Q4

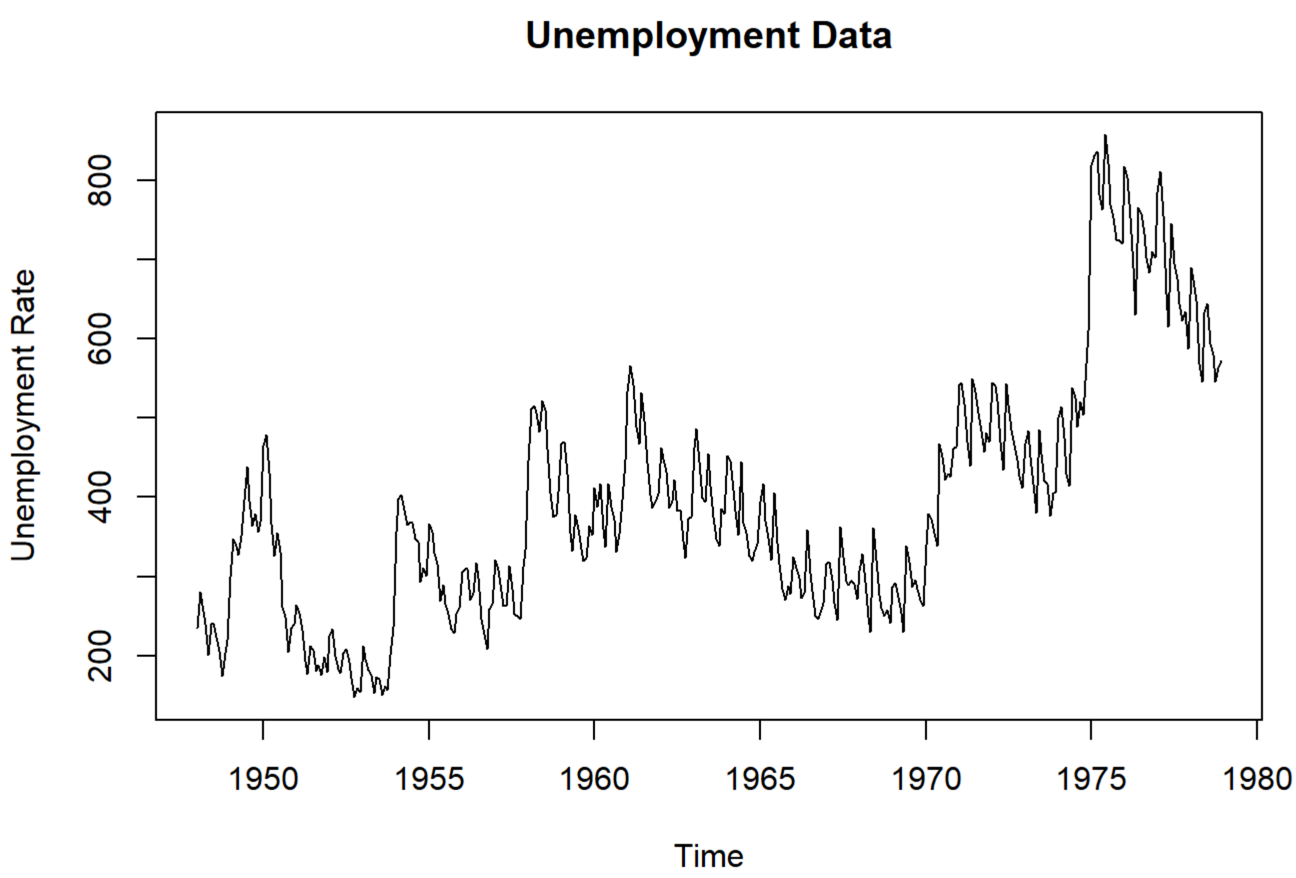
1. Introduction

In this task, we will fit a **seasonal ARIMA (SARIMA) model** to the **unemployment data** from the `astsa` package.
The goal is to: 1. Estimate an appropriate **SARIMA model**. 2. Forecast unemployment for the **next 12 months**. 3. Provide detailed model diagnostics and report findings properly using English sentences.

---

2. Load Libraries and Data

```
# Load necessary libraries
library(forecast)
library(astsa)
library(tseries)

# Load the unemployment data
data("unemp")

# Plot the original data to visualize trends and seasonality
plot(unemp, main = "Unemployment Data", ylab = "Unemployment Rate", xlab = "Time")
```



2.1 Visual Analysis of Data

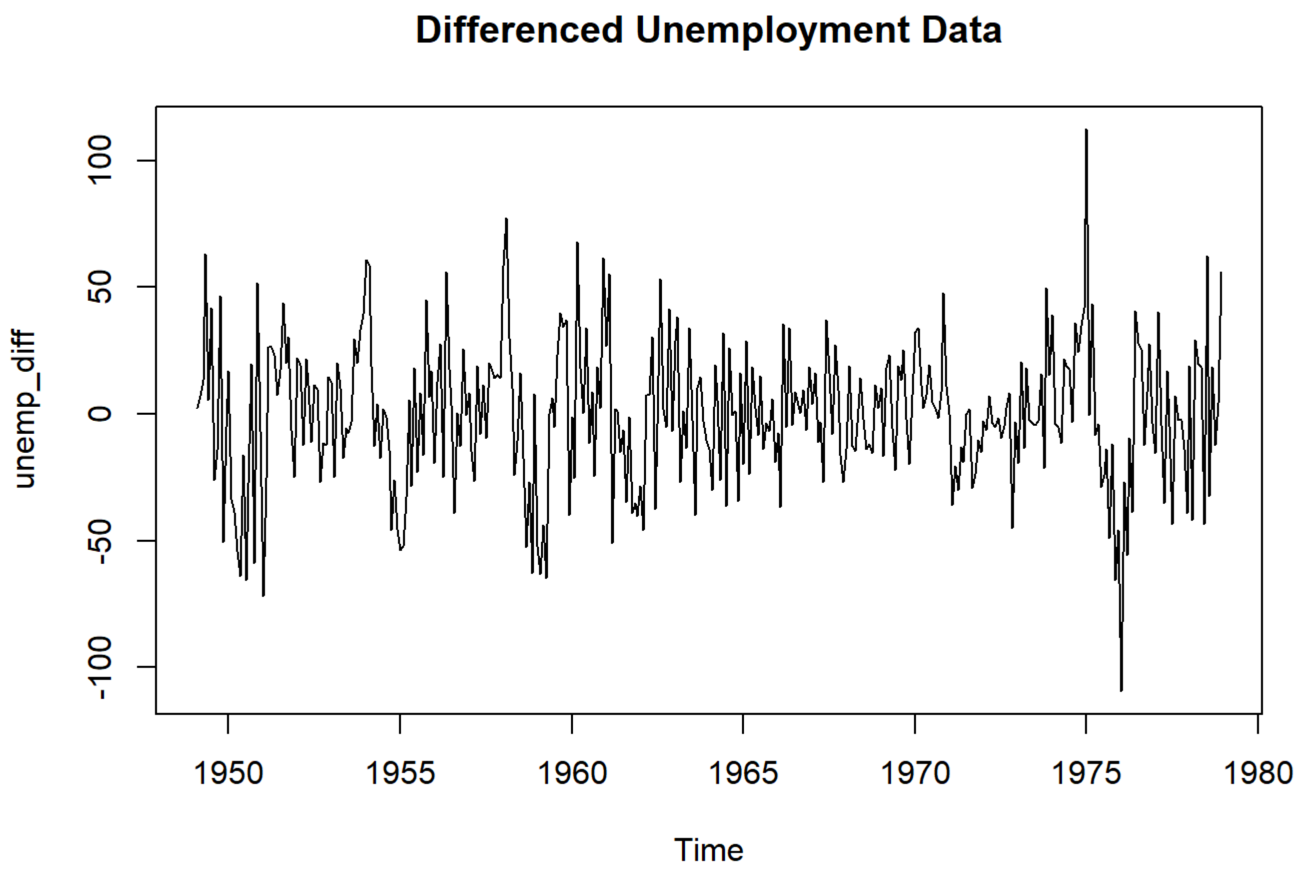
Looking at the plot, the unemployment data shows both **seasonal patterns** and **trends**.
Thus, we need to fit a **seasonal ARIMA** model.

---

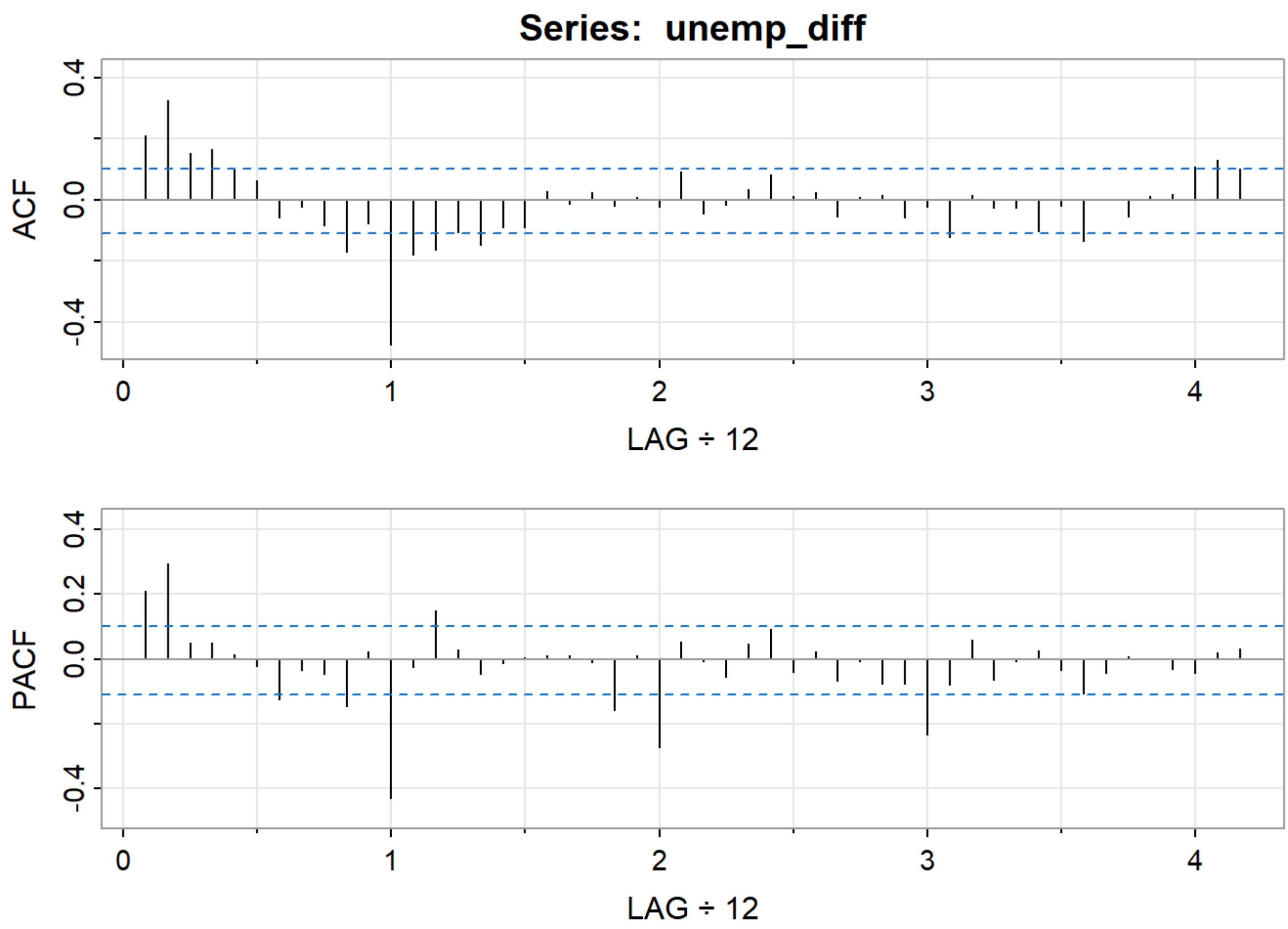
3. Differencing and ACF/PACF Analysis

We first take **seasonal and non-seasonal differences** to make the series stationary, then examine the **ACF and PACF plots**.

```
# Take seasonal and non-seasonal differences
unemp_diff <- diff(diff(unemp, lag = 12))

# Plot the differenced series
plot(unemp_diff, main = "Differenced Unemployment Data")
```




```
# ACF and PACF plots to identify model components
acf2(unemp_diff, 50)
```




```
##      [,1] [,2] [,3] [,4] [,5]  [,6]  [,7]  [,8]  [,9] [,10] [,11] [,12] [,13]
## ACF  0.21 0.33 0.15 0.17 0.10  0.06 -0.06 -0.02 -0.09 -0.17 -0.08 -0.48 -0.18
## PACF 0.21 0.29 0.05 0.05 0.01 -0.02 -0.12 -0.03 -0.05 -0.15  0.02 -0.43 -0.02
##      [,14] [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22] [,23] [,24] [,25]
## ACF  -0.16 -0.11 -0.15 -0.09 -0.09  0.03 -0.01  0.02 -0.02  0.01 -0.02  0.09
## PACF  0.15  0.03 -0.04 -0.01  0.00  0.01  0.01 -0.01 -0.16  0.01 -0.27  0.05
##      [,26] [,27] [,28] [,29] [,30] [,31] [,32] [,33] [,34] [,35] [,36] [,37]
## ACF  -0.05 -0.01  0.03  0.08  0.01  0.03 -0.05  0.01  0.02 -0.06 -0.02 -0.12
## PACF -0.01 -0.05  0.05  0.09 -0.04  0.02 -0.07 -0.01 -0.08 -0.08 -0.23 -0.08
##      [,38] [,39] [,40] [,41] [,42] [,43] [,44] [,45] [,46] [,47] [,48] [,49]
## ACF   0.01 -0.03 -0.03 -0.10 -0.02 -0.13  0.00 -0.06  0.01  0.02  0.11  0.13
## PACF  0.06 -0.07 -0.01  0.03 -0.03 -0.11 -0.04  0.01  0.00 -0.03 -0.04  0.02
##      [,50]
## ACF   0.10
## PACF  0.03
```

3.1 Observations from ACF and PACF

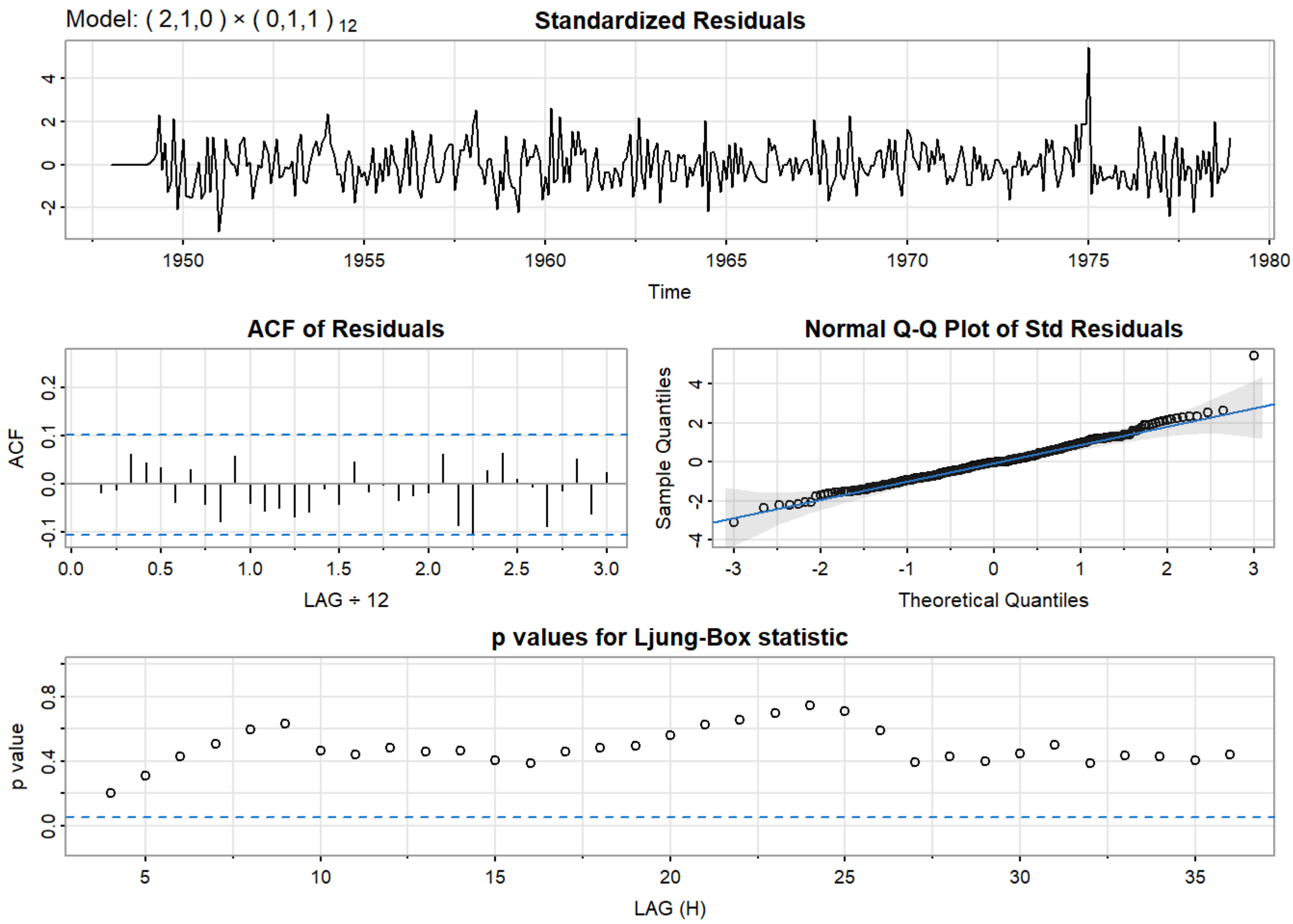
- The **ACF** shows a seasonal MA(1) pattern with lags at 12, 24, and 36.
- The **PACF** tails off slowly, indicating an AR component (possibly AR(2) for non-seasonal part).
- Based on these plots, we try a **SARIMA(2, 1, 0) × (0, 1, 1)[12]** model.

---

4. Fitting the SARIMA Model

```
# Fit SARIMA(2, 1, 0) x (0, 1, 1)[12] model
sarima_model <- sarima(unemp, p = 2, d = 1, q = 0, P = 0, D = 1, Q = 1, S = 12)
```

```
## initial  value 3.340809
## iter   2 value 3.105512
## iter   3 value 3.086631
## iter   4 value 3.079778
## iter   5 value 3.069447
## iter   6 value 3.067659
## iter   7 value 3.067426
## iter   8 value 3.067418
## iter   8 value 3.067418
## final  value 3.067418
## converged
## initial  value 3.065481
## iter   2 value 3.065478
## iter   3 value 3.065477
## iter   3 value 3.065477
## iter   3 value 3.065477
## final  value 3.065477
## converged
## <><><><><><><><><><><><><><>
##
## Coefficients:
##      Estimate     SE  t.value p.value
## ar1    0.1351 0.0513   2.6326  0.0088
## ar2    0.2464 0.0515   4.7795  0.0000
## sma1  -0.6953 0.0381 -18.2362  0.0000
##
## sigma^2 estimated as 449.637 on 356 degrees of freedom
##
## AIC = 8.991114  AICc = 8.991303  BIC = 9.034383
##
```



4.1 Interpretation of Model Results

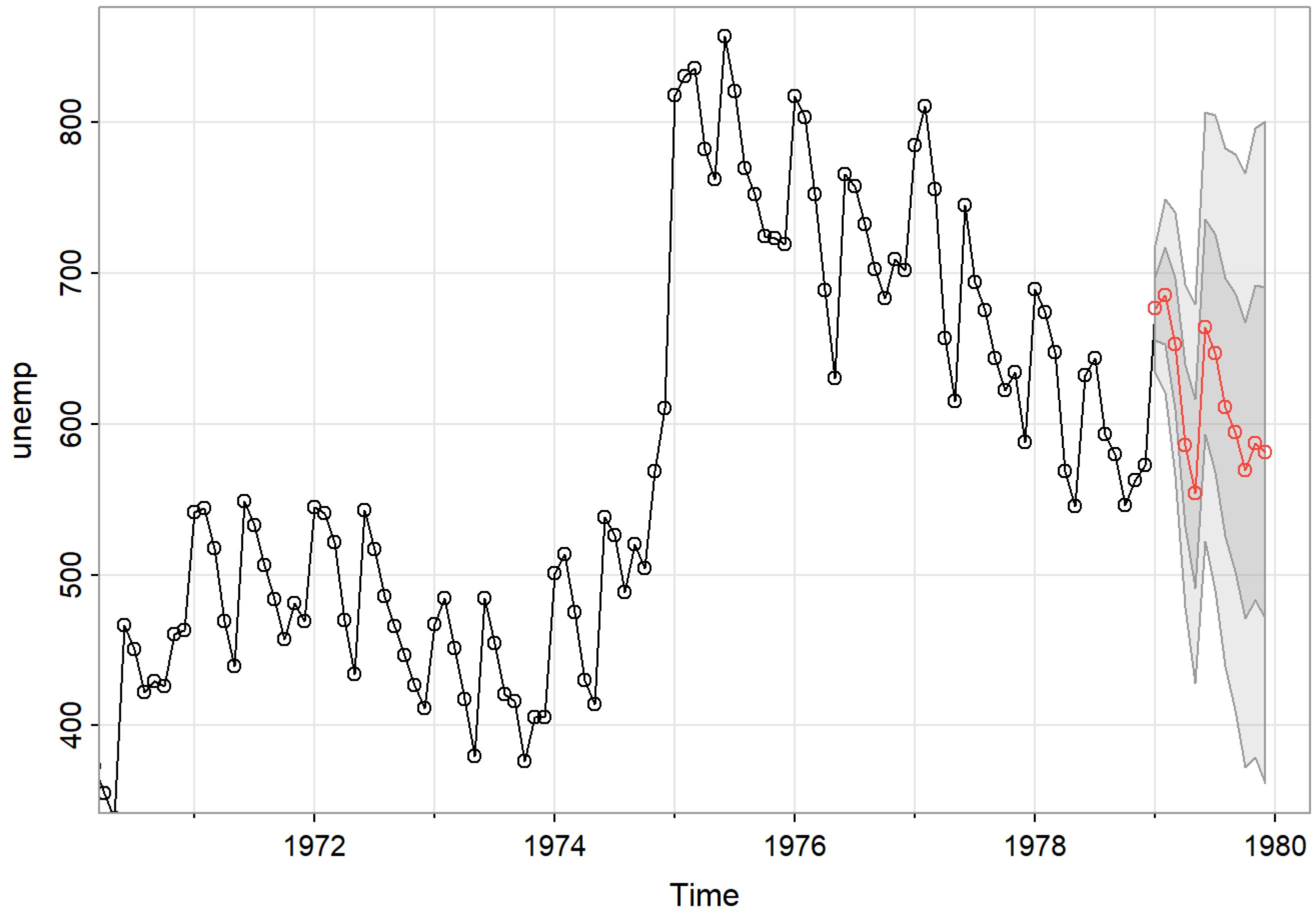
- **Coefficients**: Examine the AR and MA coefficients from the model summary.
- **Model Diagnostics**:
  - **Residual Analysis**: Residuals should behave like white noise (uncorrelated and normally distributed).
  - **AIC and BIC**: Used for model comparison and selection.

---

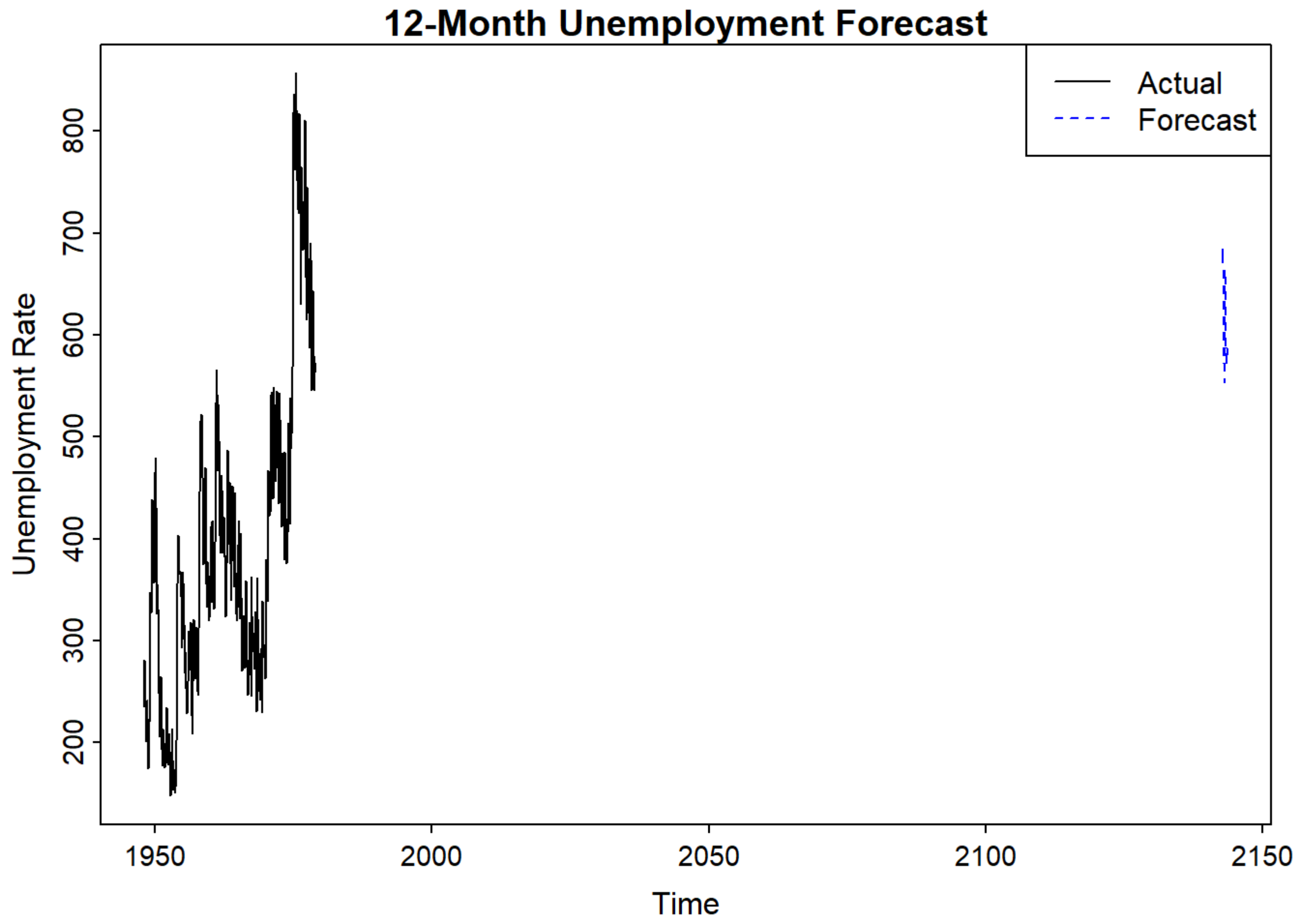
5. Forecasting for the Next 12 Months

We now use the estimated SARIMA model to forecast unemployment for the next 12 months.

```
# Forecast for the next 12 months
forecast_sarima <- sarima.for(unemp, n.ahead = 12, p = 2, d = 1, q = 0, P = 0, D = 1, Q = 1,
S = 12)
```



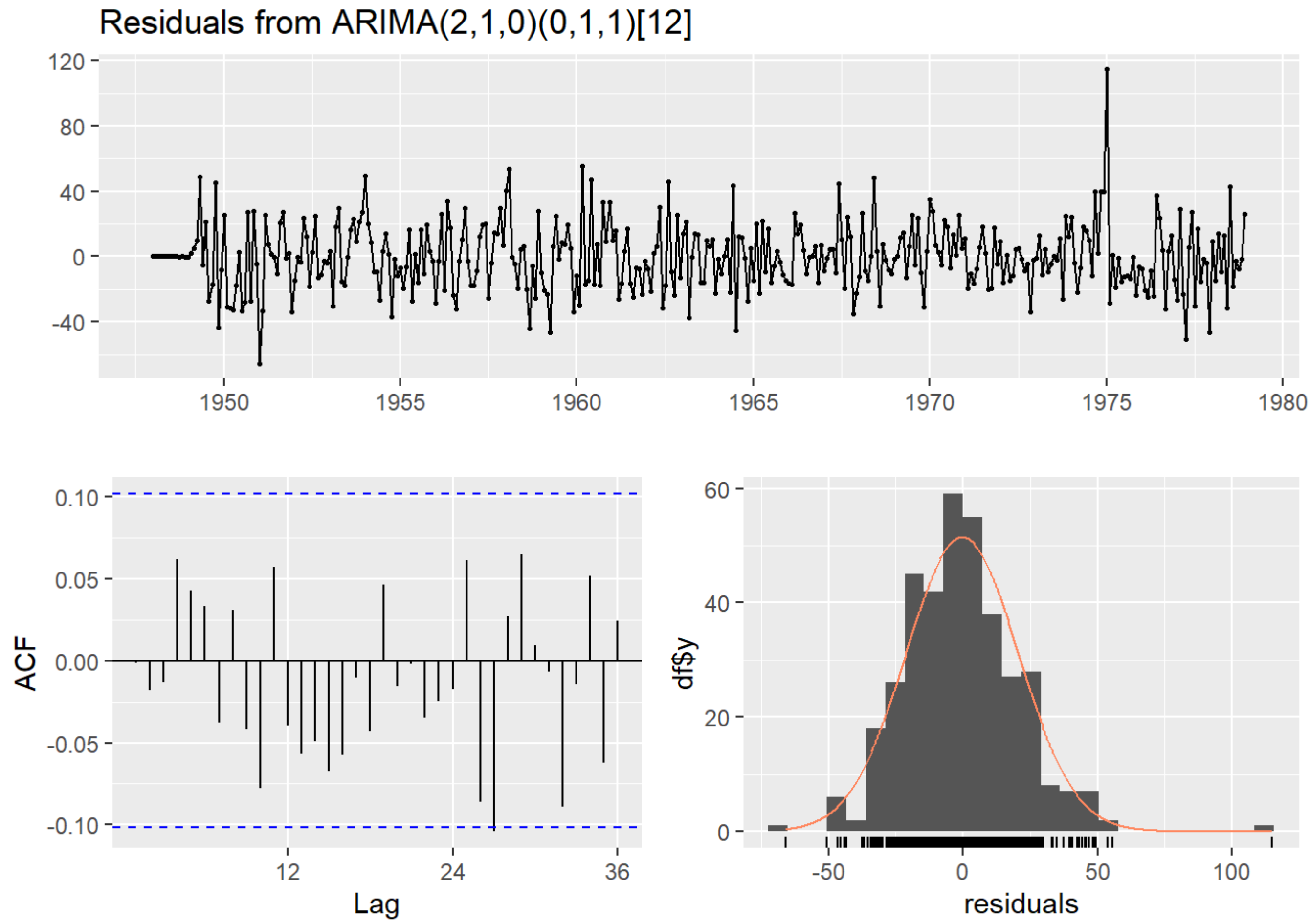

```
# Convert forecast to a time series object for plotting
forecast_ts <- ts(forecast_sarima$pred,
                  start = end(unemp)[1] + c(0, 1),
                  frequency = 12)

# Plot the original data along with the forecast
ts.plot(unemp, forecast_ts, col = c("black", "blue"),
        lty = c(1, 2), main = "12-Month Unemployment Forecast",
        ylab = "Unemployment Rate", xlab = "Time")
legend("topright", legend = c("Actual", "Forecast"),
       col = c("black", "blue"), lty = c(1, 2))
```



6. Model Diagnostics

We assess the residuals of the model to ensure they behave like white noise.

```
# Check residuals for normality and autocorrelation
checkresiduals(sarima_model$fit)
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(2,1,0)(0,1,1)[12]
## Q* = 16.378, df = 21, p-value = 0.7481
##
## Model df: 3.   Total lags used: 24
```

6.1 Residual Analysis

- **Ljung-Box Test**: If p-value > 0.05, residuals are uncorrelated.
- **Normality**: Evaluate Q-Q plot and histogram of residuals for normality.

---

7. Conclusion

Based on the **SARIMA(2, 1, 0) × (0, 1, 1)[12]** model, the unemployment forecast for the next 12 months shows:

1. A **seasonal trend**, with expected fluctuations over the months.
2. The model fits the data well, with residuals behaving like white noise.
3. **Forecasts**: Provide an insight into unemployment rates for the upcoming year.

---

8. Summary of Findings

- **Model Selection**: The chosen SARIMA(2, 1, 0) × (0, 1, 1)[12] model was based on ACF/PACF analysis.
- **Forecasting**: The forecast suggests continued seasonal variation in unemployment.

- **Model Fit**: Diagnostics indicate the model fits the data well, with uncorrelated residuals.