STAT 443: Lab 5

Aronn Grant Laurel (21232475)

31 January, 2025

Question 1

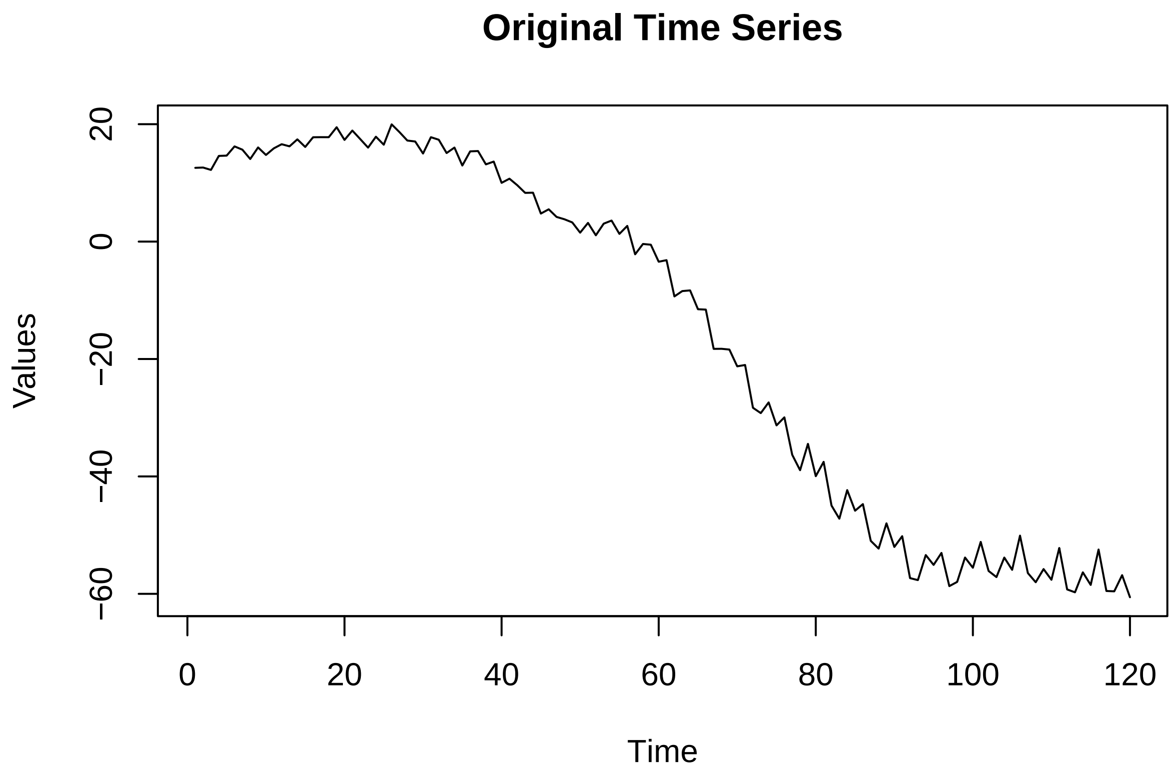
Without using any mathematical notation, describe in words what it means for a time series to be stationary.

For a time series, it's stochastic process is called stationary if its mean is constant. Another way we can determine its stationarity is by observing its plotted data and see a consistent behaviour overtime for its variance.

Question 2

```r
# this is where your R code goes
data <- read.csv("lab4data.csv", header = TRUE)
ts_data <- ts(data[,2])
```

```r
# this is where your R code goes
plot(ts_data,
     main="Original Time Series",
     xlab="Time",
     ylab="Values")
```



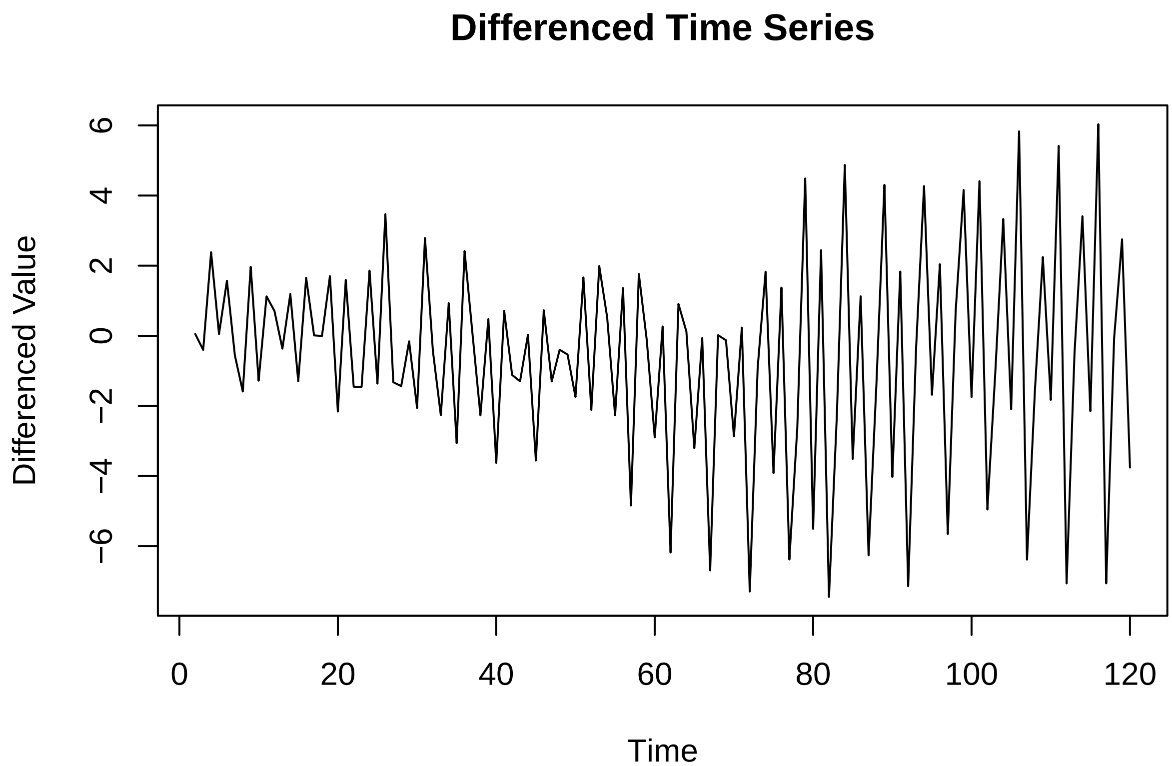

```r
# acf(ts_data, main="ACF of Original Time Series")
```

Based on our time series plot, we can see a strong negative slope / trend,hence the mean is changing over time and we may consider it as non-stationary.

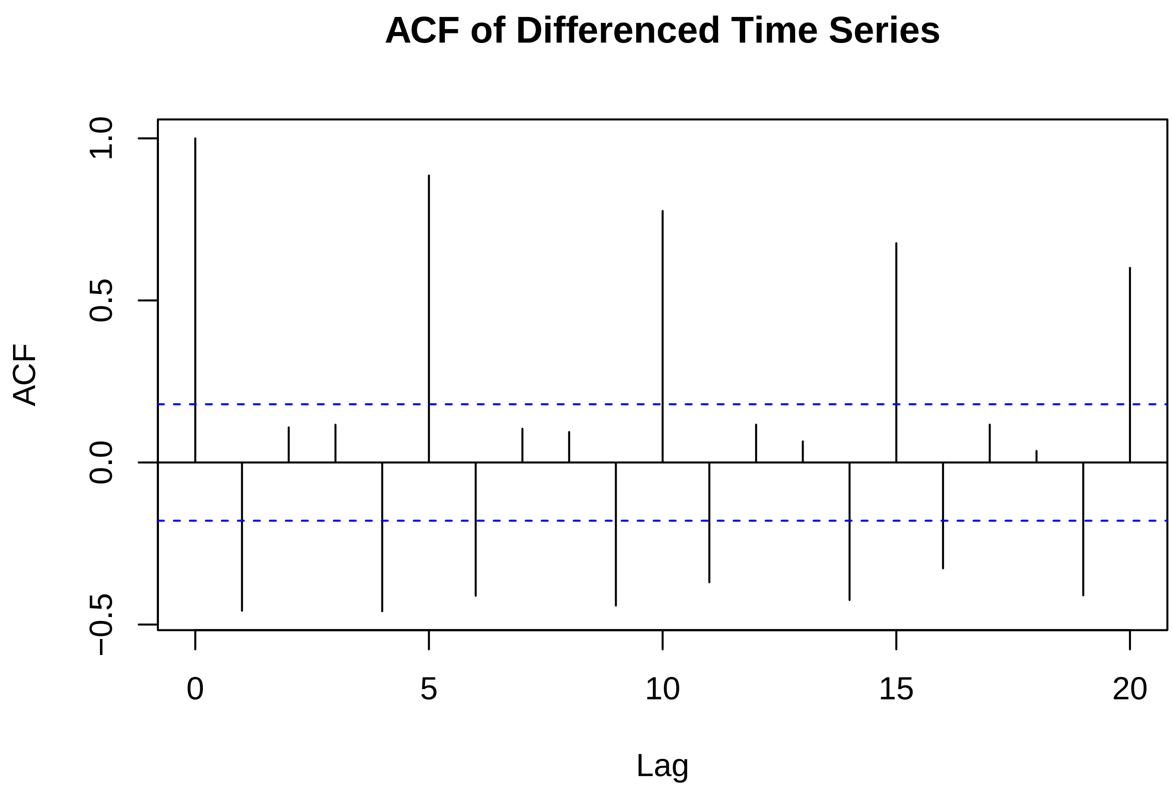
Question 3

```r
# this is where your R code goes
diff_ts <- diff(ts_data, lag=1, differences=1)

plot(diff_ts,
     main="Differenced Time Series",
     xlab="Time",
     ylab="Differenced Value")
```




```
acf(diff_ts, main="ACF of Differenced Time Series")
```



In our difference time plot, we observe values fluctuating around zero at time 0 - 60 which could indicate it being stationary. However, we can also see an increased fluctuation beyond 60. Although we removed the strong downward trend, there might still be some non-stationarity in terms of variance.

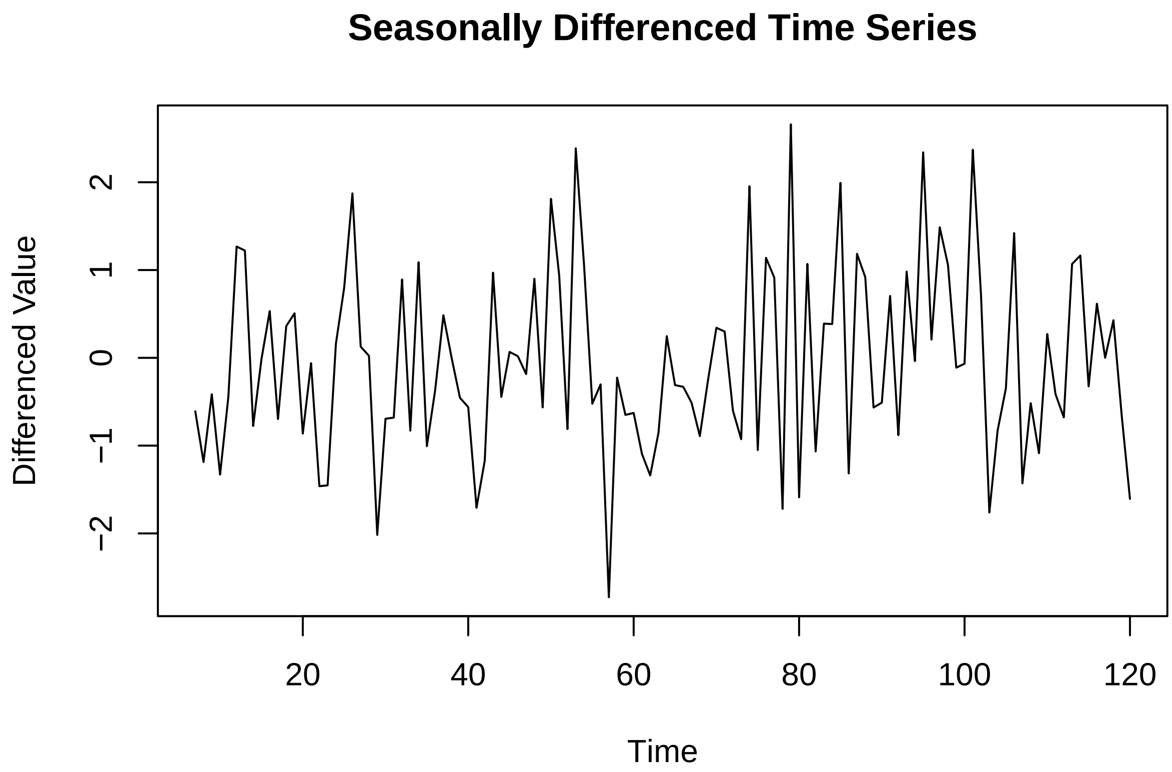
For our ACF plot, we see an osciliating pattern with significant positive spikes at several lags (lags 5, 10, 15, and 20) and significant negative spikes that are +/- one lag from the significant positive spikes. Hence, suggesting a seasonal pattern in the differenced data. Additionally, any strong autocorrelation from the non-difference data has been removed as most lags fall outside the boundary line.

Question 4

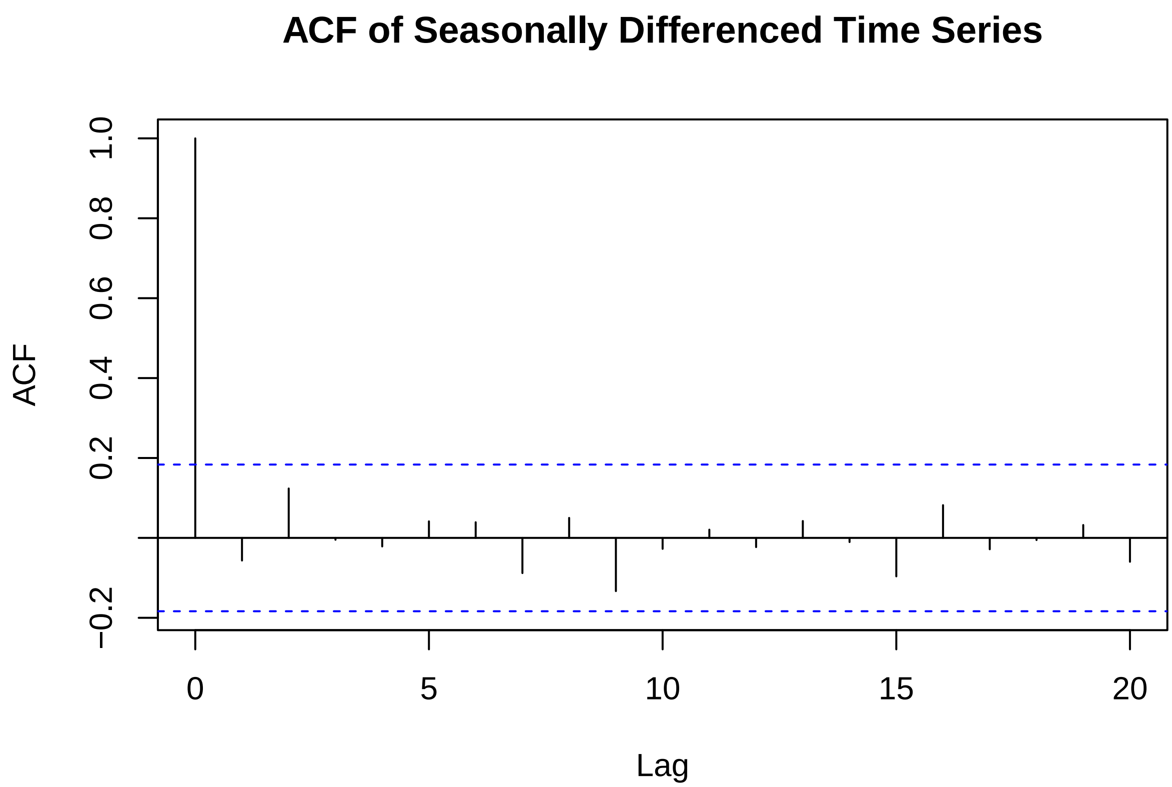
```r
# this is where your R code goes
acf_values <- acf(diff_ts, plot=FALSE)
seasonal_period <- which.max(acf_values$acf[2 : length(acf_values$acf)])

seasonal_diff <- diff(diff_ts, lag=seasonal_period, differences=1)

plot(seasonal_diff,
     main = "Seasonally Differenced Time Series",
     ylab = "Differenced Value",
     xlab = "Time")
```




```
acf(seasonal_diff, main = "ACF of Seasonally Differenced Time Series")
```



From the acf plot, we can see that it resembles a white noise process.

Question 5

From the lecture slides of Chapter 2, I would suggest a SARIMA(p,d,q) * (P,d,Q)s model where: d = difference = 1.

Question 6

(a)

Y_t = X_t - Xt_1 -(1) W_t = Y_t - Y_{t-s} - (2)

Inserting (1) in (2): W_t = Y_t - Y_{t-s} = X_t - X_{t-1} - (X_{t-s} - X_{t-s-1}) = X_t - X_{t-1} - X_{t-s} + X_{t-s-1}

(b) Y_t = X_t - X_{t-1} = (1-B) X_t

(c)

Given from Q b) Y_t = (1-B)X_t$ and Y_{t-s} = (1-B) X_{t-s} = (1-B) B^s X_t

W_t = Y_t - Y_{t-s} W_t = (1 - B)X_{t} - [(1-B) B^s X_t] W_t = (1-B)(X_{t} - B^s X_t) W_t = (1-B)(1-B^s)X_t