Analysis of Quarterly GDP and Key Economic Indicators (V1)

Kaleb Jordan

2024-12-28

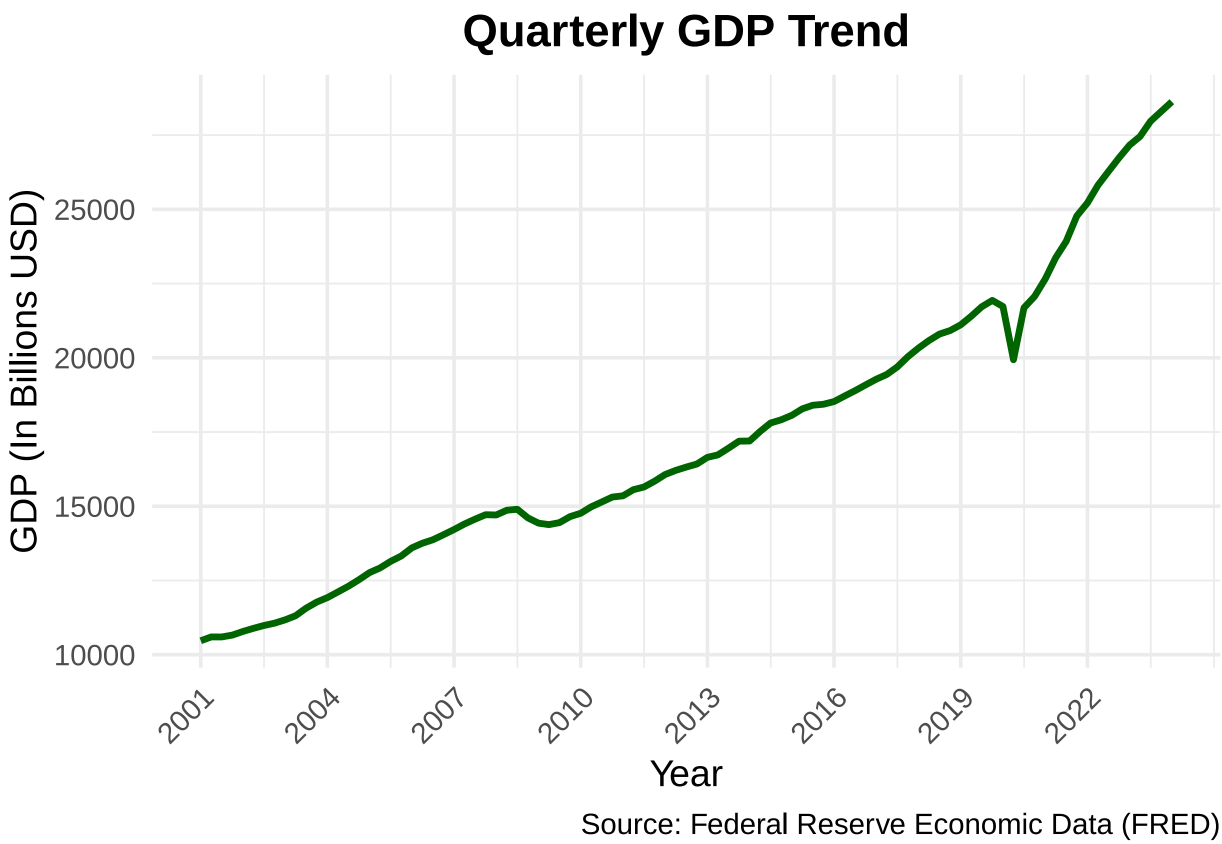
Abstract

This study investigates the quarterly United States GDP (Gross Domestic Product) from January 1, 2001 to January 1, 2024 using econometric modeling and time series analysis to forecast GDP over the next two years. Differencing and stationarity tests were applied, followed by ARIMA modeling, residual diagnostics, and forecast evaluation. Key findings suggest that ARIMA(0,2,1) is an applicable model for short-term (2-Year) GDP prediction, with the results highlighting potential implications for economic policymaking.

Introduction

GDP serves as a critical measure of a nation's economic health, providing insights into growth trends, productivity, and overall economic well-being. Accurate GDP forecasting is essential for policymakers and financial institutions to guide economic strategies. This report aims to utilize time series methods to predict future GDP values. Key macroeconomic indicators known to influence GDP are included to provide a comprehensive understanding of economic relationships and future dynamics.

The data set of United States GDP growth consists of 93 observations financial quarters from January 1, 2001-January 1, 2024. Nine other variables are in the data set and can be ignored for the ARIMA only model. *See Appendix for more details.*



From the graph, there is a strong linear increase in GDP over most of the observed period, indicating consistent economic growth. Two significant decreases are evident: the first in 2008 and the second in 2020.

- The decline in 2008 aligns with the global financial crisis and recession during the Bush (43) administration.
- The drop in 2020 corresponds to the COVID-19 pandemic and economic recession under the Trump administration.

The rising trend suggests non-stationary properties in the data. As a result, stationarity testing will be crucial before conducting time series analysis, with potential transformations (e.g., differencing or log-transformations) required to meet stationarity assumptions.

Methods

Data Collection

Quarterly GDP data from 2001–2024 were obtained alongside key macroeconomic variables such as unemployment rate, CPI growth, retail sales, job openings, and labor force participation. Data preprocessing included:

- Filtering quarterly data.
- Aggregating and renaming variables for clarity.
- Converting data into a time series format.

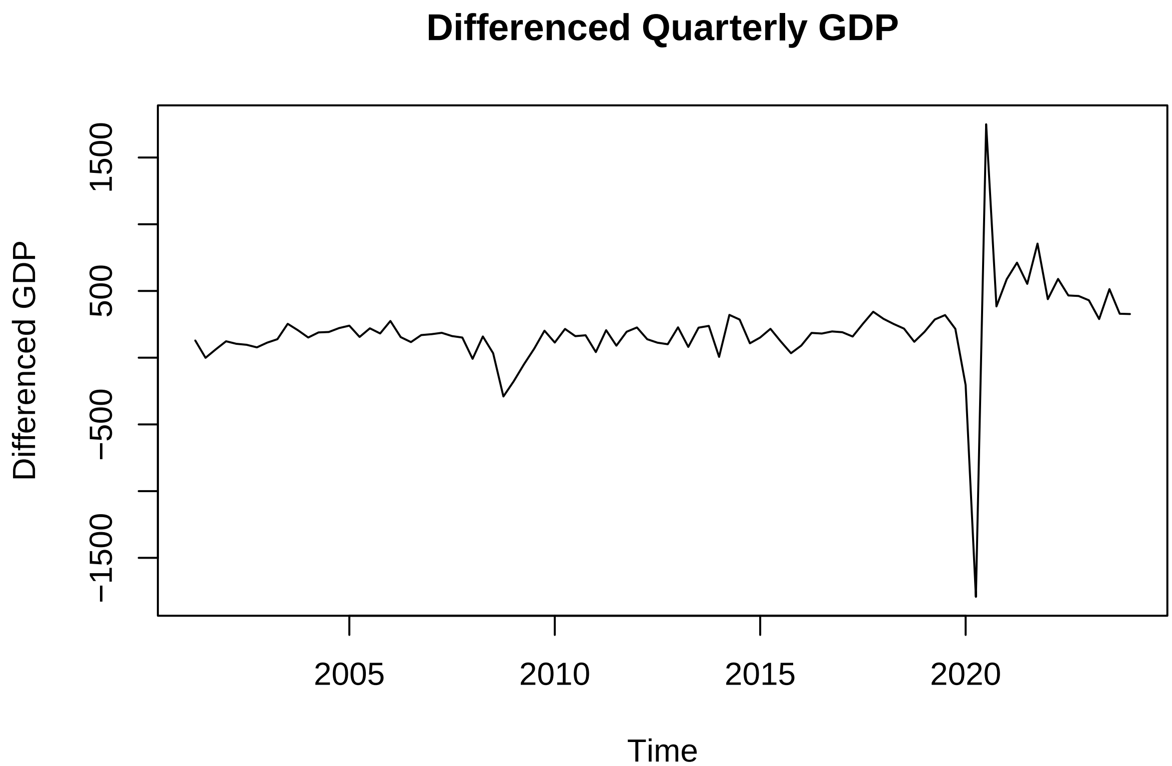
Stationarity Testing

The Augmented Dickey-Fuller (ADF) test was applied to check stationarity. Differencing was conducted where required to achieve stationarity, critical for time series modeling.

Stationary results of our overall data concluded that the data is non-stationary. A transformation is needed and concluded Differencing is the best solution. *For more information, check Appendix.*

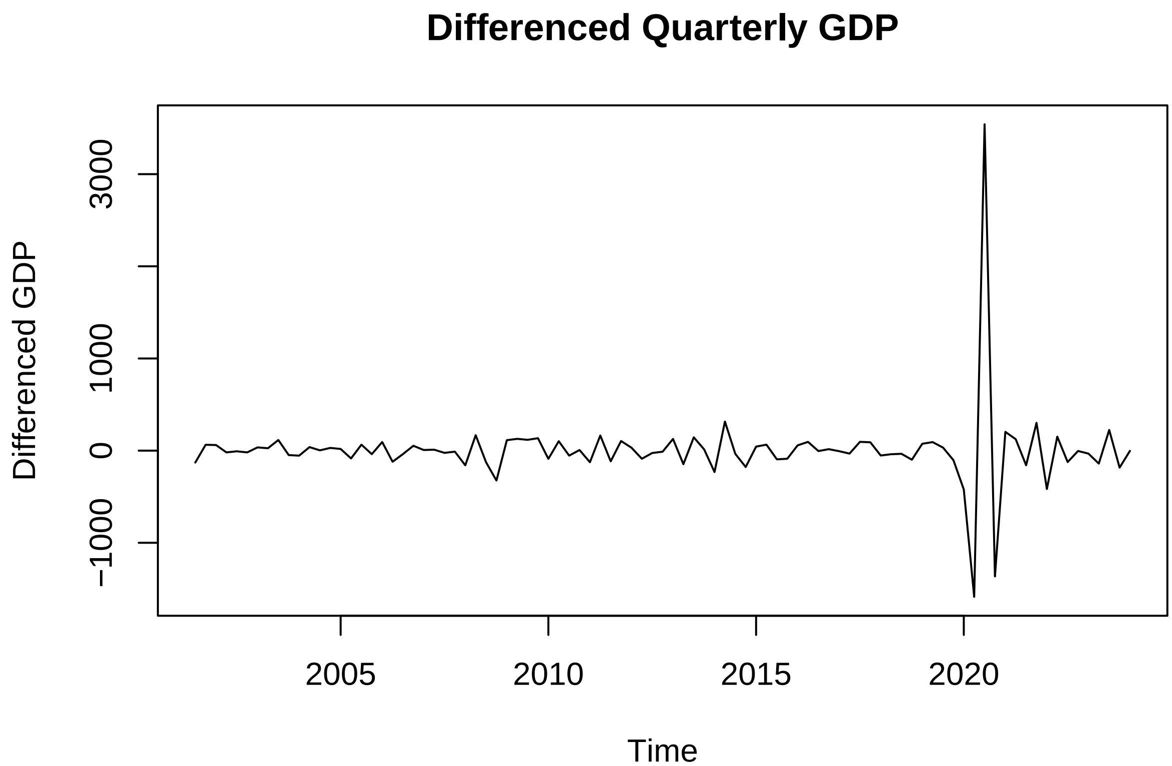
```
##
##  Augmented Dickey-Fuller Test
##
## data:  gdp_ts
## Dickey-Fuller = 0.41846, Lag order = 4, p-value = 0.99
## alternative hypothesis: stationary
```

After using differencing to transform the data, we conclude that our data is now stationary but has a p-value just less than 0.05. Consider a second differencing transformation. Here is the plot of differenced GDP values and ADF test results:



Here are the results of the second differencing transformation. We see a significantly lower p-value which ensures stationarity.

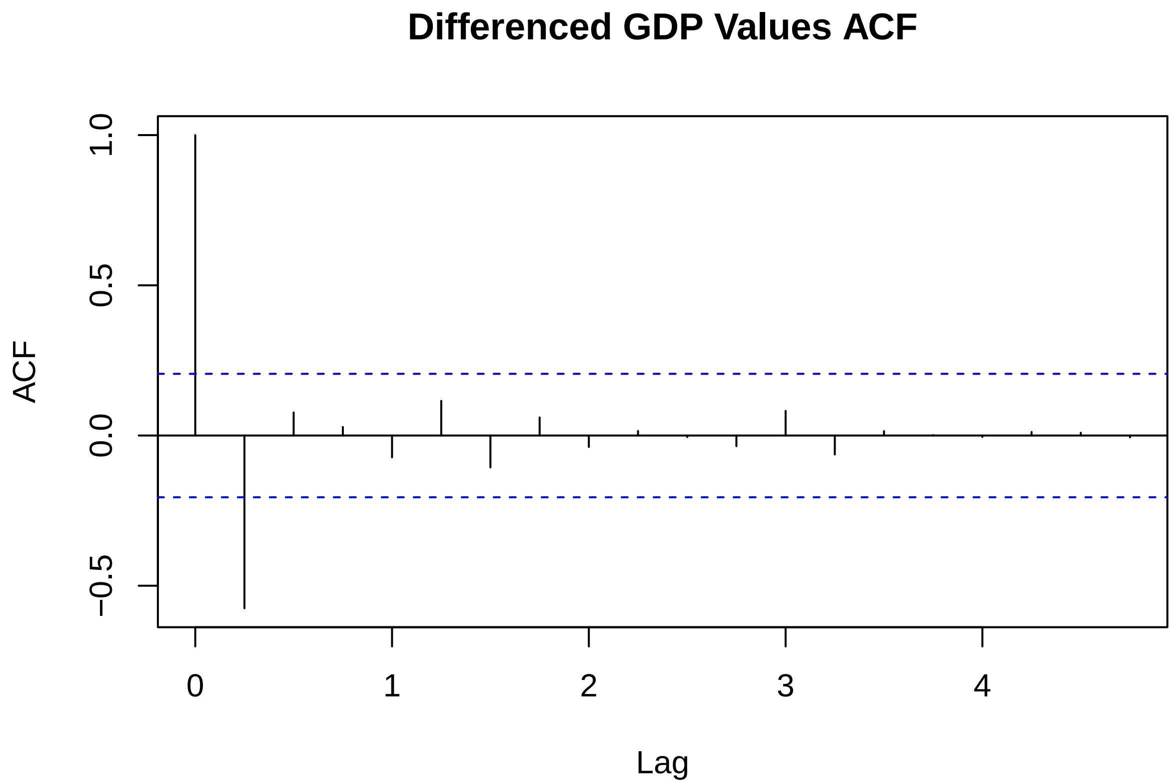
```
##
##  Augmented Dickey-Fuller Test
##
## data:  gdp_diff2
## Dickey-Fuller = -6.5063, Lag order = 4, p-value = 0.01
## alternative hypothesis: stationary
```

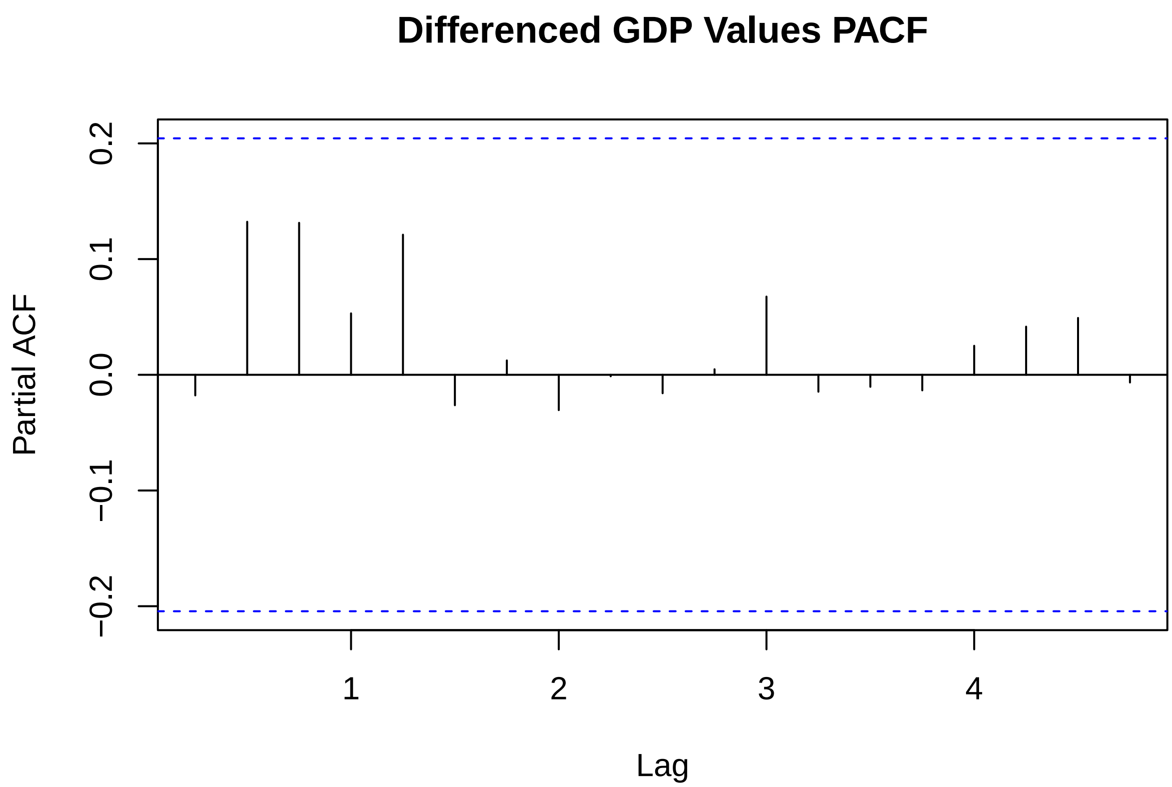

Differenced Quarterly GDP
Differenced GDP
3000
1000
0
−1000
2005
2010
2015
2020
Time

Model Selection

ACF/PACF Analysis

Identified potential autoregressive (AR) via PACF graph and moving average (MA) components via ACF graphs for the differenced values.





From the ACF graph, we see one significant lag spike which indicates that our ARIMA model will most likely conclude ARIMA(0, 2, 1).

ARIMA

Selected the optimal ARIMA model based on AIC (Akaike Information Criterion) by running an Auto-ARIMA model.

The Auto-ARIMA model confirms the model concluded by the differencing transformations and ACF/PACF graphs. *For more information on ARIMA models, see Appendix.*

Key insights include:

- ACF: 1309.07
- Root Mean Squared Error (RMSE): 308.11
- Mean Absolute Percentage Error (MAPE): 0.719%

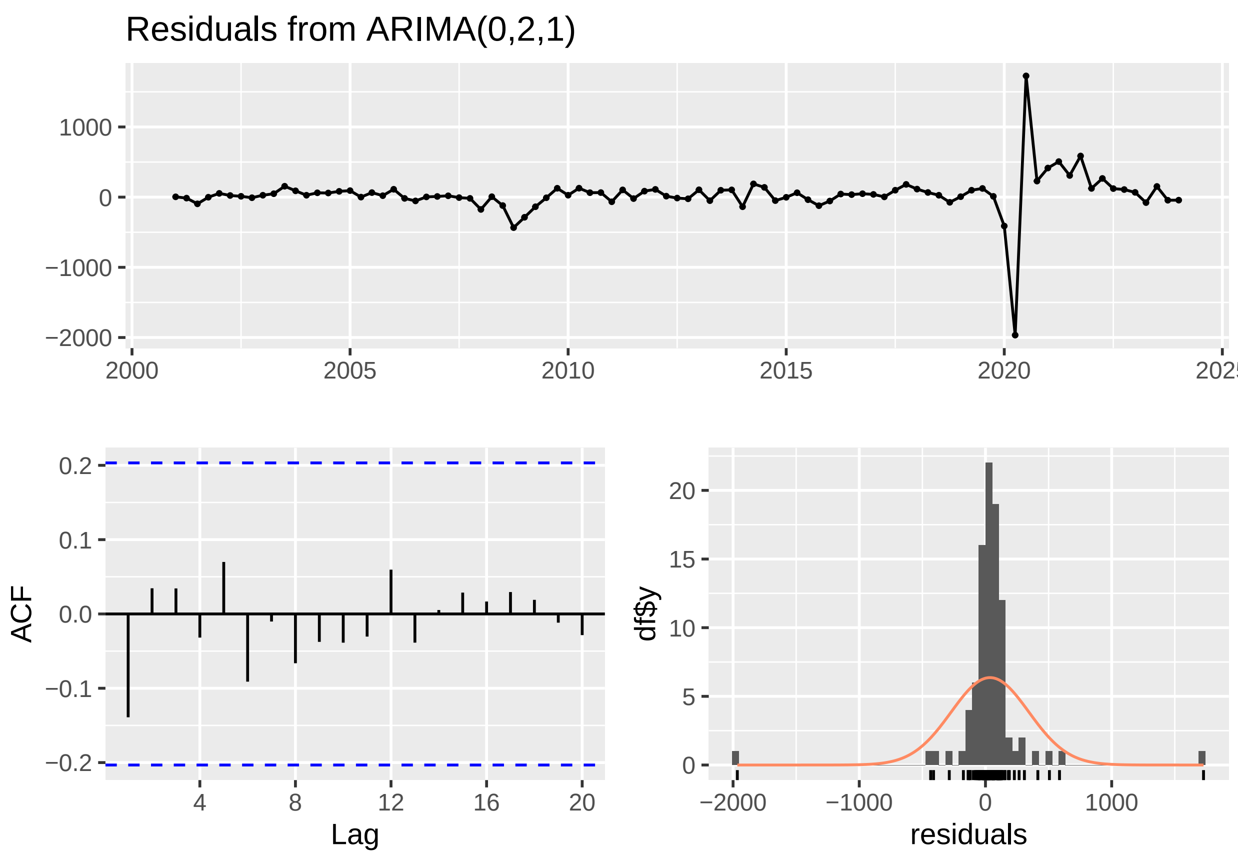
Forecasting

The chosen ARIMA model was used to predict GDP for the next eight quarters (two years). Forecast intervals (80% and 95%) were generated to assess prediction uncertainty.

Year	Quarter	Point Forecast	80% Lower Bound	80% Upper Bound	95% Lower Bound	95% Upper Bound
2024	2	28990.59	28589.21	29391.97	28376.73	29604.45
2024	3	29357.11	28767.05	29947.17	28454.69	30259.53
2024	4	29273.63	28973.22	30474.05	28575.97	30871.29
2025	1	30090.15	29191.32	30988.98	28715.51	31464.80
2025	2	30456.67	29415.31	31498.04	28864.04	32049.30
2025	3	30823.19	29642.23	32004.16	29017.06	32629.33
2025	4	31189.72	29870.41	32509.02	29172.01	33207.42
2026	1	31556.24	30098.83	33013.64	29327.33	33785.15

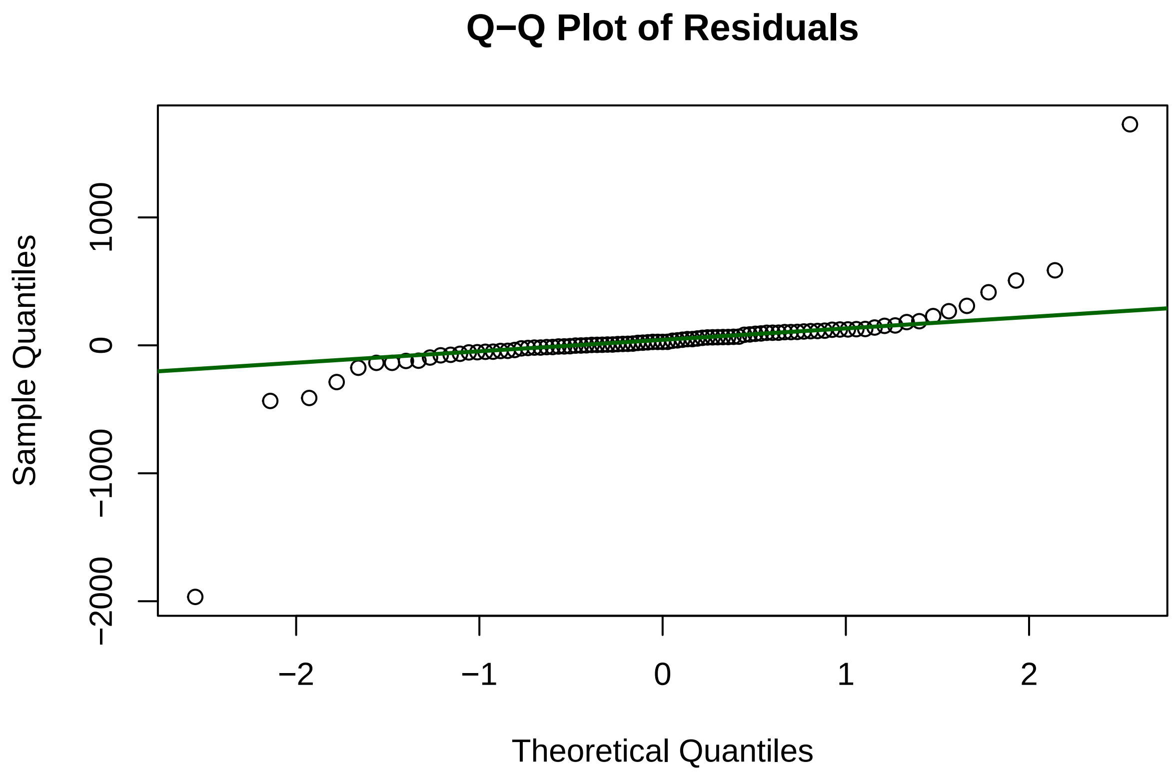
Residual Diagnostics

Residual Analysis

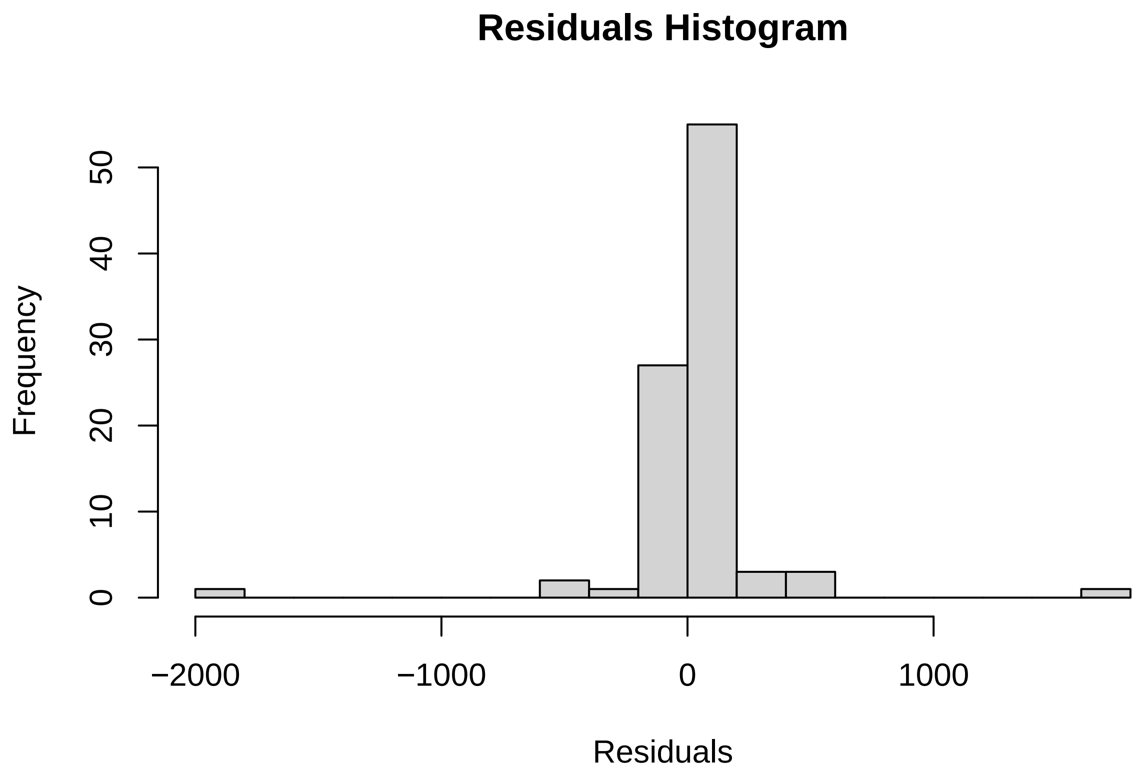



```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(0,2,1)
## Q* = 3.9896, df = 7, p-value = 0.781
##
## Model df: 1.   Total lags used: 8
```

Overall, the ARIMA(0,2,1) model performs well, as the residuals appear to be normally distributed and show minimal autocorrelation.



- While the residuals are broadly normal, the tail deviations suggest potential issues with capturing rare or extreme events in the data (e.g., anomalies around 2020).
- The model may still be reliable for forecasting general trends, but it might struggle to handle unexpected shocks.



- Residuals have a high frequency close to 0 which indicates good model performance.
- Large residuals are likely to do with extreme or unexpected events (e.g., anomalies around 2020).

Future Suggestions

- **Feature Expansion**: Include more diverse economic indicators and lagged variables for richer modeling.
- **Multivariate Models**: Apply multivariate time series techniques (e.g., VAR, VECM) to capture dynamic interactions.
- **Structural Break Analysis**: Investigate potential structural changes during economic events like financial crises.
- **Real-Time Updates**: Integrate real-time data streams to refine forecasts adaptively.

Appendix

Abstract

- All observations are found using data from the following:
  - Q1: January 1st Annually (2001-2024)
  - Q2: April 1st Annually (2001-2023)
  - Q3: July 1st Annually (2001-2023)
  - Q4: October 1st Annually (2001-2023)

Introduction

- Macroeconomic Indicators were not used in ARIMA model but will be included in future ARIMAX models for future versions of this report.

Methods

Stationarity Testing

- When conducting Augmented Dickey-Fuller (ADF) tests, null hypothesis is that the data is non-stationary and needs a transformation. Alternate hypothesis is that the data is stationary and no transformations are needed.
- When p-value < 0.05, reject the null hypothesis and conclude the data is stationary.
- When p-value > 0.05, accept the null hypothesis and conclude the data is non-stationary and a transformation is needed.

Model Selection

ARIMA

- ARIMA models combine Autoregressive observations, Differencing, and Moving Average terms.
- Autocorrelation functions (ACF) identify moving average components.
- Partial Autocorrelation functions (PACF) identify autoregressive components.

Forecasting

- Below is the code to understand how the forecasted values were generated:

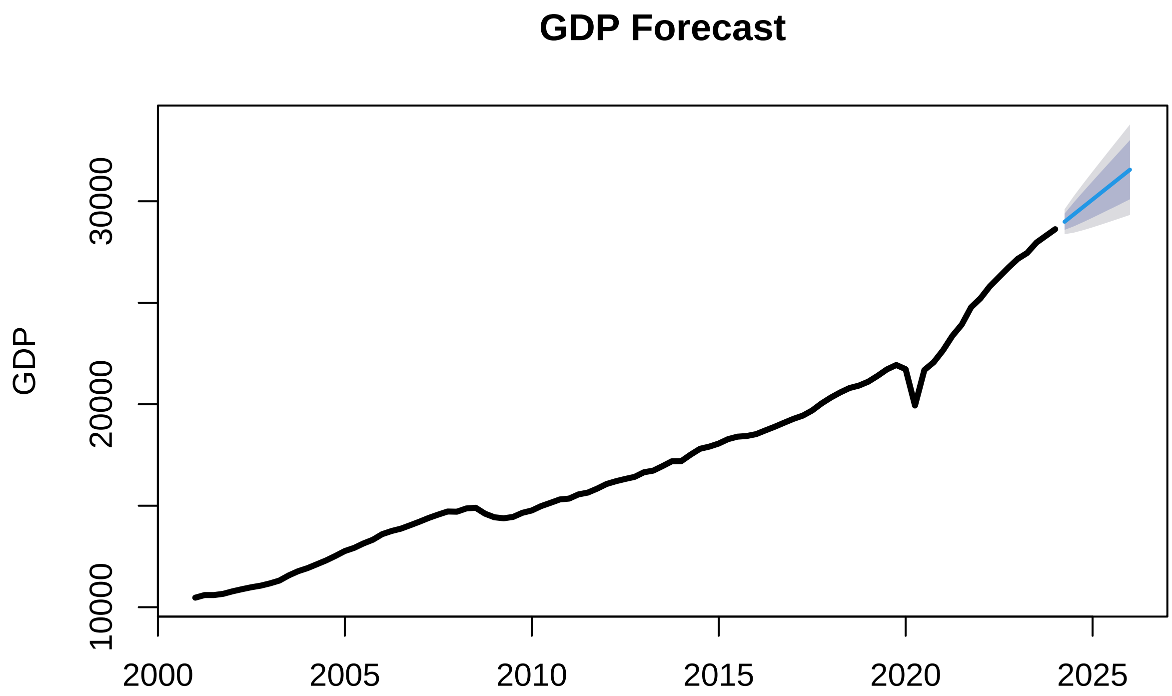
```r
# Set seed for reproducibility
set.seed(88)

# Use auto.arima model to confirm our model
arima_model = auto.arima(gdp_ts, seasonal = FALSE)
summary(arima_model)
```

```
## Series: gdp_ts
## ARIMA(0,2,1)
##
## Coefficients:
##           ma1
##       -0.9225
## s.e.   0.0423
##
## sigma^2 = 98095:  log likelihood = -652.53
## AIC=1309.07   AICc=1309.2   BIC=1314.09
##
## Training set error measures:
##                   ME     RMSE      MAE       MPE      MAPE      MASE       ACF1
## Training set 35.9848 308.1077 134.1904 0.1640608 0.7185238 0.1578936 -0.1390113
```

```r
# Forecast 8 quarters or two years of GDP growth for the future
forecast_values = forecast(arima_model, h = 8)
forecast_values
```

```
##         Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## 2024 Q2       28990.59 28589.21 29391.97 28376.73 29604.45
## 2024 Q3       29357.11 28767.05 29947.17 28454.69 30259.53
## 2024 Q4       29723.63 28973.22 30474.05 28575.97 30871.29
## 2025 Q1       30090.15 29191.32 30988.98 28715.51 31464.80
## 2025 Q2       30456.67 29415.31 31498.04 28864.04 32049.30
## 2025 Q3       30823.19 29642.23 32004.16 29017.06 32629.33
## 2025 Q4       31189.72 29870.41 32509.02 29172.01 33207.42
## 2026 Q1       31556.24 30098.83 33013.64 29327.33 33785.15
```

```r
# Plot forecasted values
plot(forecast_values, main = "GDP Forecast", ylab = "GDP", lwd = 3)
```




```r
forecast_years = c(rep(2024, 4), rep(2025, 4))
forecast_quarters = rep(1:4, times = 2)


forecast_results = data.frame(
  Year = forecast_years,
  Quarter = forecast_quarters,
  Forecasted_GDP = as.numeric(forecast_values$mean),
  Lower_80 = as.numeric(forecast_values$lower[, 1]),
  Upper_80 = as.numeric(forecast_values$upper[, 1]),
  Lower_95 = as.numeric(forecast_values$lower[, 2]),
  Upper_95 = as.numeric(forecast_values$upper[, 2])
)
# Print forecasted values (Same as table in Report)
forecast_results
```

```
##   Year Quarter Forecasted_GDP Lower_80 Upper_80 Lower_95 Upper_95
## 1 2024       1       28990.59 28589.21 29391.97 28376.73 29604.45
## 2 2024       2       29357.11 28767.05 29947.17 28454.69 30259.53
## 3 2024       3       29723.63 28973.22 30474.05 28575.97 30871.29
## 4 2024       4       30090.15 29191.32 30988.98 28715.51 31464.80
## 5 2025       1       30456.67 29415.31 31498.04 28864.04 32049.30
## 6 2025       2       30823.19 29642.23 32004.16 29017.06 32629.33
## 7 2025       3       31189.72 29870.41 32509.02 29172.01 33207.42
## 8 2025       4       31556.24 30098.83 33013.64 29327.33 33785.15
```

Source

All data used in this report were from the Federal Reserve Economic Data from the Federal Reserve Bank of St. Louis. All data was tailored to get a specific range of dates (January 1, 2001-January 1, 2024).

- **Gross Domestic Product (GDP)**
  Federal Reserve Bank of St. Louis. (2024). *Gross Domestic Product (GDP)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/GDP
- **Unemployment Rate (UNRATE)**
  Federal Reserve Bank of St. Louis. (2024). *Unemployment Rate (UNRATE)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/UNRATE
- **Job Openings (JTSJOL)**
  Federal Reserve Bank of St. Louis. (2024). *Job Openings: Total Nonfarm (JTSJOL)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/JTSJOL
- **Labor Force Participation Rate (CIVPART)**
  Federal Reserve Bank of St. Louis. (2024). *Labor Force Participation Rate (CIVPART)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/CIVPART
- **Consumer Price Index (CPI-U)**
  Federal Reserve Bank of St. Louis. (2024). *Consumer Price Index for All Items (CPALTT01USM657N)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/CPALTT01USM657N
- **Producer Price Index (PPI)**
  Federal Reserve Bank of St. Louis. (2024). *Producer Price Index for All Commodities (PPIACO)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/PPIACO
- **Sticky Price CPI**
  Federal Reserve Bank of St. Louis. (2024). *Sticky Price Consumer Price Index (CORESTICKM159SFRBATL)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/CORESTICKM159SFRBATL
- **Personal Consumption Expenditures (PCE)**
  Federal Reserve Bank of St. Louis. (2024). *Real Personal Consumption Expenditures (DPCCRV1Q225SBEA)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/DPCCRV1Q225SBEA
- **Retail Sales (RSXFS)**
  Federal Reserve Bank of St. Louis. (2024). *Advance Retail Sales: Retail and Food Services, Excluding Motor Vehicles and Parts (RSXFS)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/RSXFS
- **University of Michigan Consumer Sentiment Index (UMCSENT)**
  Federal Reserve Bank of St. Louis. (2024). *University of Michigan: Consumer Sentiment (UMCSENT)* [Dataset]. FRED. Retrieved from https://fred.stlouisfed.org/series/UMCSENT