Practical–8 : AirPassengers Time Series Analysis

Practical–8 : AirPassengers Time Series Analysis

Praveer Raj
Roll: 1
Reg. No.: 230957002

---

1. Title

Time Series Analysis of AirPassengers Data

---

2. Objective

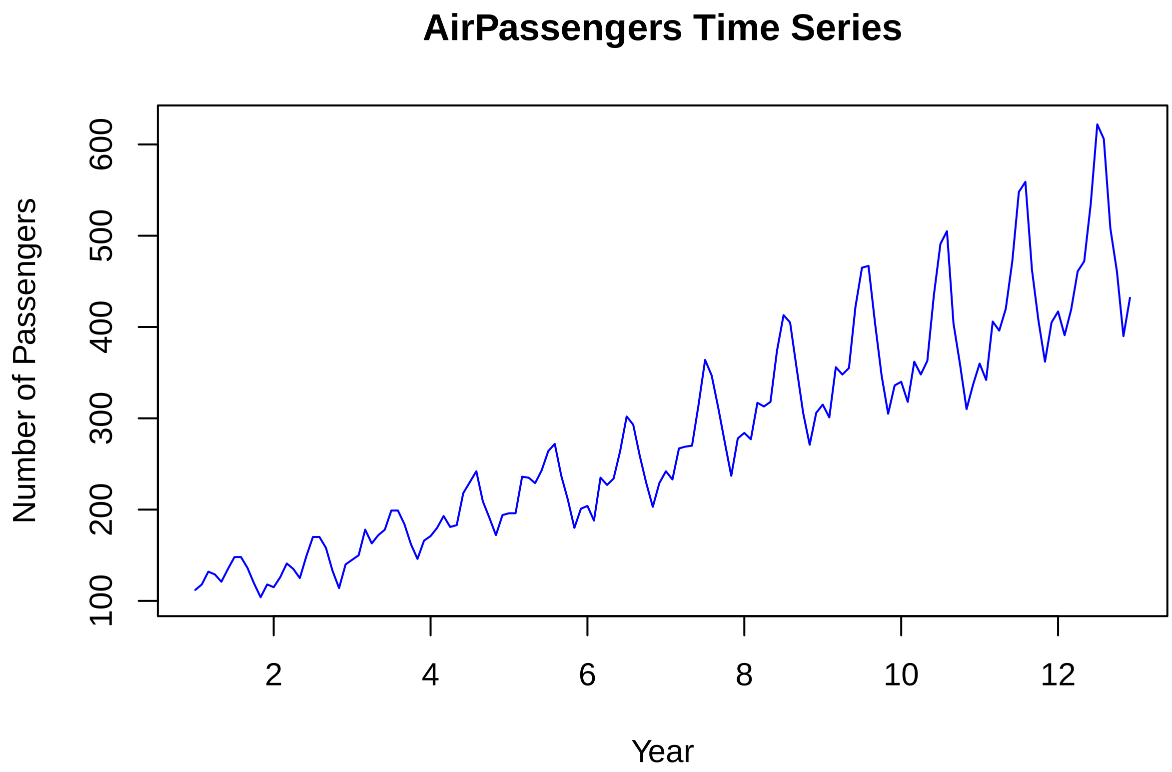
Practical-8:-

Consider the "AirPassengers" data from R library and write the R program for the following: (f) Convert the data into a time series object. (g) Plot the data to identify the dominant component. (h) Decompose the data to observe the dominating components more clearly. (i) Check stationarity or non-stationarity using ACF/PACF plot. (j) Check stationarity or non-stationarity using the KPSS test. (k) If data is non-stationary, make it stationary using an appropriate operator. (l) Based on the dominating component, select the suitable technique to fit the data. (m) Fit the data using the selected model and estimate the parameters of the model. (n) Check the goodness of fit of the model.

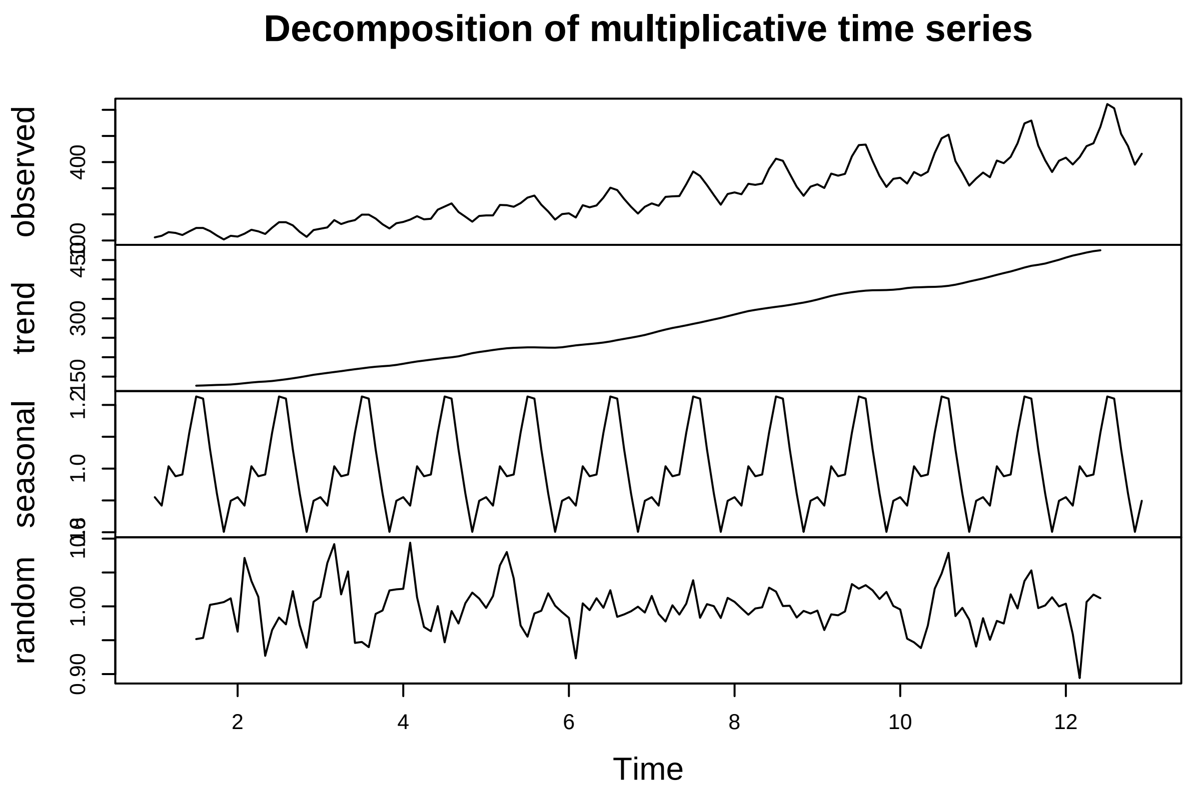
---

3. R Code

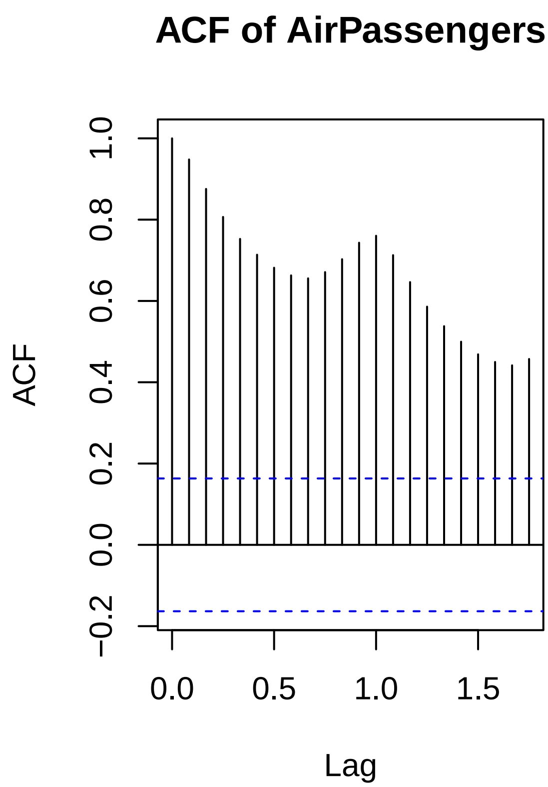
```r
############################################################
# Practical-8 : AirPassengers Time Series Analysis
############################################################

# (f) Convert the data into a time series object
data("AirPassengers")
ap_ts <- ts(AirPassengers, frequency = 12)

# (g) Plot the data to identify the dominant component
plot(ap_ts,
     main = "AirPassengers Time Series",
     xlab = "Year",
     ylab = "Number of Passengers",
     col = "blue")
```

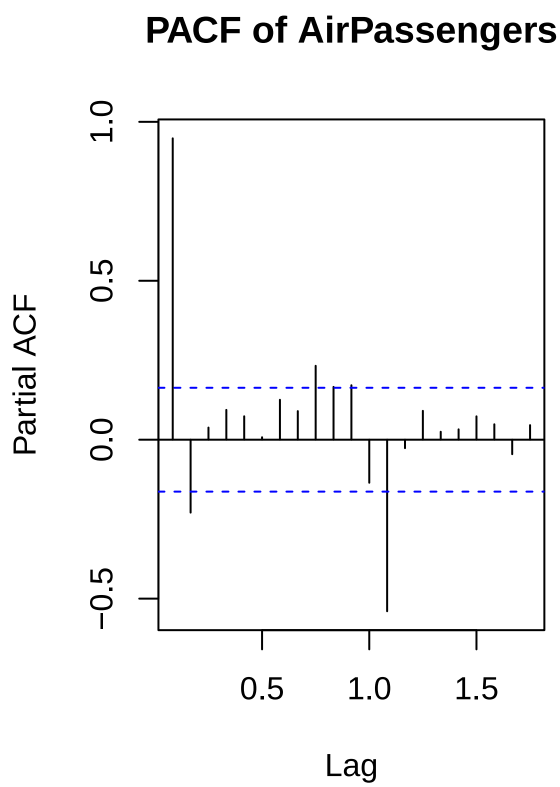



```r
# (h) Decompose the data to observe the dominating components
ap_decomp <- decompose(ap_ts, type = "multiplicative")
plot(ap_decomp)
```




```r
# (i) Check stationarity using ACF and PACF plots
par(mfrow = c(1, 2))
acf(ap_ts, main = "ACF of AirPassengers")
pacf(ap_ts, main = "PACF of AirPassengers")
```






```r
par(mfrow = c(1, 1))

# (j) Check stationarity using KPSS test
library(tseries)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

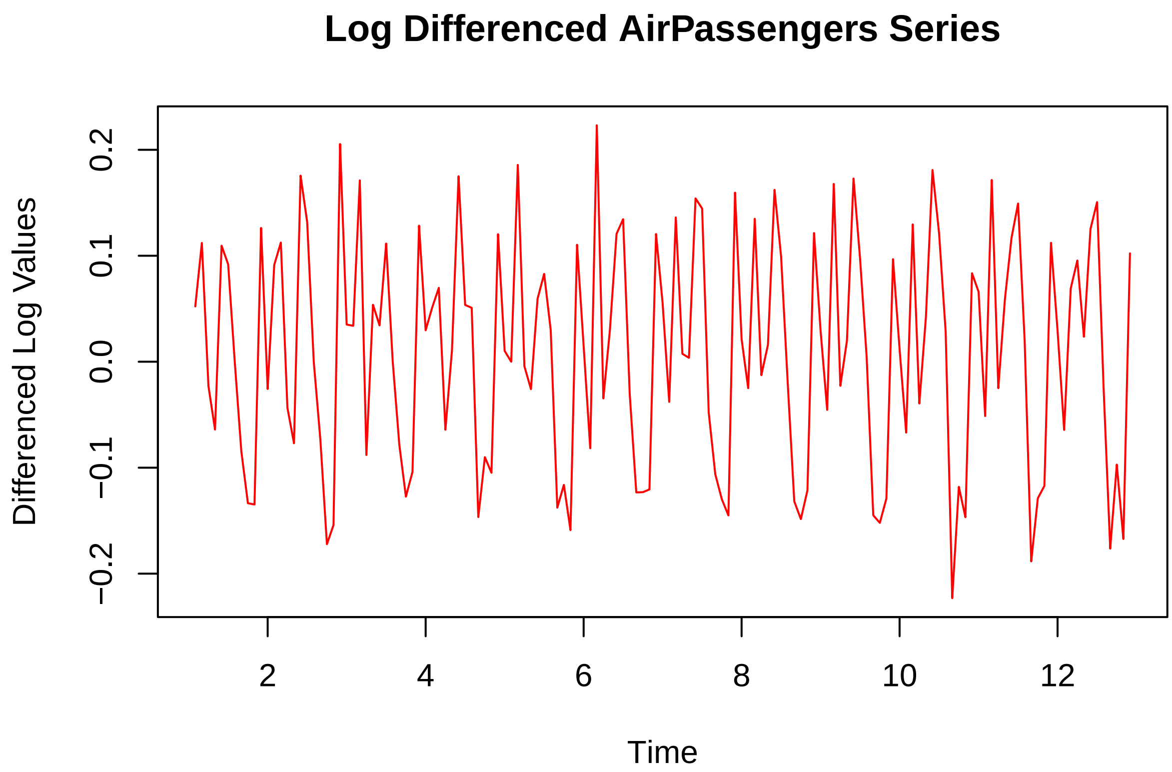
```r
kpss.test(ap_ts)
```

```
## Warning in kpss.test(ap_ts): p-value smaller than printed p-value
```

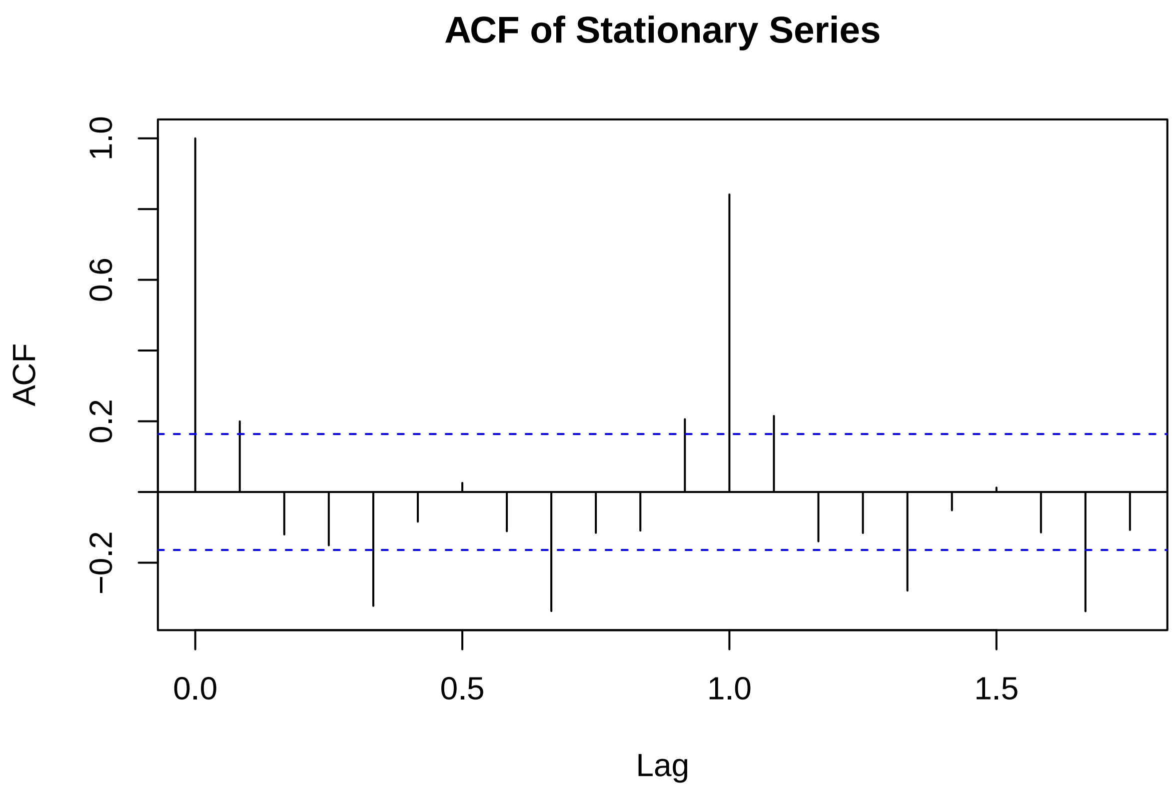
```
##
##  KPSS Test for Level Stationarity
##
## data:  ap_ts
## KPSS Level = 2.7395, Truncation lag parameter = 4, p-value = 0.01
```

```r
# (k) Make the data stationary using appropriate operator
ap_diff <- diff(log(ap_ts))

plot(ap_diff,
     main = "Log Differenced AirPassengers Series",
```

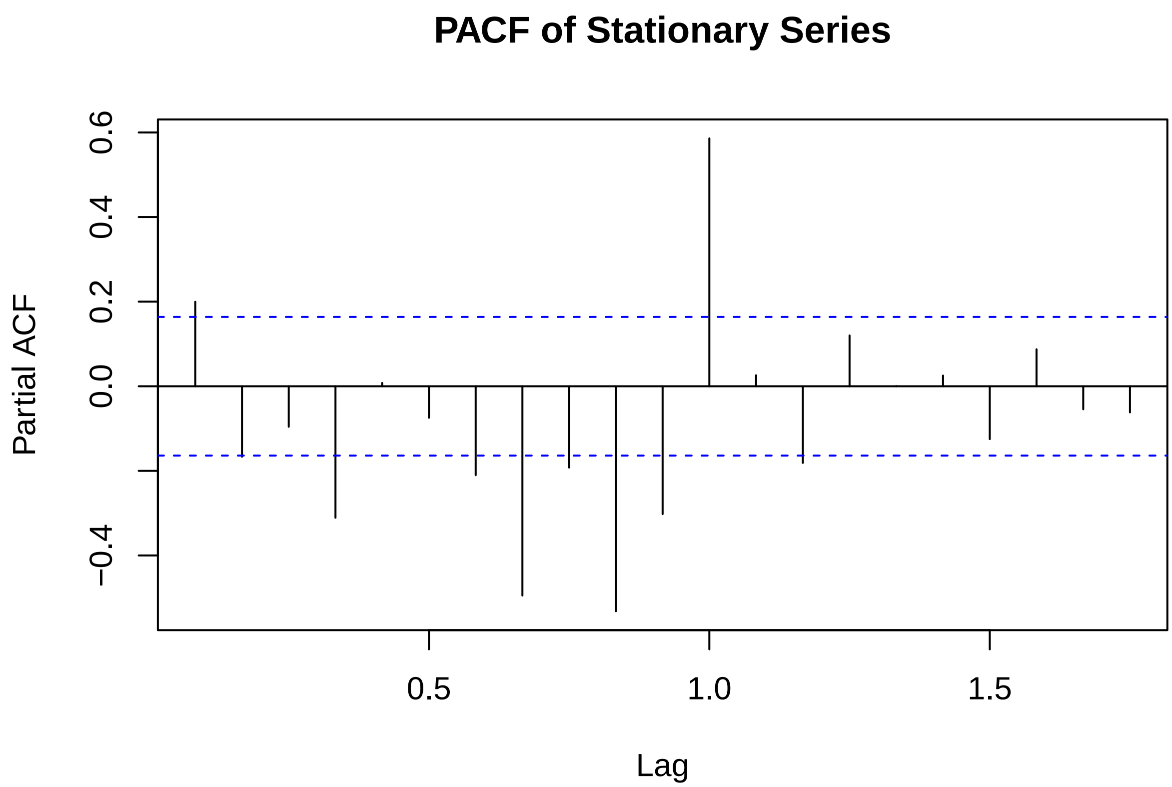
```r
    ylab = "Differenced Log Values",
    col = "red")
```




```r
acf(ap_diff, main = "ACF of Stationary Series")
```




```
pacf(ap_diff, main = "PACF of Stationary Series")
```




```
kpss.test(ap_diff)
```

```
## Warning in kpss.test(ap_diff): p-value greater than printed p-value
```

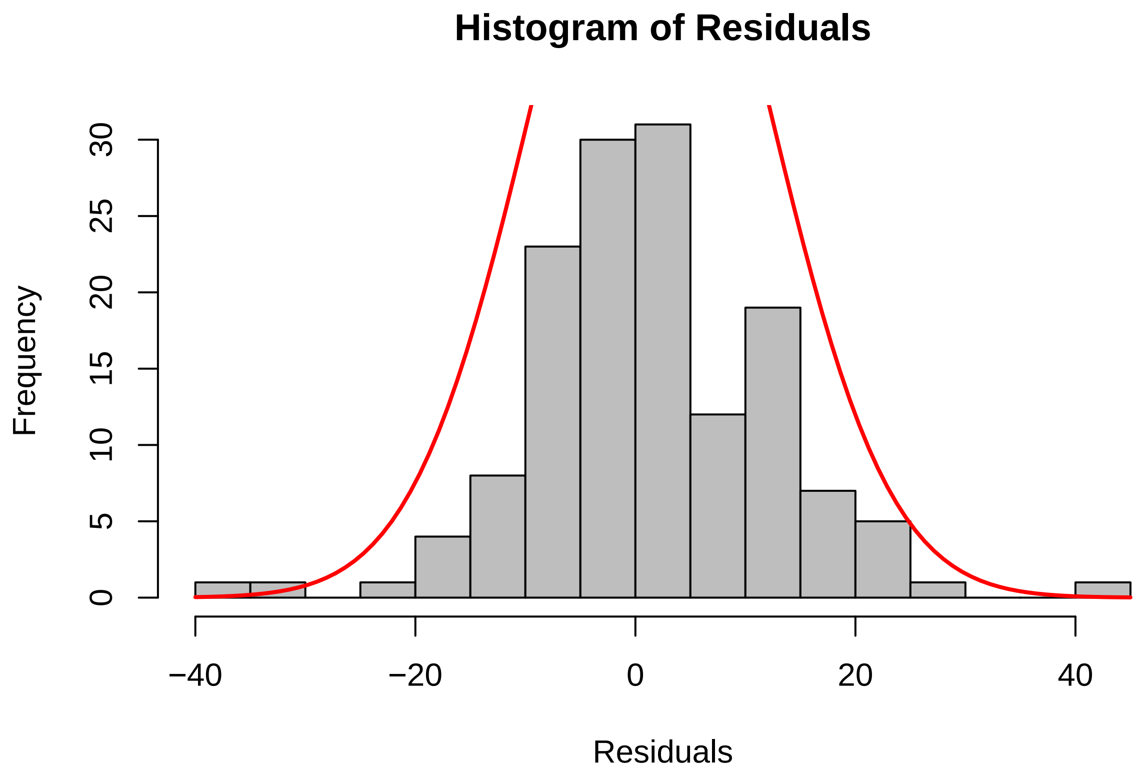
```
##
##  KPSS Test for Level Stationarity
##
## data:  ap_diff
## KPSS Level = 0.028205, Truncation lag parameter = 4, p-value = 0.1
```

```
# (l) Based on the dominating component, select suitable technique
# Presence of trend and seasonality indicates SARIMA model is appropriate

# (m) Fit the data using selected model and estimate parameters
library(forecast)
ap_model <- auto.arima(ap_ts)
summary(ap_model)
```

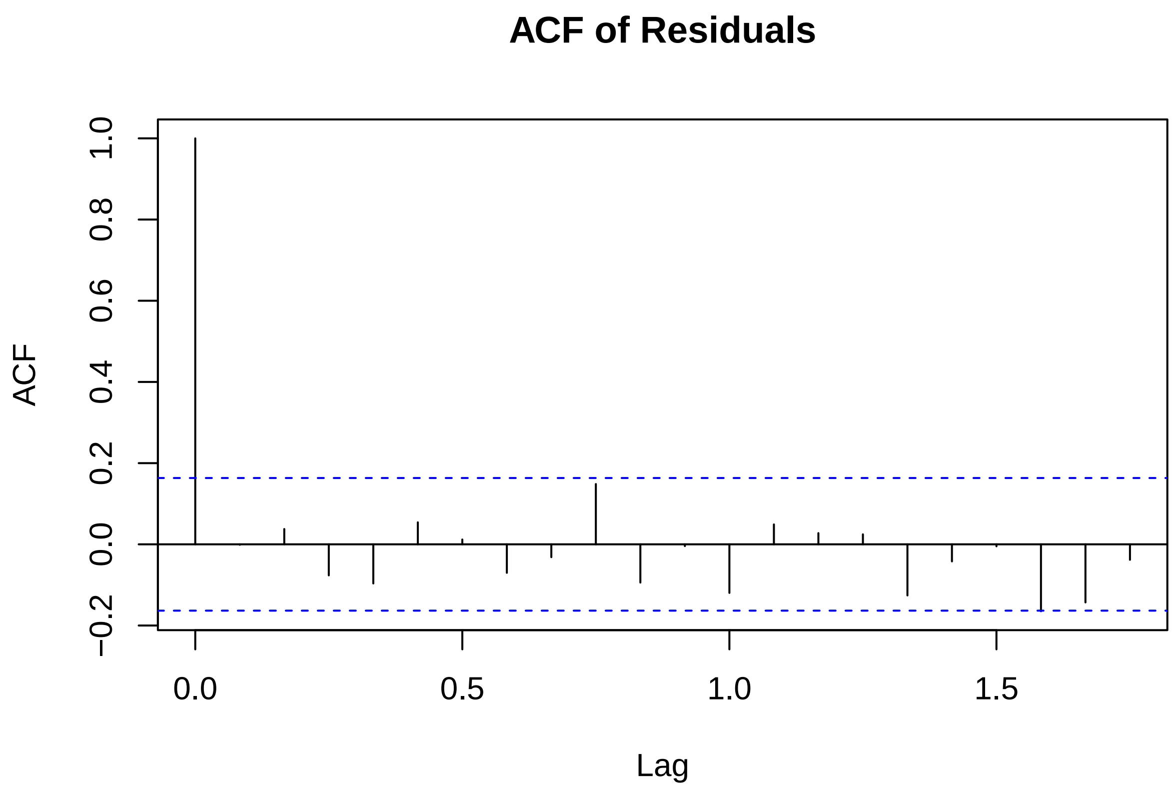
```
## Series: ap_ts
## ARIMA(2,1,1)(0,1,0)[12]
##
## Coefficients:
##          ar1     ar2      ma1
##       0.5960  0.2143  -0.9819
```

```
## s.e.  0.0888  0.0880   0.0292
##
## sigma^2 = 132.3:  log likelihood = -504.92
## AIC=1017.85   AICc=1018.17   BIC=1029.35
##
## Training set error measures:
##                    ME     RMSE     MAE      MPE     MAPE     MASE        ACF1
## Training set 1.3423 10.84619 7.86754 0.420698 2.800458 0.245628 -0.00124847
```

```r
# (n) Check the goodness of fit of the model

# Extract residuals
res <- residuals(ap_model)

# Histogram of residuals with Frequency on Y-axis
hist(res,
     breaks = 20,
     freq = TRUE,
     col = "grey",
     main = "Histogram of Residuals",
     xlab = "Residuals",
     ylab = "Frequency")

# Overlay normal curve
curve(dnorm(x, mean(res), sd(res)) * length(res) *
        diff(hist(res, plot = FALSE)$breaks)[1],
      add = TRUE, col = "red", lwd = 2)
```




```r
# Residual diagnostics
acf(res, main = "ACF of Residuals")
```




```r
# Ljung-Box test
Box.test(res, lag = 12, type = "Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  res
## X-squared = 10.973, df = 12, p-value = 0.5312
```

```r
############################################################
# 4. Output
# The time series plot shows a clear upward trend and strong
# seasonal pattern in the AirPassengers data.
# Decomposition confirms the presence of trend and seasonality.
# ACF and PACF plots indicate non-stationarity.
# KPSS test confirms the series is non-stationary.
# Log transformation and first differencing make the series stationary.
# A SARIMA model is selected using AIC.
# Residuals show no significant autocorrelation.
# Histogram of residuals is approximately normal.
# Ljung-Box test confirms residuals behave like white noise.
#
# 5. Conclusion
# The AirPassengers data exhibits non-stationary behavior due
# to trend and seasonality. Stationarity is achieved using
```

```
# log transformation and differencing. A SARIMA model is
# successfully fitted and validated through diagnostic checks.
# The model provides a good fit and is suitable for analyzing
# and forecasting the given time series.
##########################################################
```