- To assess the accuracy of GDP predictions for the year 2023, this project aims to conduct exploratory data analysis (EDA) followed by model development using multiple regression techniques. The primary objective is to determine if the actual GDP aligns with the predicted values derived from the model, Subsequently through model development using training and testing data, the project seeks to determine the accuracy of predicting

the actual GDP based on factors such as GDP per capita, real GDP growth, government expenditure as a percentage of GDP, import and export values, total tax rates, unemployment rates, minimum wage, and population demographics.

➤ **Dataset / resource link / file:**

Dataset- [Countries GDP \(kaggle.com\)](https://www.kaggle.com/datasets/alexisbisa/countries-gdp)

File-1



ORIGINAL DATA.csv

(Original Data)

File-2



Log Transform.csv

(Logarithmic Transformed Data)

**Note:** Here we transformed our original Data into **Logarithmic transformation** to improve the overall performance and interpretability of the multiple linear regression model.

**Why we need transformation:**

When the coefficient of determination (R-squared) is low, it indicates that the model does not explain much of the variability in the data. In such cases, transforming the data can sometimes improve the fit of the model by addressing issues such as heteroscedasticity (unequal variances across the range of predictors) or nonlinearity.

Here are some reasons why we might transform data when the multiple  $R^2$  is low:

- Nonlinearity
- Heteroscedasticity
- Outliers
- Assumption Violations
- Interpretability

➤ **Code:**

```
# Load required libraries
```

```
library(dplyr)
```

```
library(ggplot2)
```

```
library(reshape2)
```

```
library(rpart)
```

```
library(corrplot)
```

```
library(caret)
```

**#CODE FOR EDA**

```
#Accessing Data
```

```
data = read.csv("C:\\Users\\ABHISHEK\\OneDrive\\Desktop\\LM Project\\Log  
Transform.csv")
```

```
# View the structure of your dataset
```

```
str(data)
```

```
# Summary statistics
```

```
summary(data)
```

```
# Correlation matrix
```

```
correlation_matrix = cor(data[,c("GDP", "Import...", "Export...",  
                                "Population", "Unemployment.rate...", "Minimum.Wage...month.",  
                                "GDP.per.capita..in...", "Real.GDP.Growth..in..",  
                                "Gov.Expenditure...of.GDP", "Tax.total.rate...")])
```

```
corrplot(correlation_matrix, method = "color")
```

```
# Histogram for each variable
```

```
par(mfrow=c(3, 3))
```

```
for(i in 2:10) {
```

```
  hist(data[,i], main = colnames(data)[i], xlab = colnames(data)[i], col = "skyblue", border =  
  "darkblue")
```

```
}
```

```
# Box plot for each variable
```

```
par(mfrow=c(3, 3))
```

```
for(i in 2:10) {
```

```
boxplot(data[,i], main = colnames(data)[i], col = "skyblue", border = "darkblue")
}
```

### #CODE FOR MODEL DEVELOPMENT

#### # MULTIPLE LINEAR REGRESSION

```
#storing GDP_data in a dataframe called 'data'
```

```
dataframe = as.data.frame(data)
```

```
colnames(data)
```

#### # Multiple Linear Regression

```
model = lm(dataframe$GDP ~ dataframe$GDP.per.capita.in...
+ dataframe$Real.GDP.Growth..in..
+ dataframe$Gov.Expenditure...of.GDP + dataframe$Import...
+ dataframe$Export...+ dataframe$Tax.total.rate...
+ dataframe$Unemployment.rate...+
dataframe$Minimum.Wage...month. + dataframe$Population)
```

#### # Summary of the regression model

```
summary(model)
```

#### # Visualization of Residuals

```
residuals = resid(model)
```

```
fitted = fitted(model)
```

#### # Plotting Residuals vs Fitted Values

```
ggplot(data, aes(x = fitted, y = residuals)) +
geom_point() +
geom_hline(yintercept = 0, linetype = "dashed", color = "red") +
labs(title = "Residuals vs Fitted Values", x = "Fitted Values", y = "Residuals")
```

#### # Plotting Normal Q-Q Plot

```
qqnorm(residuals)
```

```
qqline(residuals)
```

#### # Plotting Residuals vs Predictors

```
par(mfrow = c(3, 3)) # Arrange plots in a 3x3 grid
```

```
for(i in 2:length(model$coefficients)) {
```

```
plot(data[, i], residuals, xlab = names(data)[i], ylab = "Residuals")
```

```

}

# COMPAIRISON OF ACTUAL VS FITTED

# Split the data into training and testing sets
set.seed(123) # for reproducibility

train_index = createDataPartition(data$GDP, p = 0.8, list = FALSE)
train_data = data[train_index, ]
colnames(train_data)

test_data = data[-train_index, ]

# Train the linear regression model
lm_model = lm(GDP ~ ., data = train_data)

# Train the multiple linear regression model
lm_model = lm(GDP ~ GDP.per.capita..in... + Real.GDP.Growth..in.. +
Gov.Expenditure...of.GDP
+ Import... + Export... + Tax.total.rate... + Unemployment.rate...
+ Minimum.Wage...month. + Population, data = train_data)

# Predict GDP for training data
train_predictions = predict(lm_model, newdata = train_data)

# Predict GDP for testing data
test_predictions = predict(lm_model, newdata = test_data)

# Compare actual GDP with predicted GDP for training data
train_comparison = data.frame(Actual_GDP = train_data$GDP, Predicted_GDP =
train_predictions)

# Compare actual GDP with predicted GDP for testing data
test_comparison = data.frame(Actual_GDP = test_data$GDP, Predicted_GDP =
test_predictions)

# Output
cat("\nTesting Data Comparison:\n")
print(head(test_comparison))

# Mean Absolute Error (MAE) for test data
test_MAE = mean(abs(test_comparison$Actual_GDP - test_comparison$Predicted_GDP))

# Mean Squared Error (MSE) for test data
test_MSE = mean((test_comparison$Actual_GDP - test_comparison$Predicted_GDP)^2)

# Root Mean Squared Error (RMSE) for test data

```

```

test_RMSE = sqrt(test_MSE)
# R-squared for test data
test_Rsquared = summary(lm_model)$adj.r.squared
cat("\nTest Data:\n")
cat("Mean Absolute Error (MAE):", test_MAE, "\n")
cat("Mean Squared Error (MSE):", test_MSE, "\n")
cat("Root Mean Squared Error (RMSE):", test_RMSE, "\n")
cat("R-squared:", test_Rsquared, "\n")

```

### ➤ Output:

#### EDA,

```

> # Summary statistics
> summary(data)

```

| GDP                   | Import...              | Export...               | Population     | Unemploym |
|-----------------------|------------------------|-------------------------|----------------|-----------|
| ent.rate...           |                        |                         |                |           |
| Min. : 8.775          | Min. : 5.502           | Min. : 4.434            | Min. : 4.580   | Min. : -  |
| 0.1524                |                        |                         |                |           |
| 1st Qu.: 10.051       | 1st Qu.: 8.997         | 1st Qu.: 8.669          | 1st Qu.: 6.362 | 1st Qu.:  |
| 1.7901                |                        |                         |                |           |
| Median : 10.603       | Median : 9.424         | Median : 9.344          | Median : 6.953 | Median :  |
| 2.3091                |                        |                         |                |           |
| Mean : 10.665         | Mean : 9.460           | Mean : 9.311            | Mean : 6.872   | Mean :    |
| 2.2622                |                        |                         |                |           |
| 3rd Qu.: 11.375       | 3rd Qu.: 9.943         | 3rd Qu.: 10.191         | 3rd Qu.: 7.412 | 3rd Qu.:  |
| 2.8846                |                        |                         |                |           |
| Max. : 13.299         | Max. : 11.593          | Max. : 12.824           | Max. : 9.145   | Max. :    |
| 4.6290                |                        |                         |                |           |
| Minimum.Wage...month. | GDP.per.capita...in... | Real.GDP.Growth...in... | Gov.Expend     |           |
| iture...of.GDP        |                        |                         |                |           |
| Min. : -2.2007        | Min. : 2.831           | Min. : -3.000           | Min. : -0      |           |
| .9929                 |                        |                         |                |           |
| 1st Qu.: -1.4186      | 1st Qu.: 3.844         | 1st Qu.: -1.683         | 1st Qu.: -0    |           |
| .6639                 |                        |                         |                |           |
| Median : -1.2522      | Median : 4.272         | Median : -1.420         | Median : -0    |           |
| .5169                 |                        |                         |                |           |
| Mean : -1.2508        | Mean : 4.173           | Mean : -1.369           | Mean : -0      |           |
| .5375                 |                        |                         |                |           |
| 3rd Qu.: -1.0004      | 3rd Qu.: 4.634         | 3rd Qu.: -1.252         | 3rd Qu.: -0    |           |
| .4137                 |                        |                         |                |           |
| Max. : -0.6306        | Max. : 5.296           | Max. : 0.000            | Max. : -0      |           |
| .2192                 |                        |                         |                |           |
| Tax.total.rate...     |                        |                         |                |           |
| Min. : -1.0969        |                        |                         |                |           |
| 1st Qu.: -0.5161      |                        |                         |                |           |



Median :-0.4214  
Mean :-0.4239  
3rd Qu.:-0.3028  
Max. : 0.3416

Based on the summary statistics of the dataset, we can draw the following conclusions:

Economic Indicators:

The GDP ranges from **8.775\$** to **13.299\$**, with a mean of **10.665\$**.

Import and export values vary between **5.502\$** to **11.593\$** and **4.434\$** to **12.824\$** respectively.

Population ranges from **4.580\$** to **9.145\$**.

Labor Market:

The unemployment rate ranges from **-0.1524** to **4.629%**, with a mean of **2.2622**.

Government and Fiscal Policy:

Minimum wage varies between **-2.2007** to **-0.6306**.

Real GDP growth ranges from **-3.000%** to **0.000%**.

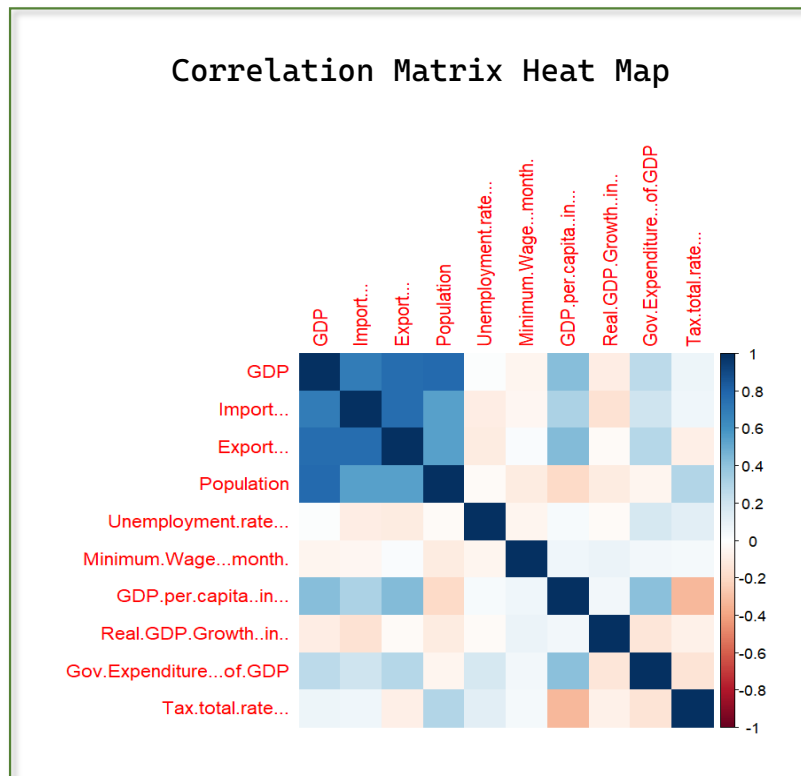
Government expenditure as a percentage of GDP ranges from **-0.9929%** to **-0.2192%**.

Tax total rate ranges from **-1.0969** to **0.3416**.

Income and Economic Well-being:

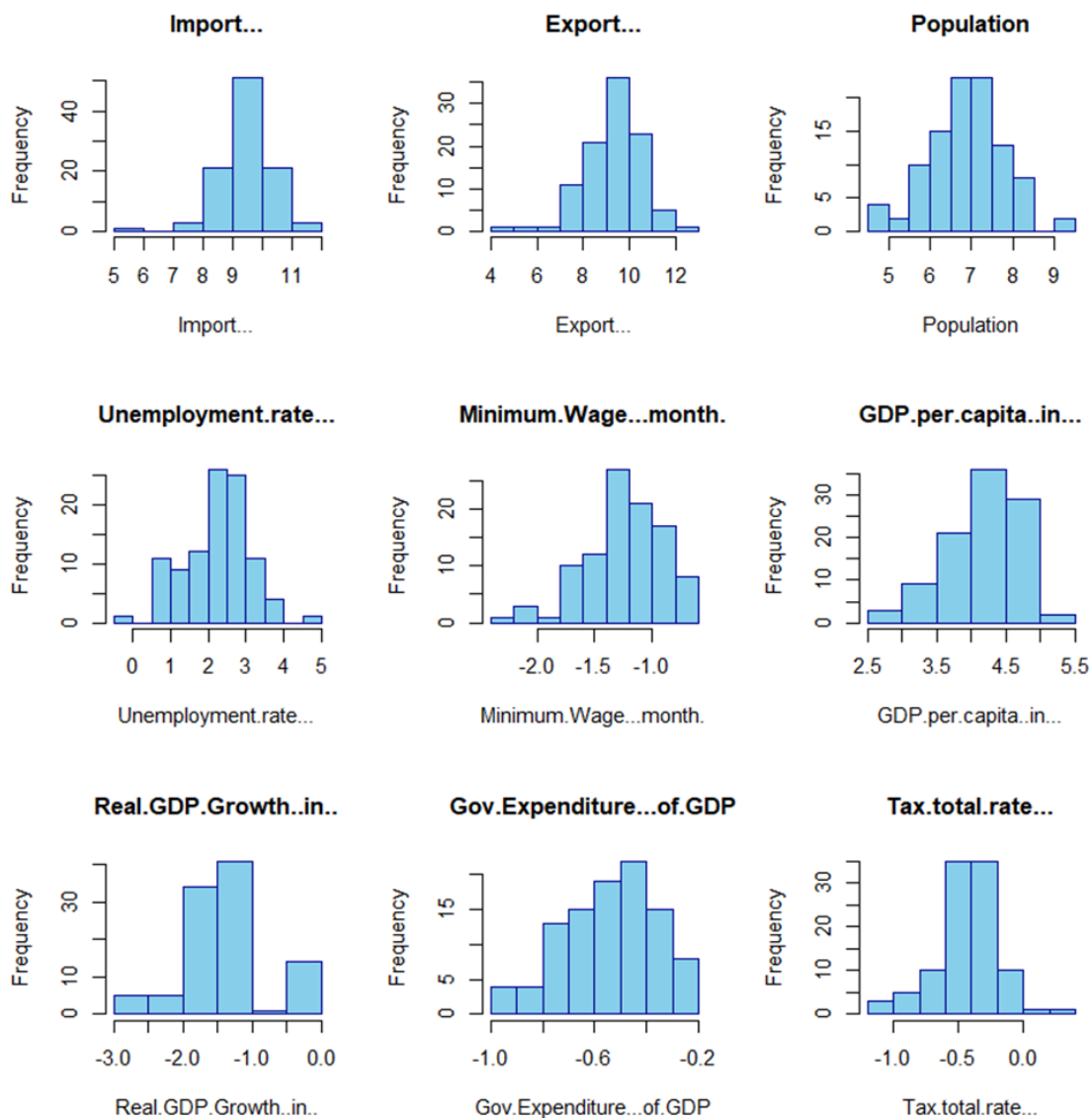
GDP per capita ranges from **2.831** to **5.296**.

Overall, the dataset provides insights into various economic indicators, labor market conditions, government policies, and their impacts on the economy.



- There is a **positive correlation** between a country's GDP and its Imports and Exports. This means that as the Imports and Exports increases, The GDP of a country tends to increase as well.
- There is a **weak positive correlation** between GDP and Population This means that there is a slight tendency for countries with higher populations to have larger GDP.
- There is a **negative correlation** between Real GDP Growth and its unemployment rate, Government Expenditure & Tax Total rate. This means that as the unemployment, Government Expenditure & Tax Total rate of a country decreases, the GDP of a country tends to increase.

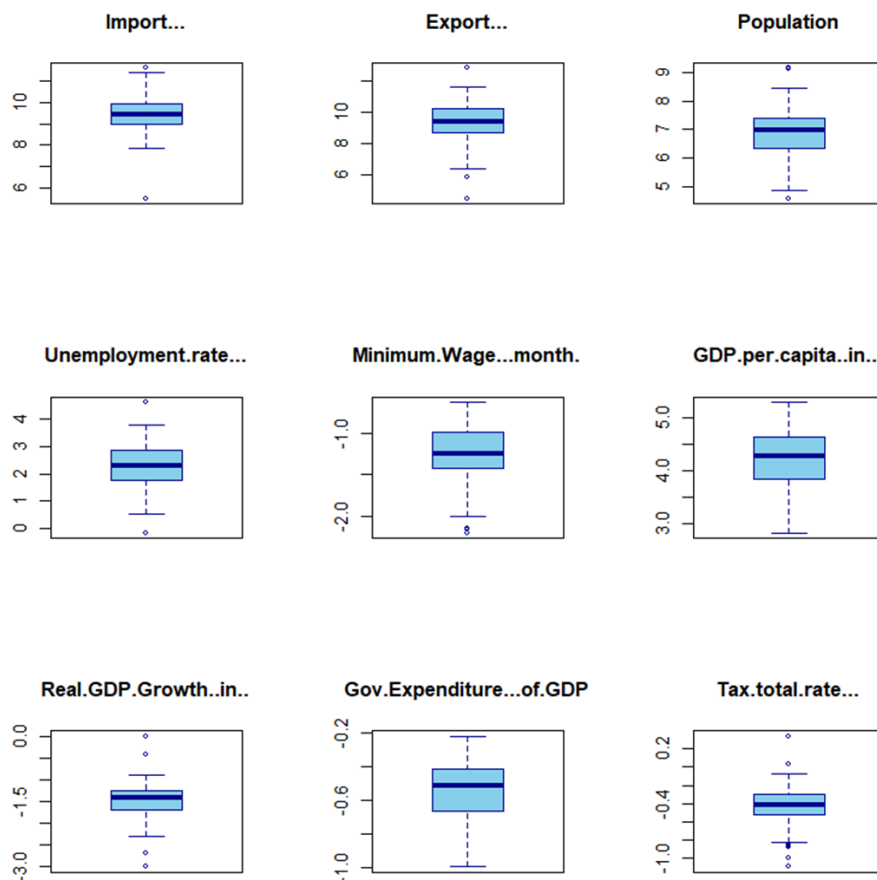
## Individual Histogram For Each Independent Variables



- The Histogram for **Import** and **Export** appears to be symmetrical with a peak in the center and tapering tails towards the left and right sides. This suggests that most countries have import and export values close to the average, with fewer countries having very high or very low import/export values.
- The Histogram for **Population** appears to be symmetrical with a peak in the centre. It is suggested that There are somewhat more countries with lower populations than with higher populations.
- The histogram for **unemployment rate** appears skewed to the right this suggest that their might be more countries with higher unemployment rate than countries with lower unemployment rate.

- The Histogram for **Minimum Wage** is difficult to decipher because it seems to have multiple peaks. It is possible that there are multiple clusters of countries with minimum wage around certain values.
- The histogram for **GDP per capita** is skewed to the right. There might be more countries with lower GDP per capita than with higher GDP per capita.
- The histogram for **Real GDP Growth** appears centered around zero with a slight skew to the right. There might be slightly more countries with negative or zero GDP growth than with positive GDP growth.
- The histogram for **Government Expenditure** is difficult to decipher because it seems to have multiple peaks. It's possible that there are multiple clusters of countries with government expenditure around certain values.
- The histogram for **Tax Total Rate** appears skewed to the right. There might be more countries with lower tax total rates than with higher tax total rates.

## Individual Box plot for each independent variables



## Central Tendencies:

- Import, export, unemployment rate, and tax total rate exhibit a central tendency, with most countries clustered around a central point for these variables.
- Population and GDP per capita show a wider range, with a central tendency but also significant variation across countries.

## Variations:

- Boxplots show variations in values for all variables, with outliers on both ends for most, indicating some countries have extreme values compared to the rest.
- Minimum Wage, Government Expenditure, and possibly Real GDP Growth show complex distributions, potentially with multiple clusters of countries with similar values.

## Economic Relationships:

- While it's difficult to definitively determine relationships between variables from just the boxplots, similar trends in import and export medians suggest a possible correlation between these economic activities.

## Model Development,

Our Hypothesis is:

$H_0$  = There is no significant relationship between the independent variables.

$H_1$  = There is a significant relationship between the independent variables.

```
> #CODE FOR MODEL DEVELOPMENT
> # MULTIPLE LINEAR REGRESSION
> # Multiple Linear Regression
> model = lm(dataframe$GDP ~ dataframe$GDP.per.capita..in...
+           + dataframe$Real.GDP.Growth..in..
+           + dataframe$Gov.Expenditure...of.GDP + dataframe$Import...
+           + dataframe$Export...+ dataframe$Tax.total.rate...
+           + dataframe$Unemployment.rate...+
+           dataframe$Minimum.Wage...month. + dataframe$Population)
>
> # Summary of the regression model
> summary(model)
```

Call:

```
lm(formula = dataframe$GDP ~ dataframe$GDP.per.capita..in... +
    dataframe$Real.GDP.Growth..in.. + dataframe$Gov.Expenditure...of.GDP +
    dataframe$Import... + dataframe$Export... + dataframe$Tax.total.rate... +
    dataframe$Unemployment.rate... + dataframe$Minimum.Wage...month. +
    dataframe$Population)
```

Residuals:

| Min      | 1Q       | Median   | 3Q      | Max     |
|----------|----------|----------|---------|---------|
| -0.59664 | -0.12045 | -0.02491 | 0.10534 | 0.81025 |

Coefficients:

|                                     | Estimate | Std. Error | t value | Pr(> t )   |
|-------------------------------------|----------|------------|---------|------------|
| (Intercept)                         | -0.15942 | 0.34412    | -0.463  | 0.6443     |
| dataframe\$GDP.per.capita..in...    | 0.98310  | 0.05471    | 17.970  | <2e-16 *** |
| dataframe\$Real.GDP.Growth..in..    | -0.03779 | 0.03273    | -1.155  | 0.2513     |
| dataframe\$Gov.Expenditure...of.GDP | 0.32952  | 0.13372    | 2.464   | 0.0156 *   |
| dataframe\$Import...                | 0.03162  | 0.03859    | 0.819   | 0.4147     |
| dataframe\$Export...                | 0.01292  | 0.03175    | 0.407   | 0.6852     |
| dataframe\$Tax.total.rate...        | 0.05379  | 0.10564    | 0.509   | 0.6119     |
| dataframe\$Unemployment.rate...     | 0.01446  | 0.02571    | 0.563   | 0.5752     |
| dataframe\$Minimum.Wage...month.    | -0.01624 | 0.06388    | -0.254  | 0.7999     |
| dataframe\$Population               | 0.93100  | 0.03710    | 25.094  | <2e-16 *** |

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2051 on 90 degrees of freedom

Multiple R-squared: 0.9584, Adjusted R-squared: 0.9542

F-statistic: 230.1 on 9 and 90 DF, p-value: < 2.2e-16

The Multiple linear regression model suggests that:

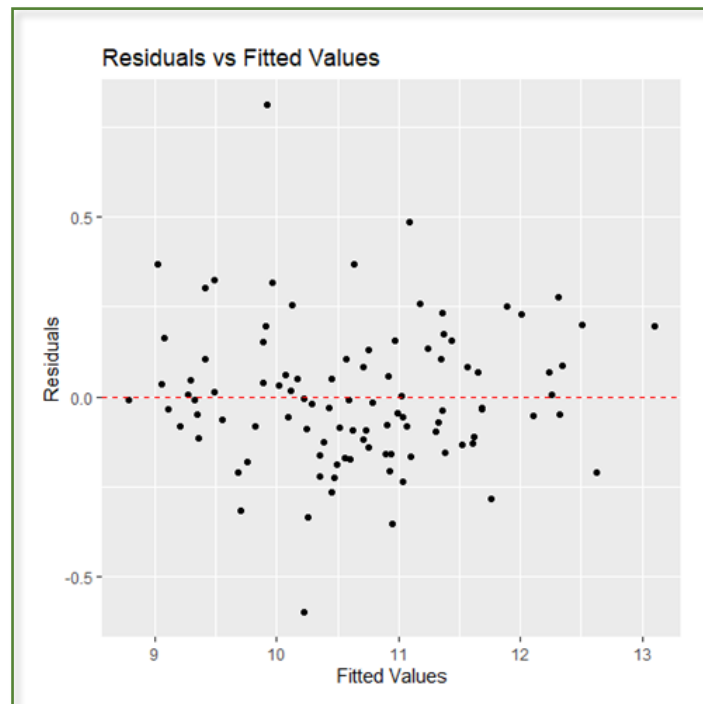
- GDP is significantly positively influenced by GDP per capita and population.
- Government expenditure as a percentage of GDP also has a significant positive impact on GDP.
- Real GDP growth, imports, exports, tax total rate, unemployment rate, and minimum wage do not have a statistically significant impact on GDP.
- The model explains approximately 95.8% of the variance in GDP.
- Population and GDP per capita are the most significant predictors of GDP.

Model also provides the values for **Multiple R-Squared** and **Adjusted R-Squared** as **0.9584** and **0.9542** respectively. That means the model explains approximately **95.8%** of the variance in GDP.

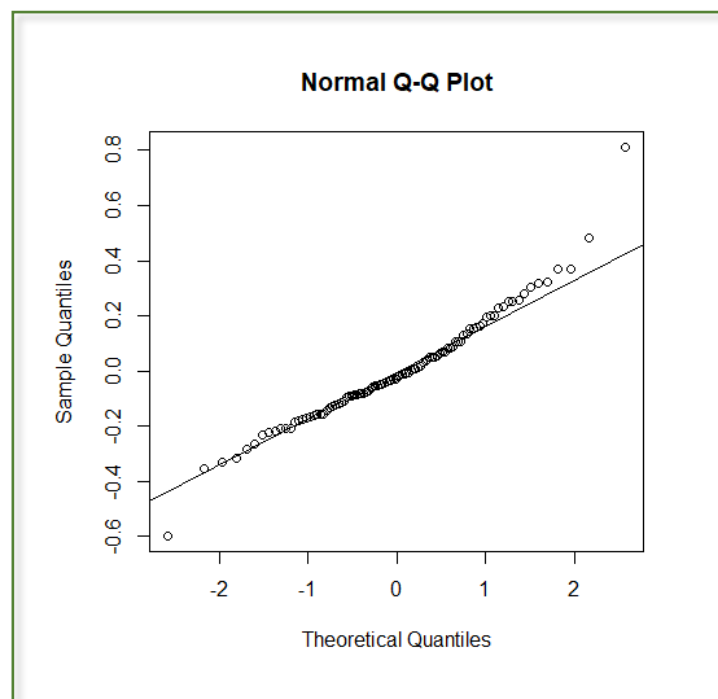
Since **P-Value** is  **$2.2e^{-16}$**  which is less than the **L.O.S. ( $\alpha$ )** which is **0.05**. Hence, we reject the null hypothesis.

**Conclusion:**

There is a significant relationship between the independent variables.

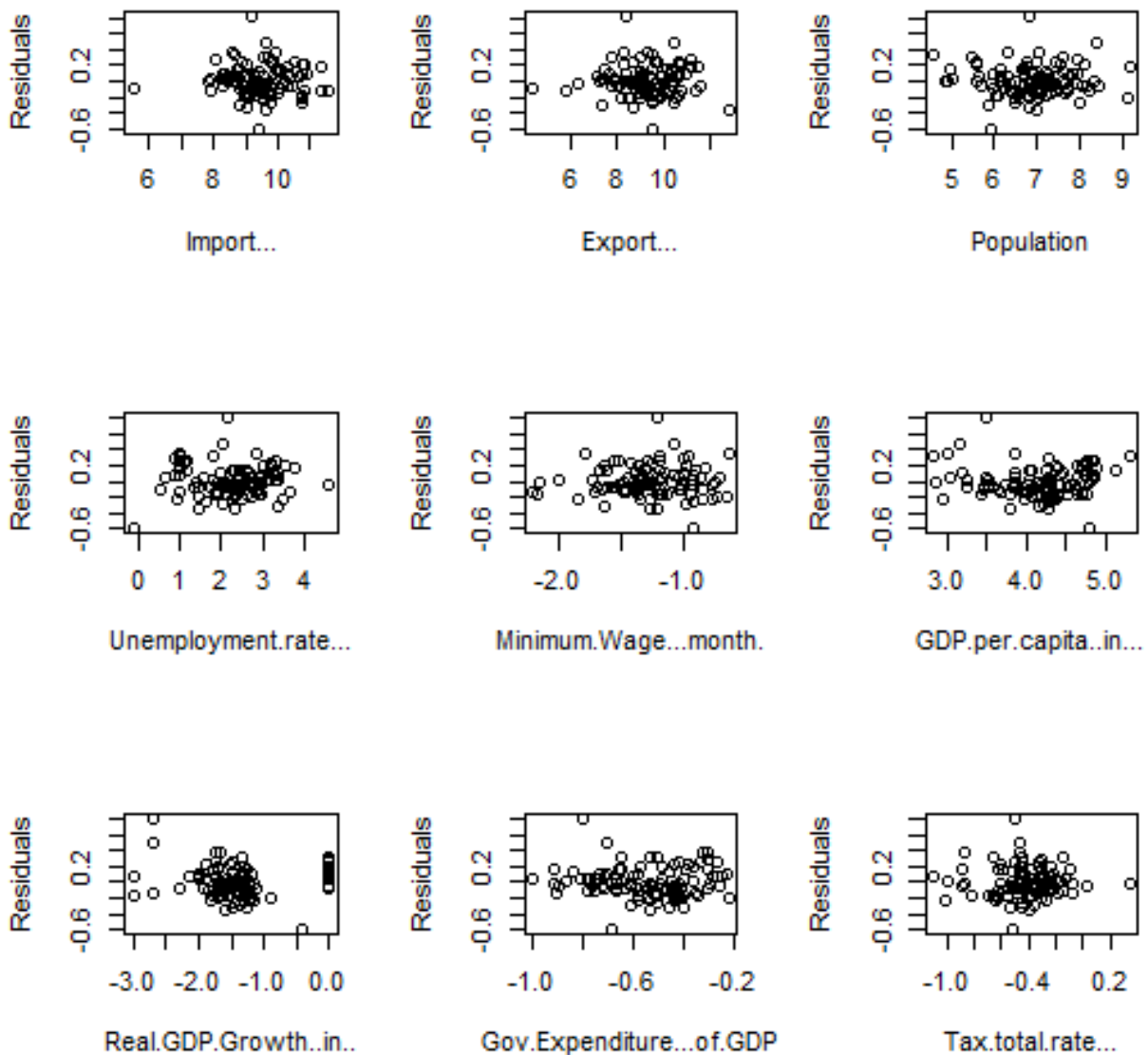


- The residuals appear to be scattered around zero, which is a good sign. This suggests that there is no consistent bias in the model's predictions.
- There is no clear pattern in the residuals. This suggests that the model's errors are random and not systematically related to the fitted values.



- The Normal Q-Q plot compares the observed residuals to the residuals that would be expected if the data were normally distributed.
- Here the residuals are normally distributed since the points on the Q-Q plot fall approximately along a straight diagonal line.
- Based on the Normal Q-Q plot, we can conclude that the residuals from the linear regression model approximately follow a normal distribution.
- This suggests that the assumption of normality for the residuals is reasonable

## Plotting Residuals vs Predictors





- The residual vs. predictors plot for the given linear regression model shows a random scatter of points around the horizontal line at zero, indicating that the linearity assumption is met.
- The spread of the points appears to be approximately constant across all levels of the fitted values, suggesting that the assumption of homoscedasticity is met.
- There is no noticeable pattern or trend in the plot, indicating that the independence assumption is not violated.

### Comparison of Actual Vs Fitted,

```
> # COMPAIRISON OF ACTUAL VS FITTED
> # Split the data into training and testing sets
> set.seed(123) # for reproducibility
> train_index = createDataPartition(data$GDP, p = 0.8, list = FALSE)
> train_data = data[train_index, ]
> colnames(train_data)
[1] "GDP" "Import..." "Population"
[3] "Export..." "Minimum.Wage...month."
[5] "Unemployment.rate..." "Real.GDP.Growth..in.."
[7] "GDP.per.capita..in.." "Tax.total.rate..."
[9] "Gov.Expenditure...of.GDP"
> test_data = data[-train_index, ]
>
> # Train the linear regression model
> lm_model = lm(GDP ~ ., data = train_data)
>
> # Train the multiple linear regression model
> lm_model = lm(GDP ~ GDP.per.capita..in... + Real.GDP.Growth..in.. + Gov.Exp
enditure...of.GDP
+ Import... + Export... + Tax.total.rate... + Unemployment.ra
te...
+ Minimum.Wage...month. + Population, data = train_data)
>
> # Predict GDP for training data
> train_predictions = predict(lm_model, newdata = train_data)
>
> # Predict GDP for testing data
> test_predictions = predict(lm_model, newdata = test_data)
>
> # Compare actual GDP with predicted GDP for training data
> train_comparison = data.frame(Actual_GDP = train_data$GDP, Predicted_GDP =
train_predictions)
>
> # Compare actual GDP with predicted GDP for testing data
> test_comparison = data.frame(Actual_GDP = test_data$GDP, Predicted_GDP = te
st_predictions)
>
> # Output
>
> cat("\nTesting Data Comparison:\n")
```

Testing Data Comparison:

```

> print(head(test_comparison))
  Actual_GDP Predicted_GDP
1   10.281064      9.849134
2   10.184069     10.445883
6    9.237483      9.118459
13  10.586295     10.615463
18   9.274069      9.294846
30   9.297070      9.342214
>
> # Mean Absolute Error (MAE) for test data
> test_MAE = mean(abs(test_comparison$Actual_GDP - test_comparison$Predicted_
GDP))
>
> # Mean Squared Error (MSE) for test data
> test_MSE = mean((test_comparison$Actual_GDP - test_comparison$Predicted_GDP
)^2)
>
> # Root Mean Squared Error (RMSE) for test data
> test_RMSE = sqrt(test_MSE)
>
> # R-squared for test data
> test_Rsquared = summary(lm_model)$adj.r.squared
>
> cat("\nTest Data:\n")

Test Data:
> cat("Mean Absolute Error (MAE):", test_MAE, "\n")
Mean Absolute Error (MAE): 0.1353863
> cat("Mean Squared Error (MSE):", test_MSE, "\n")

```

### Mean Absolute Error (MAE):

- MAE measures the average magnitude of the errors in a set of predictions, without considering their direction.
- In this case, the MAE for the test data is approximately **0.135\$**.
- It means that, on average, the model's predictions are off by about **0.135\$** units of GDP.

### Mean Squared Error (MSE):

- MSE measures the average of the squares of the errors or deviations.
- In this case, the MSE for the test data is approximately **0.0294\$**.
- It means that, on average, the squared error of the model's predictions is about **0.0294\$** units of GDP.

### Root Mean Squared Error (RMSE):

- RMSE is the square root of the MSE and it measures the average magnitude of the errors in a set of predictions, giving a relatively high weight to large errors.
- In this case, the RMSE for the test data is approximately **0.1715\$**.

- It means that, on average, the model's predictions are off by about **0.1715\$** units of GDP.

These metrics give us a measure of how well the model is performing. Lower values of MAE, MSE, and RMSE indicate better performance of the model in predicting GDP.

### ➤ **Conclusions:**

- Understanding how accurate GDP predictions are is crucial for making informed decisions.
- Our analysis sheds light on the strengths and weaknesses of forecasting methods.
- It provides valuable insights for policymakers, economists, and investors.
- By learning from past predictions, we can improve future economic planning.
- Ultimately, this research helps us navigate the complex world of economics more effectively.
- Ultimately, by acknowledging the discrepancies between predicted and actual GDP values and leveraging the insights gained from this analysis, stakeholders can make more informed decisions and better prepare for the challenges and opportunities that lie ahead in an ever-changing economic environment.

### ➤ **Reference:**

- Dataset: We collected dataset from,  
(Ctrl+ click on the link to open the website)  
[India GDP Growth Rate 1961-2024 | MacroTrends](#)  
[Economy of India - Wikipedia](#)  
[Heatmap - Economic Indicators By Country \(tradingeconomics.com\)](#)  
[GDP per capita \(current US\\$\) | Data \(worldbank.org\)](#)
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introduction to Linear Regression Analysis. John Wiley & Sons.
- Software used:
  - I. R-Studio
  - II. Ms- Excel
  - III. Word Document