PREDICTIVE ANALYTICS

HW2

Chris Cirelli

02/15/2020

Question 1

Chris Cirelli

Time Series Analysis HW 2

1.) Let $w_t : t = 0, 1 \ldots$ be a white noise process with variance σ^2_w and let $|\phi| <$ be a constant.

Consider the process $x_0 = w_0$ and

$$x_t = \phi x_{t-1} + w_t, \quad t = 1, 2 \ldots$$

a.) Show that $x_t = \sum_{j=0}^{t} \phi^j w_{t-j}$ for any $t = 1$.

$$x_t = \phi(\phi x_{t-2} + w_{t-1}) + w_t$$

$$x_t = \phi^2 x_{t-2} \phi w_{t-1} + w_t$$

$$x_t = \phi^k x_{t-k} + \sum_{j=0}^{k-1} \phi^j w_{t-j}$$

assuming $|\phi| < 1$ as $j \to \infty$ we can represent the AR(1) process as a linear one given by

$$\sum_{j=0}^{\infty} \phi^j w_{t-j}$$

b.) Find $E(x_t)$

$$E(x_t) = \sum_{j=0}^{\infty} \phi^j E(w_{t-j}) = 0$$

as $j \to \infty$ the value of $x_t \to w_t$ and $E(w_t) = 0$.

c) Show that for $t = 0, 1 \ldots$

$$var(x_t) = \frac{\sigma^2_w}{1-\phi^2}\left(1 - \phi^{2(t+1)}\right)$$

$$x_t = \phi x_{t-1} + w_t$$

$$x_t = \Sigma_{j=0}^{\infty} \phi^j w_t - 1$$

$$var(x_t) = \Sigma_{j=0}^{\infty} \phi^j \cdot \underbrace{var(w_t - 1)}_{= \sigma^2}$$

assumes $a \perp b$ so we can just add

assumes w_t $N(0, \sigma^2)$

$$var(x_t) = \Sigma_{j=0}^{\infty} \phi^j \cdot \sigma^2$$

product across an infinite series.

$$var(x_t) = \sigma^2 (1 + \phi^1 + \phi^2 + \phi^3 \ldots)$$

$$= \frac{\sigma^2_w}{1-\phi^2}$$

d.) Show that for $h \geq 0$ $\quad cov(x_t, x_{t+h}) = \phi^h var(x_t)$

Question? why do we use expectation w/ covariance? because we est μ?

$$E\left[E(x_t - \hat{x}) \cdot E(x_{t-1} - \hat{x}_{t-1})\right]$$

TBD

HW 2. Question 1

e.) Is x_t Stationary?

x_t stationary $\mid$ $|\phi| < 1$.

$AR(1) = \quad x_t = \phi x_{t-1} + w_t$

if $\phi = 1 \Rightarrow \quad x_t = 1 \cdot x_{t-1} + E(w_t)$

$x_t = x_{t-1} + 0$

$\therefore$ x_t is only dependent on the prior observation & w/ assume a random walk.

Furthermore, $Var(x_t) = t \times \sigma^2_w$ meaning that variance will increase unchecked as $t \Rightarrow \infty$

if $\phi > 1$ then the past observations of x_t grow as t increases, meaning x is progressively more dep on older observations. Logically, this does not make sense, and as a geometric series does not allow for a finite sum.

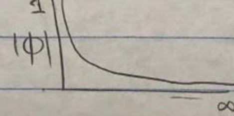
x_t is stationary $\mid$ $|\phi| < 1$

Allows for an infinite geometric series.

As $t \Rightarrow \infty$ $|\phi| \Rightarrow 0$.

f.) Argue that as $t \Rightarrow \infty$ the process becomes stationary.

- $|\phi| < 1$ leads to a decaying function
- As $t \Rightarrow \infty$ $|\phi|$ will approach but



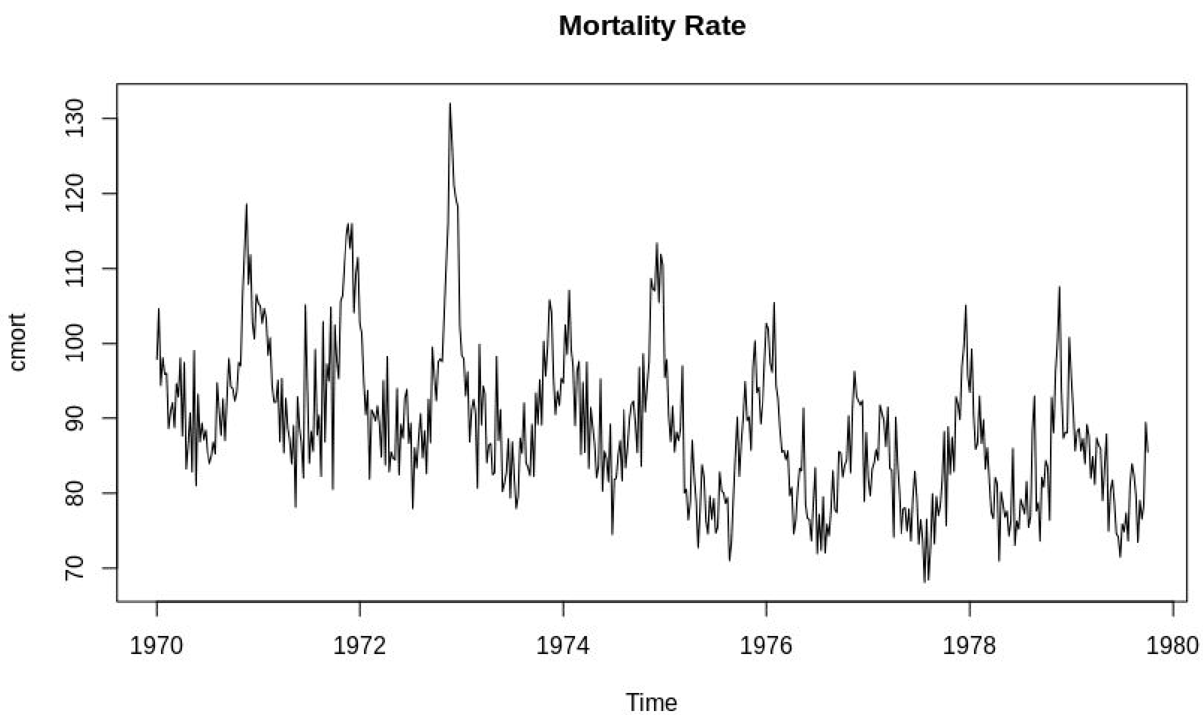
never reach zero. Therefore, past observations of x_t w/ decrease in importance.

and given $x_t = \sum_{j=0}^{\infty} \phi^j w_{t-j}$ $\quad x_t \approx w$ which we know is stationary when $w = N(0, \sigma^2)$ & iid.

Question 2

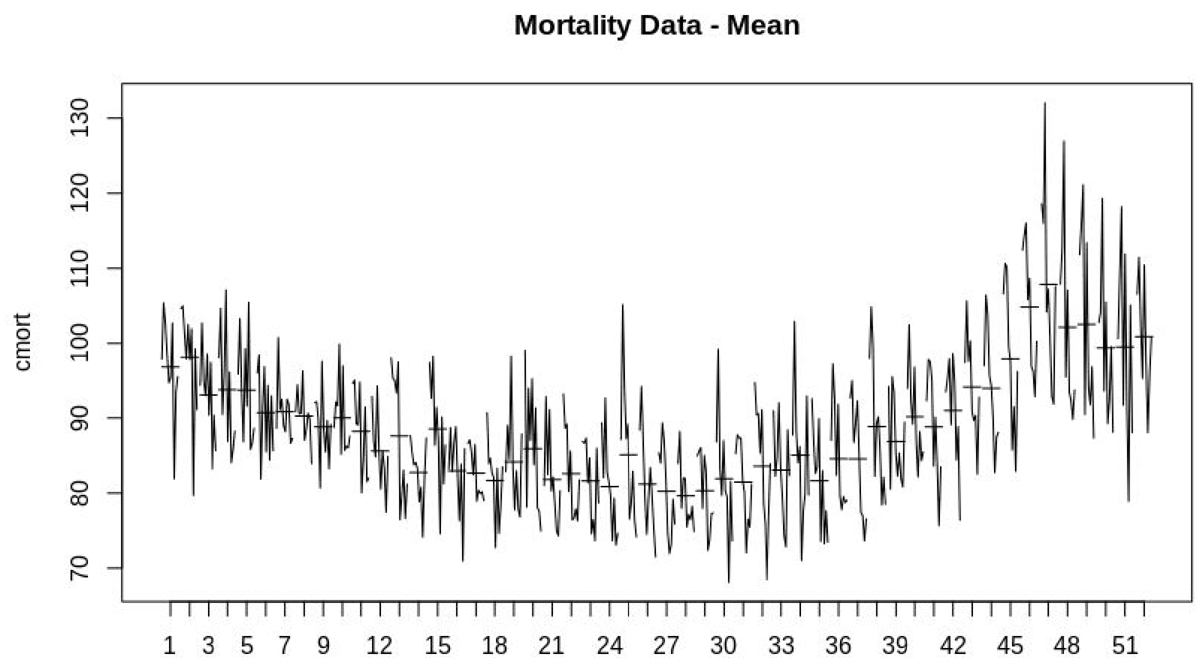
a.) Plot the data and ACF, PACF. Comment on the stationarity and discuss which model you would use.

plot(cmort)

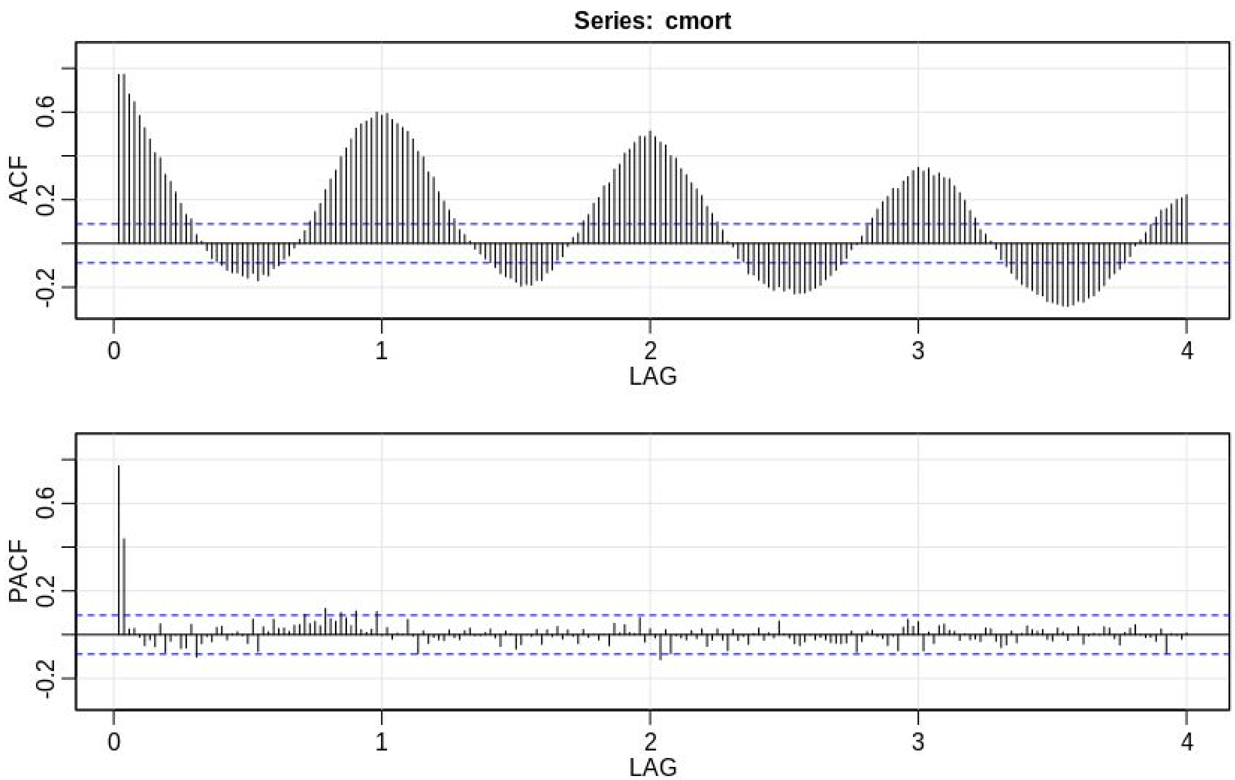


The data does not appear to be stationary. The mean appears to decline over time "trend" and there appears to be yearly seasonality.

```
cmort1  = cmort[0: 100]
cmort2  = cmort[408 : 508]
cmort1_mu = mean(cmort1)
cmort2_mu = mean(cmort2)
mu1_2 = c(cmort1_mu, cmort2_mu)
barplot(mu1_2, main='Barplot Mu1 and Mu2')
monthplot(cmort, main='Mortality Data - Mean')
```



'Plot ACF & PACF'
par(mfrow=c(2,1))
acf(cmort)
pacf(cmort)



Observations:

- The ACF and PACF plots indicate that this is an AR model with some seasonality at increments of 1. In the case of the ACF, the phi value progressively decreases from 1 as the lag increases. If this were an MA model we would see a single bar and then nothing. In the case of the PACF plot, we see two bars and then minimal ones thereafter. If this were an MA model we would see progressively declining bars as is the case with the ACF plot.

Question 2(b)
Fit an AR(2) model to xt using the method of moments Yule Walker equation

```
cmort.diff <- diff(log(cmort))
ar.yw <- ar.yw(cmort.diff, order=2)
ar.yw

Call:
ar.yw.default(x = cmort.diff, order.max = 2)
Coefficients:
      1        2
-0.5678  -0.0945

Order selected 2  sigma^2 estimated as  0.004183
```

Matrix

```
            [,1]        [,2]
[1,] 0.001966395 0.001020007
[2,] 0.001020007 0.001966395
```

```
out = acf(cmort.diff)
rho = out$acf[1:5]
G2 = matrix(c(rho[1],rho[2],rho[2],rho[1]),2)
g2 = c(rho[2],rho[3])
solve(G2)%*%g2
[1,] -0.56775711
[2,] -0.09453654
```

Question 3

a.) $x_t = 0.8x_{t-1} - 0.15x_{t-2} + w_t - 0.3w_{t-1}$

$x_t - 0.8x_{t-1} + 0.15x_{t-2} = w_t - 0.3w_{t-1}$

$(1-0.8B)x_t + (1-0.15B^2)x_t = (1-0.3B)w_t.$

$(1-0.8B+0.15B^2)x_t = (1-0.3B)w_t.$

ARMA (2,1)

$\phi(B) = (1-0.8B+0.15B^2) = (1-0.5B)(1-0.3B)$

$1-0.3B-0.5B+0.15B^2$

$1-0.8B+0.15B^2$

- common factor = $(1-3B)$
- Therefore ARMA(1,1) model.

b.) $x_t = x_{t-1} - 0.5x_{t-2} + w_t - w_{t-1}$

$x_t - x_{t-1} + 0.5x_{t-2} = w_t - w_{t-1}$

$(1-B+0.5B^2)x_t = (1-B)w_t.$

- Common factor = $(1-B)$
- Therefore ARM(1,1) model.

Causality: $\phi(z) \neq 0$ for $|z| \leq 1$

ϕ & $\theta \neq 0$

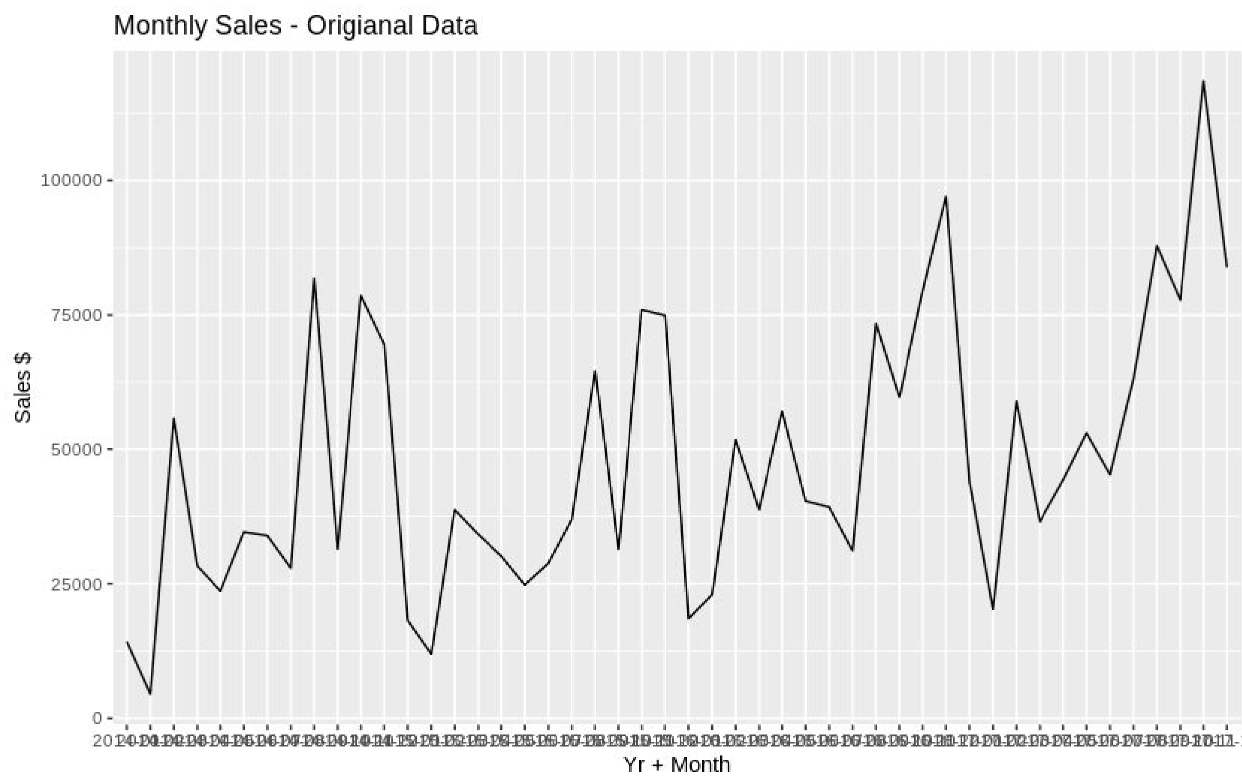
Both are causal & invertible

Question 4

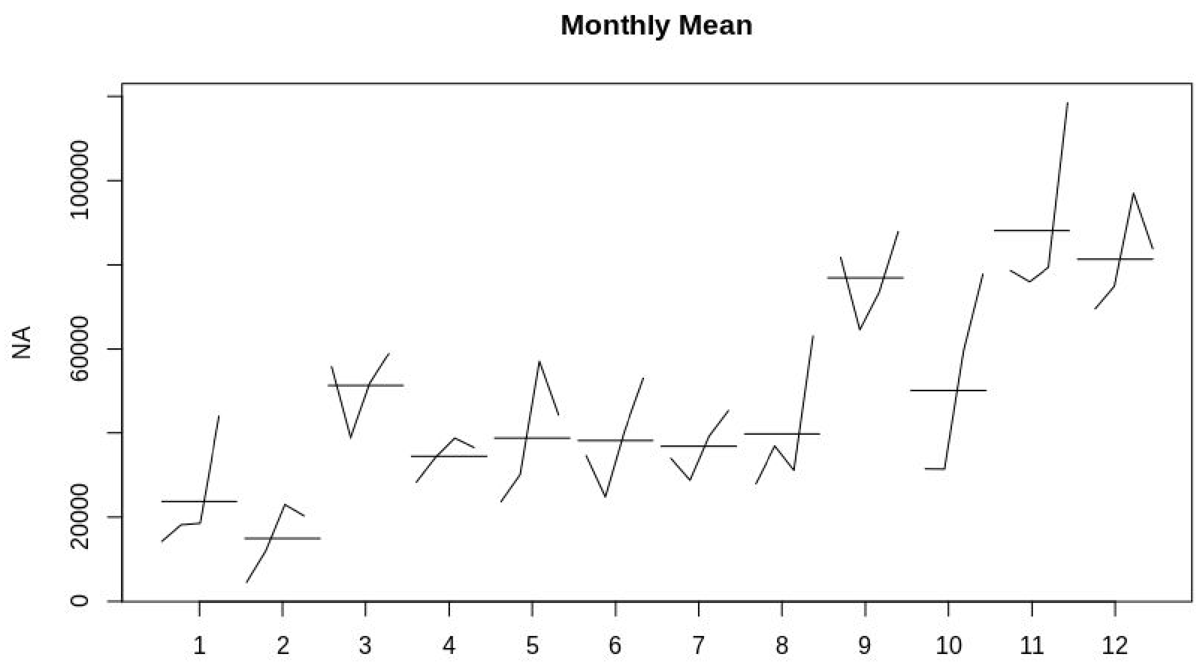
```
# Clear Namespace
rm(list=ls())
dev.off(dev.list())

# Load Libraries
library(readxl)
library(plyr)
library(ggplot2)
library(astsa)
library(tseries)

## DATA PREPARATION -----------------------------------------------------

# Load Data & Get Monthly Sales
sales <- read_excel("Desktop/repositories/Time_Series/hw2/sales.xls")
sales$yr <- format(sales$`Order Date`, format='%Y')
sales$yr_month <- format(sales$`Order Date`, format='%Y-%m')
sales[c('Sales', 'yr_month')]
df.sum <- ddply(sales, c('yr_month'), summarize, Sales = sum(Sales))
```

```
> df.sum
   yr_month      Sales
1   2014-01  14236.895
2   2014-02   4519.892
3   2014-03  55691.009
4   2014-04  28295.345
5   2014-05  23648.287
6   2014-06  34595.128
7   2014-07  33946.393
8   2014-08  27909.468
9   2014-09  81777.351
10  2014-10  31453.393
11  2014-11  78628.717
12  2014-12  69545.621
13  2015-01  18174.076
```

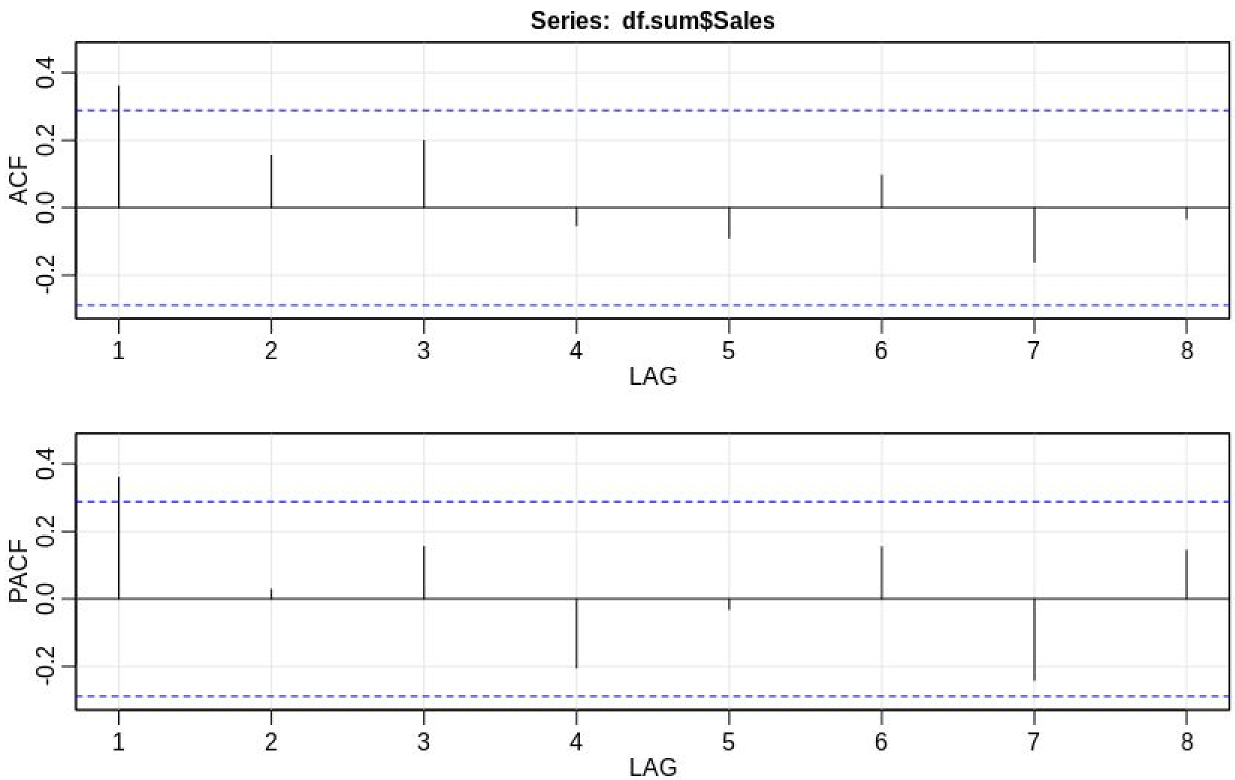
```
# Plot Monthly Sales
' Mean and variance do not appear to be constant'
ggplot(df.sum, aes(x=df.sum$yr_month, y=df.sum$Sales, group=1)) + ggtitle('Monthly Sales - Origianal Data') +
  xlab('Yr + Month') + ylab('Sales $') + geom_line()
```



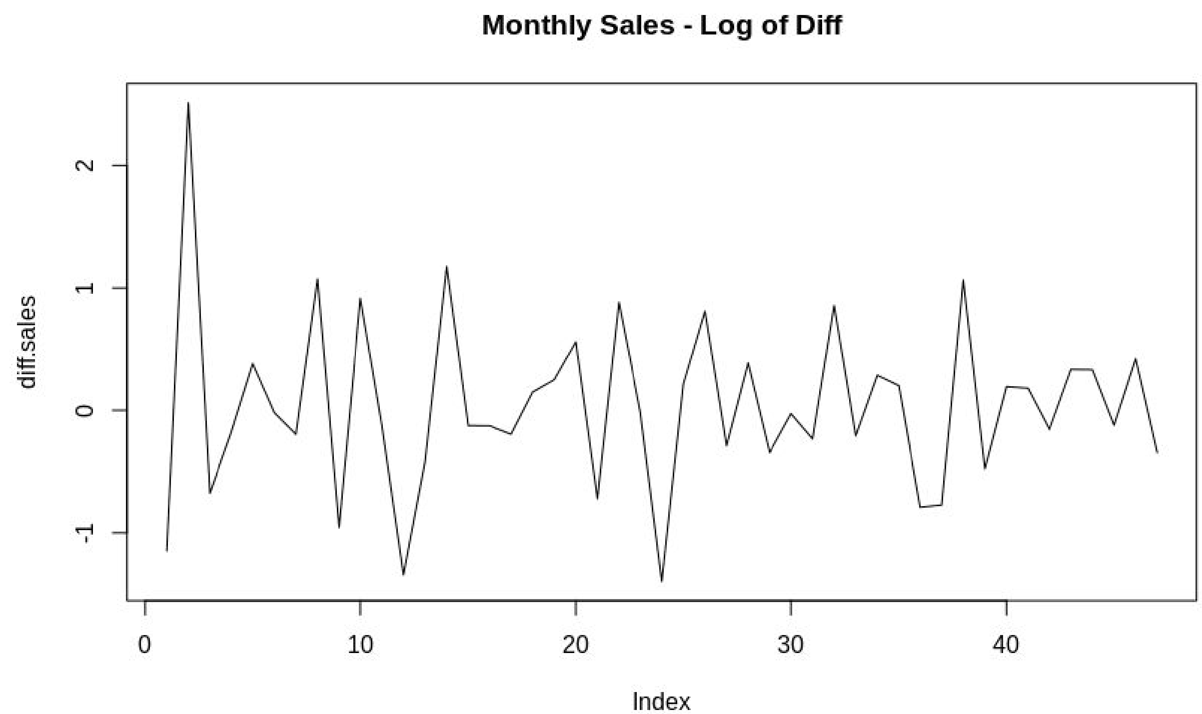

```
# Check If Mean Constant
' Mean is not constant and therefore, the timeseries is not stationary'
monthplot(df.sum$Sales, main='Monthly Mean')
```



ACF & PACF - Original Data
' ACF Looks like a slow decaying function up to lag=4 & the PACF drops from 0.36 to 0.03 or close to zero.   This looks like an AR(1) process'
acf2(df.sum$Sales)



Take diff and log of data
' Seems to e centered about the mean now. Some change in variance'
diff.sales <- diff(log(df.sum$Sales))
plot(diff.sales, type='l', main='Monthly Sales - Log of Diff')




```
# Take ACF & PACF - Differenced Data
' Interpretation:  This looks like an AR model based on the ACF and order of 2 based on the PACF'
acf2(diff.sales, max.lag=12)


## FIT AR MODELS ---------------------------------------------------

# YW Fit
' Note: |phi| < 1'
sales.yw <- ar.yw(diff.sales, order=2)
sales.yw$x.mean
sales.yw.mse <- sum(sales.yw$resid[3: length(sales.yw$resid)]^2) / length(sales.yw$resid[3 : length(sales.yw$resid)])
plot(sales.yw$resid, type='l')
sales.yw$ar # coefficients
sales.yw$asy.var.coef # covariance matrix

# MLE Fit
sales.mle <- ar.mle(diff.sales, order=2)
sales.mle$x.mean
sales.mle.mse <- sum((sales.mle$resid[3: length(sales.mle$resid)])^2) / length(sales.mle$resid)
```

```
# YW Prediction
sales.pr <- predict(sales.yw, n.ahead=10)
comb <- c(diff.sales, sales.pr$pred)
ts.plot(comb, type='l', main = 'Plot YW Prediction')  # Plot w/ prediction
lines(sales.pr$pred + sales.pr$se, lty=2) # conf intv
lines(sales.pr$pred - sales.pr$se, lty=2) # conf intv

# MLE Prediction
sales.pr <- predict(sales.mle, n.ahead=10)
comb <- c(diff.sales, sales.pr$pred)
ts.plot(comb, type='l', main='Plot MLE Prediction')  # Plot w/ prediction
lines(sales.pr$pred + sales.pr$se, lty=2) # conf intv
lines(sales.pr$pred - sales.pr$se, lty=2) # conf intv

# Compare MSE
' MLE has the lower MSE'
paste('MLE MSE => ', sales.mle.mse)
paste('YW MSE  => ', sales.yw.mse)


## FIT ARIMA MODELS -------------------------------------------------------

# Arima Model
sales.mle.ar1 <- arima(diff.sales, c(1,0,0))
sales.mle.ar2 <- arima(diff.sales, c(2,0,0))


# MA Model
sales.mle.ma1 <- arima(diff.sales, c(0,0,1))
sales.mle.ma2 <- arima(diff.sales, c(0,0,2))

# AR(1) AR(2) Predictions
ar1.pr <- predict(sales.mle.ar1, n.ahead=10)
ar1.mse <- sum(ar1.pr$se^2) / length(ar1.pr$se)
ar2.pr <- predict(sales.mle.ar2, n.ahead=10)
ar2.mse <- sum(ar2.pr$se^2) / length(ar2.pr$se)

# MA(1) MA(2) Predictions
ma1.pr <- predict(sales.mle.ma1, n.ahead=10)
ma1.mse <- sum(ma1.pr$se^2) / length(ma1.pr$se)
dma2.pr <- predict(sales.mle.ma2, n.ahead=10)
ma2.mse <- sum(ma2.pr$se^2) / length(ma2.pr$se)
```

```
# Compare Results
mse.results <- c(ar1.mse, ar2.mse, ma1.mse, ma2.mse)
mse.labels <- c('AR1', 'AR2', 'MA1', 'MA2')
barplot(mse.results, names.arg = mse.labels, main= 'MSE RESULTS')
```