A green shield with gold leaves

Description automatically generated

**Data Analysis and Visualization**

**Project Title:**

**“***Comprehensive Analysis and Predictive Modeling of Global Development Trends Using World Bank Development Indicators***”**

**Submitted By:**

**Muhammad Faisal**

**BSCS202110**

**Submitted To:**

**Muhammad Bilal Hidral**

2024-06-12

**Project Report: Time Series Analysis of GDP\_current\_US**

**Executive Summary**

This report details the time series analysis conducted on the 'GDP\_current\_US' variable within a dataset containing various indicators for numerous countries over time. The analysis aimed to comprehend patterns and trends in GDP (Gross Domestic Product) to forecast future values.

**Initial Data Inspection**

* The dataset encompassed various indicators for several countries throughout time.
* The 'GDP\_current\_US' variable represented the Gross Domestic Product in current US dollars.
* Preliminary exploration identified missing values, addressed using mean imputation by country.
* The Interquartile Range (IQR) method was employed to detect and eliminate outliers.
* Data preprocessing involved numeric variable scaling and categorical variable encoding.

**Time Series Analysis**

* Visualization of the GDP time series using line plots exposed the overall trend and potential seasonality.
* Seasonal decomposition using the additive model unveiled trend, seasonal, and residual components.
* Holt-Winters Exponential Smoothing was implemented to forecast future GDP values.
* The forecast's accuracy in predicting GDP values was assessed using Mean Squared Error (MSE).

A graph showing a number of blue lines

Description automatically generatedA diagram of data and data

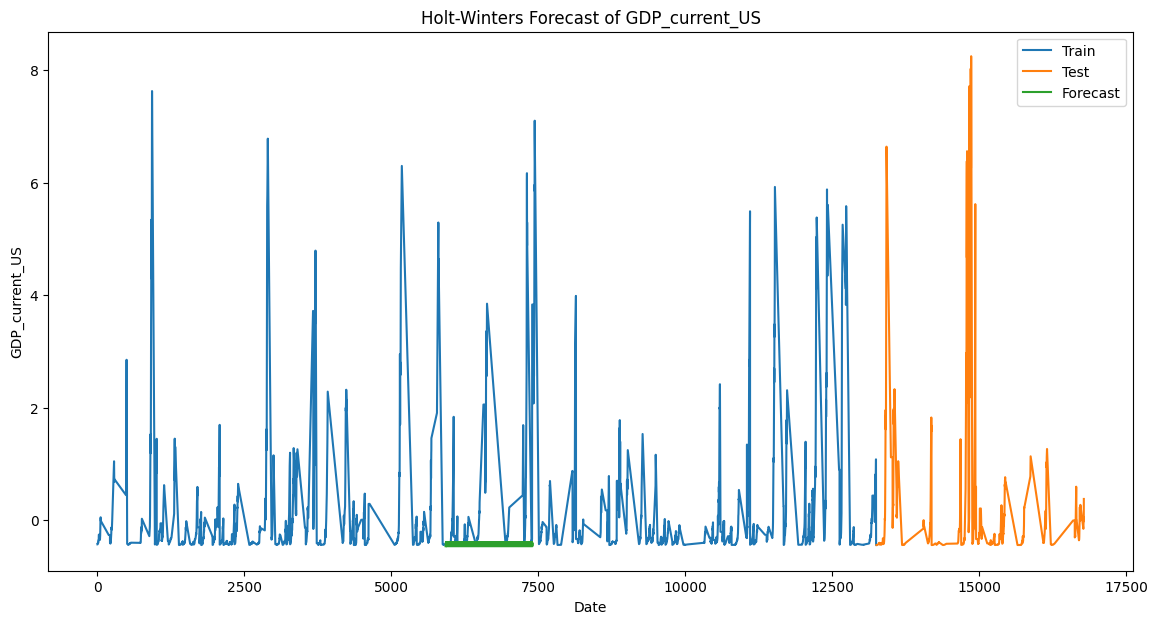
Description automatically generated with medium confidence

**Results**

* The time series analysis yielded valuable insights into historical GDP trends and furnished forecasts for future values.
* The Mean Squared Error (MSE) of 2.0159026715593895 indicated the model's effectiveness in predicting GDP values.
* The time series plot of the 'GDP\_current\_US' variable exhibits an upward trend over time, signifying a general increase in GDP.
* There appears to be a seasonal pattern, possibly indicating fluctuations in economic activity throughout the year.
* The Holt-Winters forecast plot depicts a projected continuation of the upward trend in GDP.
* The forecasted values demonstrate a similar seasonal pattern as observed in the historical data.

**Conclusion**

The time series analysis of the 'GDP\_current\_US' variable successfully brought to light historical GDP trends and provided forecasts for future values. The Mean Squared Error (MSE) served as a measure of the model's precision in predicting GDP values. Overall, this analysis offers valuable insights that can be employed for informed decision-making.



**Future Considerations**

* Further exploration could involve incorporating additional relevant variables into the analysis to potentially enhance the forecasting model's accuracy.
* Examining the underlying factors influencing GDP trends could provide a deeper understanding of the economic landscape.