STAD57: Time Series Analysis
Problem Set 2 Solutions

1. Use R to generate 400 observations from the following models:

  a. $X_t = W_t + 2W_{t-1} - W_{t-2}$ (MA)

  b. $X_t = -.8X_{t-1} + .4X_{t-2} + W_t$ (AR)

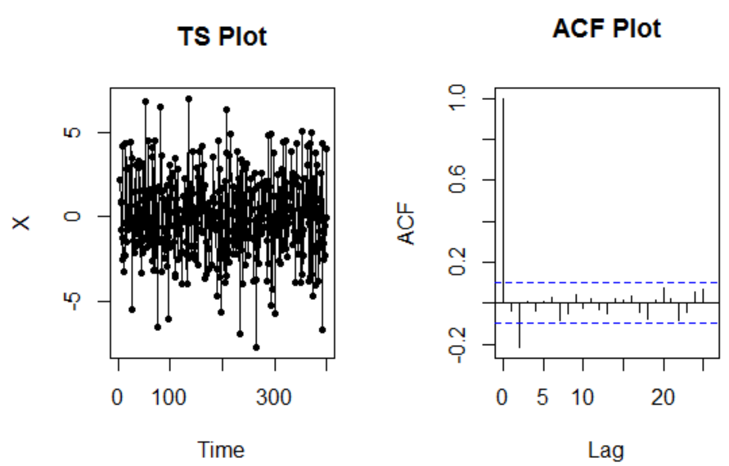
  c. $X_t = .01 + X_{t-1} + W_t$ (Random walk)

(Hint: first generate a Normal white noise sequence $\{W_t\}$ with function `rnorm` and then use function `filter`). Create time series plots and ACF plots for all the series.

SOL:

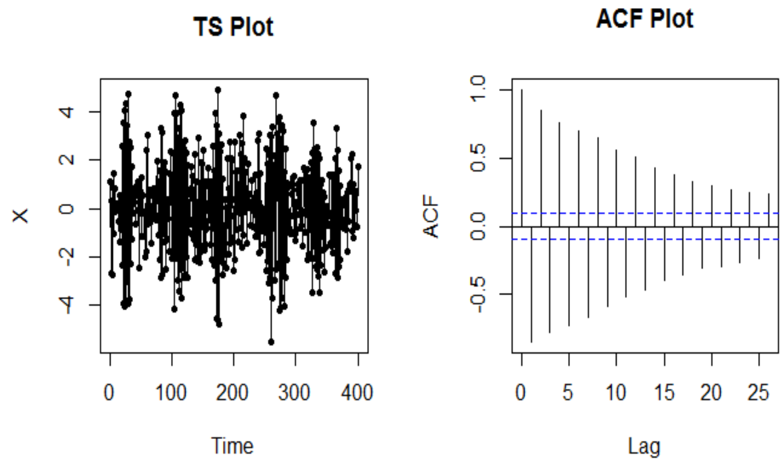
a.

```
W=rnorm(n=400, mean=0, sd=1)
X=filter(W, c(1,2,-1), sides=1)
par(mfrow=c(1,2))
plot(X,type='o',pch=20, main="TS Plot")
acf(X, na.action = na.pass, main="ACF Plot")
```



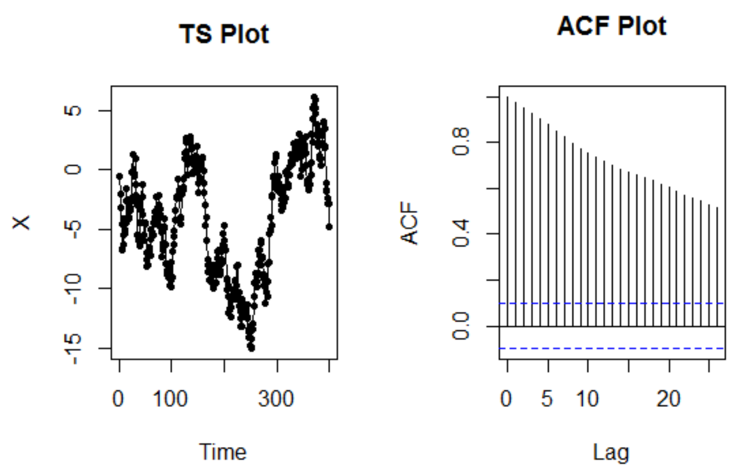
b.

```
X=filter(W, c(-.5,.4), method = "recursive")
plot(X,type='o',pch=20, main="TS Plot")
acf(X, na.action = na.pass, main="ACF Plot")
```



c.

```
X=filter(W+.01, c(1), method = "recursive")
plot(X,type='o',pch=20, main="TS Plot")
acf(X, na.action = na.pass, main="ACF Plot")
```



2. Load the following TS data sets in R and pre-process them (using de-trending, differencing, transformations etc) so that they become as close to stationary as possible.
   a. Monthly Canadian reserves (in $)
   b. Monthly car sales in Quebec (in # cars)
   c. Daily average temperatures in Toronto (in °C)

   Create time series plots and ACF plots for all of the original and processed series.

SOL:

There are typically more than one ways to model a time series, so the following answers are not strictly right (or wrong). R code is given at the end.

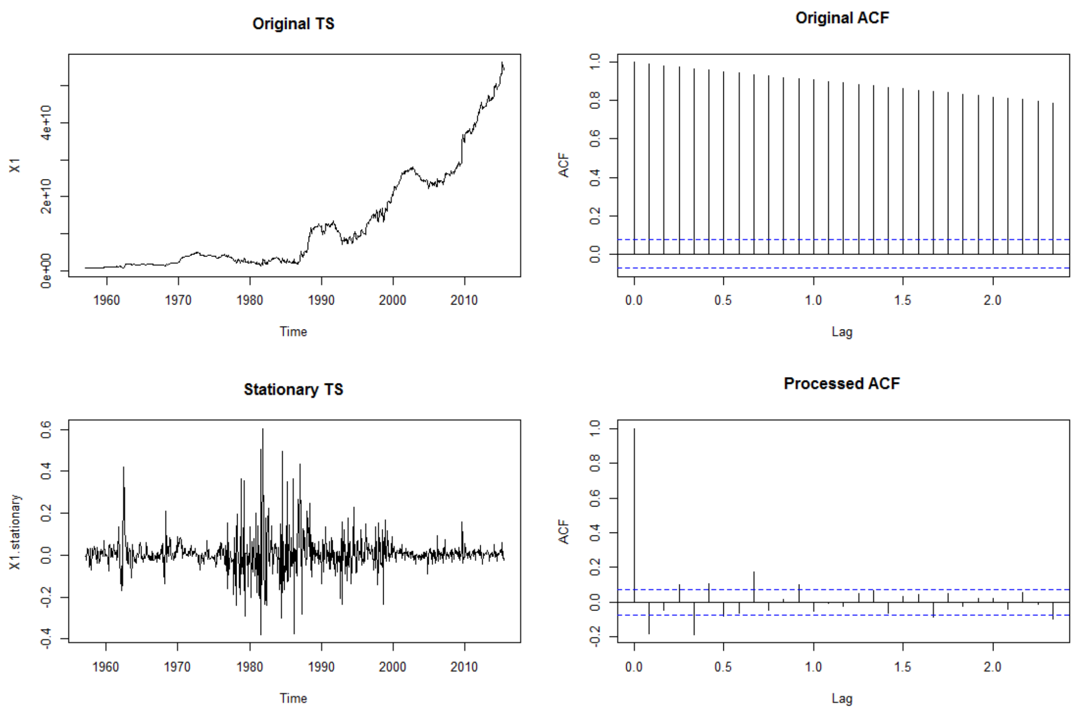
a.

I used first order differences of the logarithm of the series, i.e.:

$$Y_t = \nabla \log(X_t) = \log(X_t) - \log(X_{t-1}) = \log\left(X_t / X_{t-1}\right)$$

These are sometimes called log-returns, or continuously compounded returns (https://en.wikipedia.org/wiki/Rate_of_return#Logarithmic_or_continuously_compounded_return) and they are approximately equal to usual returns b/c

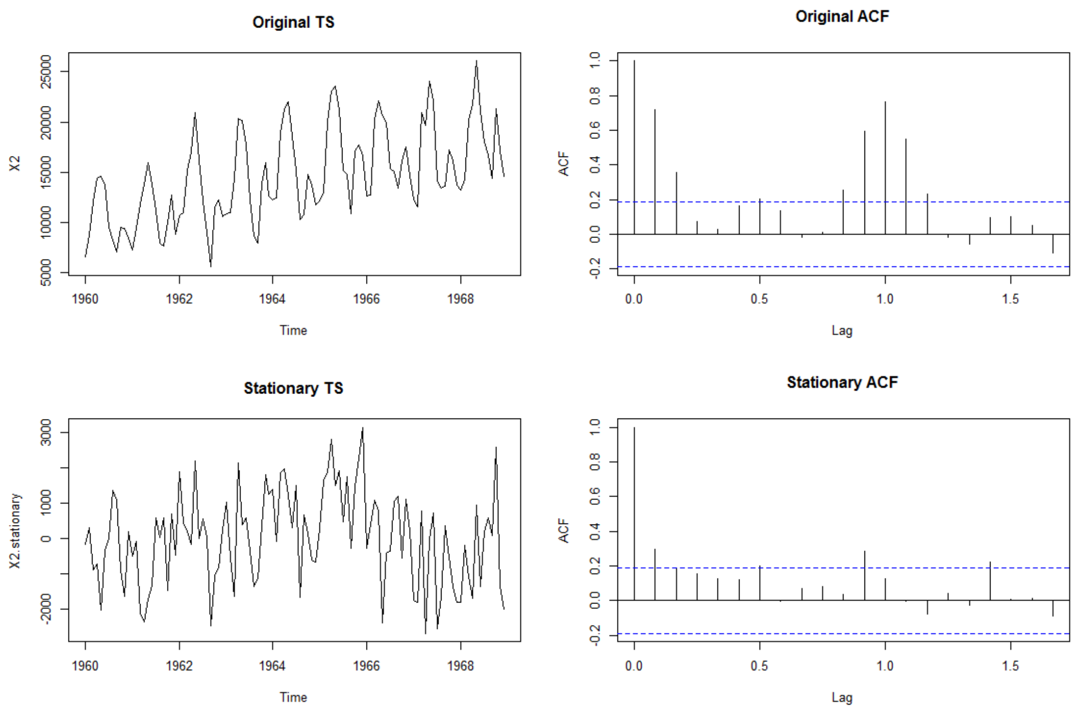
$\log(X_t) - \log(X_{t-1}) \approx \log'(X_{t-1}) \nabla X_t = \dfrac{\nabla X_t}{X_{t-1}} = \dfrac{X_t - X_{t-1}}{X_{t-1}}$ (by using the 1st order Taylor expansion of the log function $\log(x+h) - \log(x) \underset{\text{for } h \to 0}{\approx} \log'(x) \times h$). The log-return (i.e. diff-log) transformation is very common for financial data (and has interesting modeling implications).



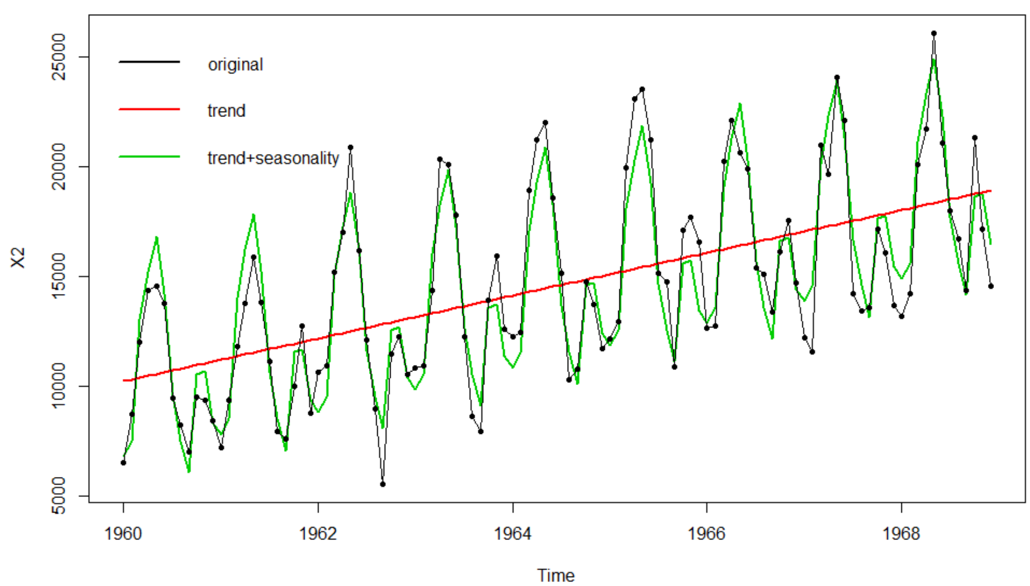
b.

I used detrending w/ linear trend and additive seasonality: $Y_t = X_t - T_t - S_t$,

where $T_t = \hat{\beta}_0 + \hat{\beta}_1 t$ from the regression model $X_t = \beta_0 + \beta_1 t + \varepsilon_t$, and $S_t = \hat{\mu}_{(t \bmod 12)}$ from the ANOVA model $X_t = \mu_{(t \bmod 12)} + \eta_t$

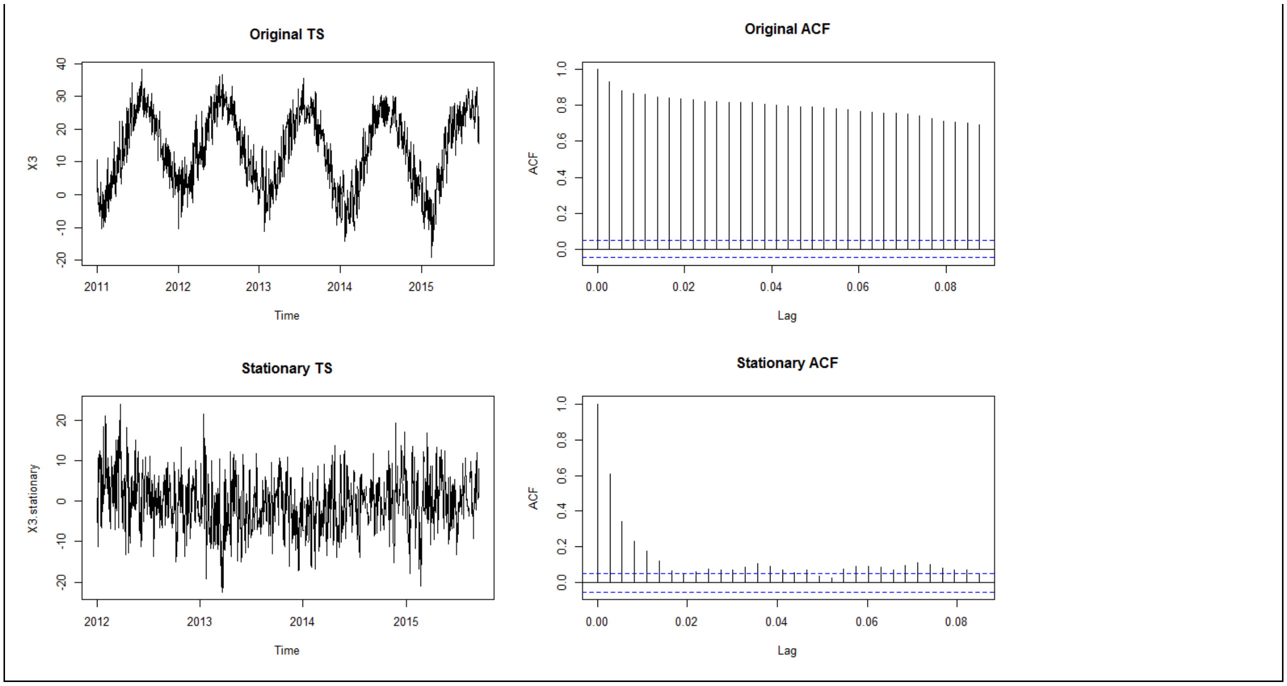


The plot of the fitted linear trend & seasonality is give below:



c.

Since these are daily temperature data, they should have an annual seasonal pattern. So, I just used differencing at lag 365 (i.e. 1 year): $Y_t = X_t - X_{t-365}$



Original TS
X3
Time
Original ACF
ACF
Lag
Stationary TS
X3.stationary
Time
Stationary ACF
ACF
Lag