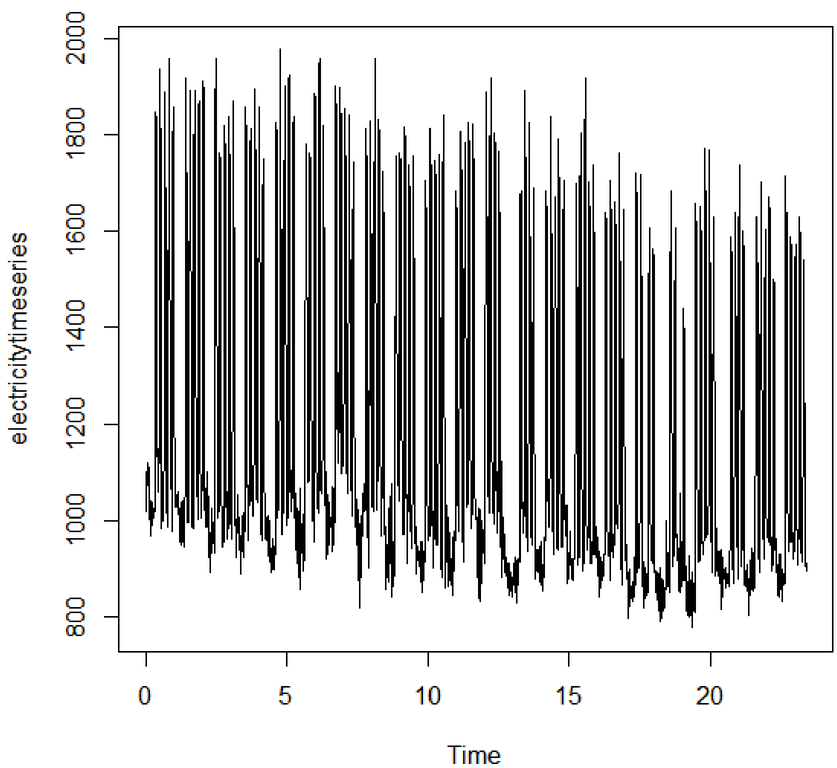
ST3233 Assignment 1
Group 31
A0102680R Ethan Koh
A0127222U Chan Yan Jia
A0124817E Lim Wei Qi
A0131386H Abigail Teo Si Min
A0105533R Wang Jiabao

Exercise 1 (Electricity forecast)

#part 1
> data=read.table("D:/NUS/ST3233 Applied Time Series Analysis/datasets/electricity_load.dat")
> electricitytimeseries=ts(data$V1,frequency=168,start=c(0))

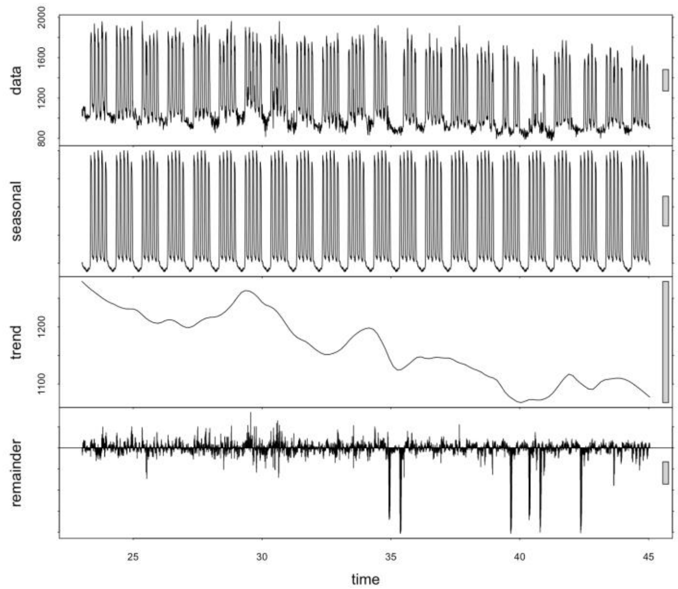
#part 2
> plot(electricitytimeseries)

#appropriate seasonal period can be weekly

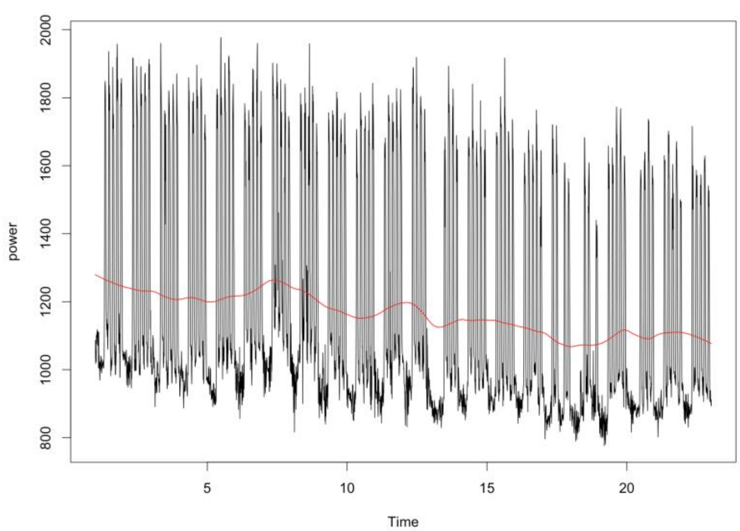


#part 3
> electricity.decomp=stl(electricitytimeseries,s.window="periodic",robust=TRUE)

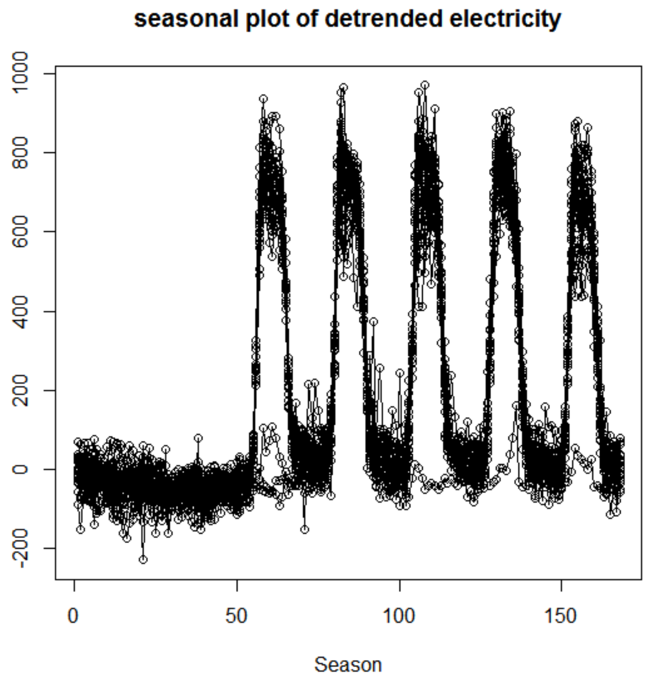
```
> plot(electricity.decomp)
```




```
> trend=electricity.decomp$time.series[,"trend"]
> plot(electricitytimeseries)
> lines(trend,col="red")
```



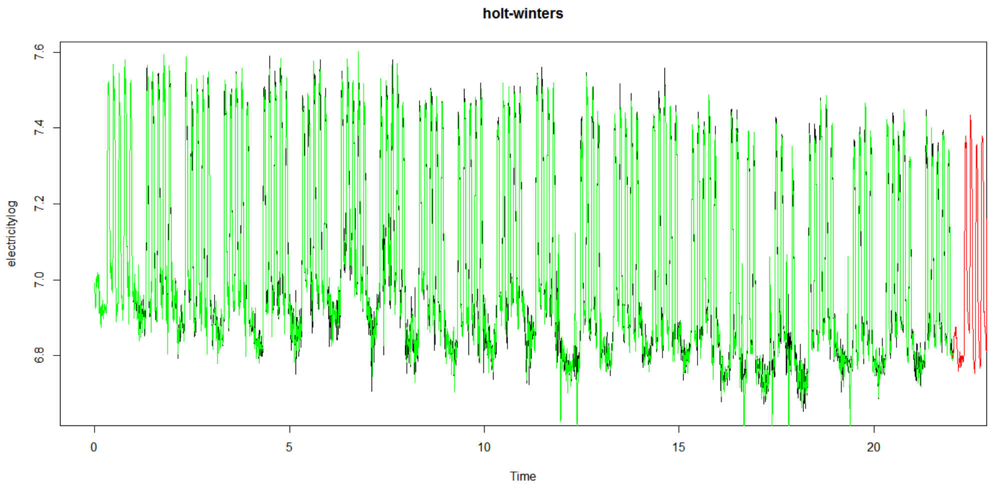

```
#part 4
> library(fpp)
> seasonplot(electricitytimeseries-trend,s=168,main="seasonal plot of detrended electricity")
```



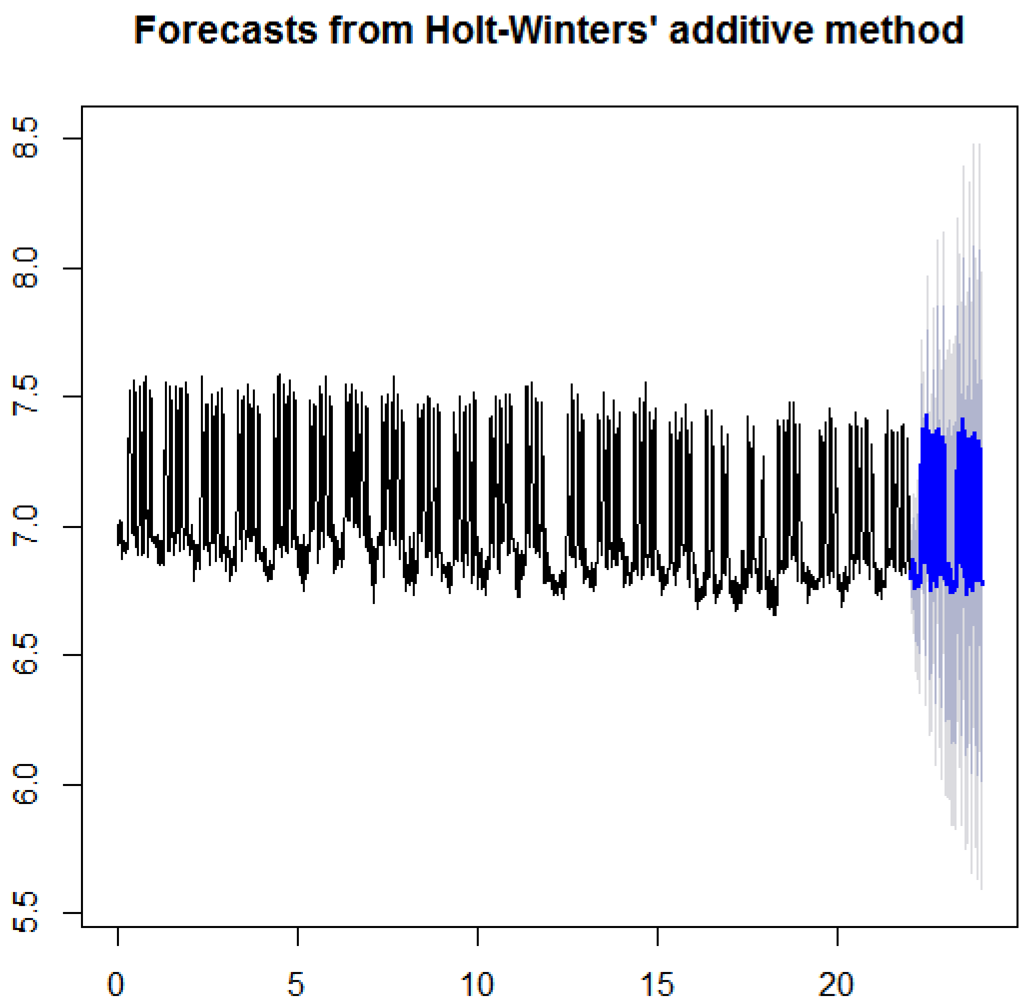
For a time series to be perfectly seasonal, each period needs to have the same pattern. From the graph above, we can see that the graph have different patterns. Certain data follows a concave shape while other data follows a straighter line. The seasonal pattern is not a fixed and known period in this case.

#part 5

```
> electricitylog=log(electricitytimeseries)
> electricitysmoothing=hw(electricitylog,h=336,initial="simple",seasonal="additive")
> plot(electricitylog,main="holt-winters")
> lines(fitted(electricitysmoothing),col="green")
> lines(electricitysmoothing$mean,col="red")
```



> plot(electricitysmoothing)



Exercise 2 (Bootstrap estimate)

(i) $X_k = W_k - (5/6)\, W_{k-1} + (1/6)\, W_{k-2}$

$X_k = (1-(5/6)B + (1/6)B^2)\, W_k$

$ø(x) = 1-(5/6)x + (1/6)x^2$

$ø(x) = 0$

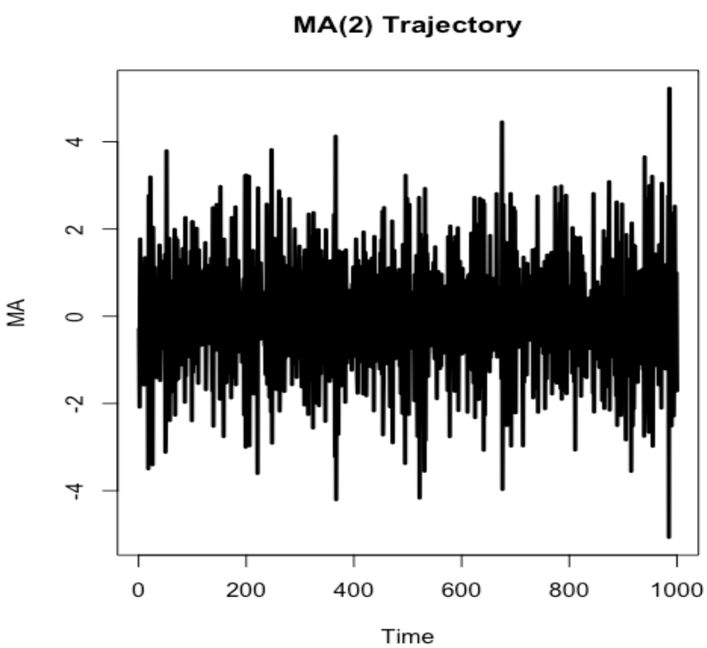
$1-(5/6)x + (1/6)x^2 = 0$

$x = 2$ or 3

Since the roots of the polynomial characteristic are both larger than the absolute value of 1, the MA(2) process is invertible.

(ii) Using R,

```
> set.seed(123)
> T = 10^3
> MA = arima.sim(list(ma=c(-⅚,⅙)), T)
> plot(MA, lwd=2, type="l", main="MA(2) Trajectory")
```




```
> arima(MA, order =
```

c(0,0,2))

```
Call:
arima(x = MA, order = c(0, 0, 2))

Coefficients:
          ma1     ma2  intercept
      -0.8627  0.1635     0.0043
s.e.   0.0314  0.0310     0.0094

sigma^2 estimated as 0.981:  log likelihood = -1409.75,  aic = 2827.51
```

From the results above, the estimate of α is -0.8627 and the estimate of β is 0.1635. To double check that the estimation procedure provides a good estimate of $\alpha = -5/6$ and $\beta = 1/6$, we shall calculate the confidence interval of α and β.

95% confidence interval for α: -0.8627 ± 1.96*0.0314 => (-0.924244, -0.801156)
95% confidence interval for β: 0.0314 ± 1.96*0.031 => (-0.02936, 0.09216)

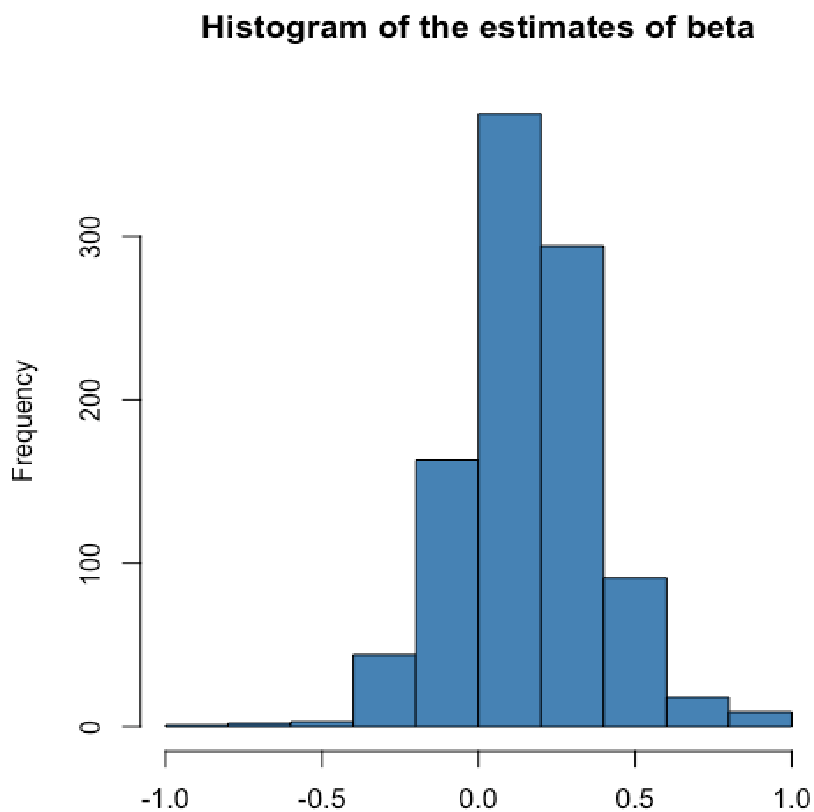
Since both the estimates of α and β falls within their 95% confidence interval, we can conclude that the estimation procedure provided good estimates of α and β.

(iii)

```
> set.seed(123)
> T = 40
> num_experiments = 1000
> beta_list = rep(0, num_experiments)
> for (i in 1:num_experiments){
  MA2 = arima.sim(list(ma=c(-5/6,1/6)),T)
  beta_list[i] = arima (MA2, order = c(0,0,2))$coef[2]
}
```

(iv)

> hist(beta_list, col="steelblue", xlim = c(-1,1), xlab = "Estimates of beta", main = "Histogram of the estimates of beta")



> mean(beta_list)

[1] 0.1556237

> var(beta_list)

[1] 0.04703875

The mean and variance of these estimates are 0.1556 and 0.04704 respectively rounded off to 4 significant figures.

(v)

> set.seed(123)

> T = 10^4

> MA = arima.sim(list(ma=c(3)),T)

> arima(MA, order = c(0,0,1))

```
Call:
arima(x = MA, order = c(0, 0, 1))

Coefficients:
         ma1  intercept
      0.3183    -0.0092
s.e.  0.0095     0.0395

sigma^2 estimated as 8.971:  log likelihood = -25159.65,  aic = 50325.31
```

From the results above, we know that the estimate for α is 0.3183.

For $X_k=W_k+3W_{k-1}$, the autocorrelation function is

ρ_k = 1 k=0

= $3/(1+3^2)=3/10$ k=1

= 0 k≥2

For $X_k=W_k+1/3W_{k-1}$, the autocorrelation function is

ρ_k = 1 k=0

= $(1/3)/(1+(1/3)^2)=3/10$ k=1

= 0 k≥2

The value of the estimate for α is very different from the true value of α = 3 because it is possible for 2 MA models to have the same autocorrelation function. As seen from the workings above, both the MA(1) model with α=1/3 and α=3 have the same autocorrelation function. In command **arima**, we assume the MA(1) model is invertible which indicates that the estimate of α is less than 1, which is 1/3 in this case. Therefore one obtains the estimate of α□ ≈ ⅓.

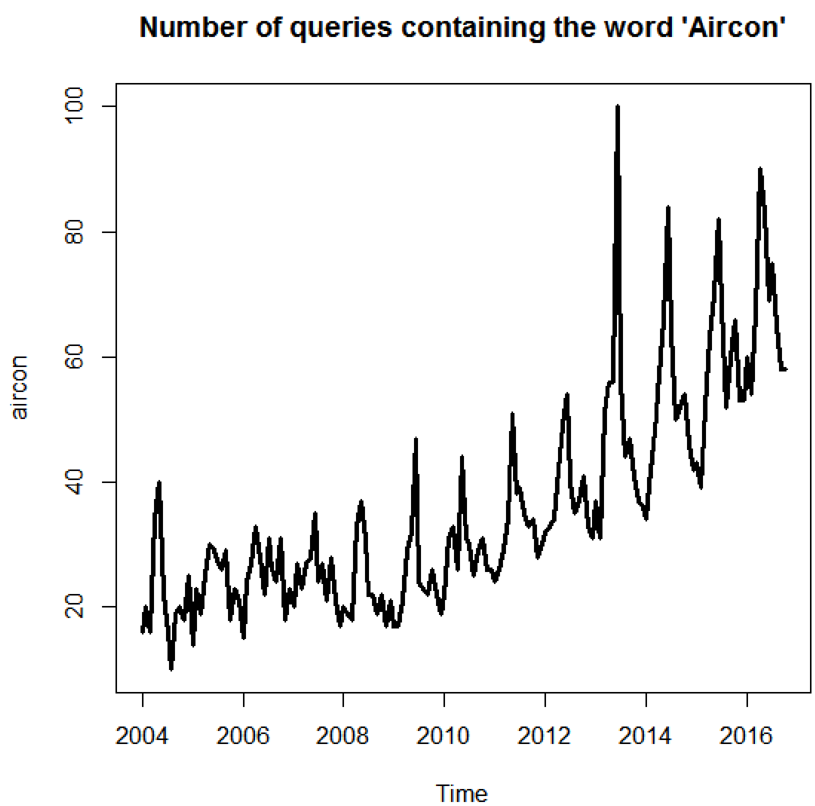
Exercise 3 (Sale forecast)

```
> aircon_data = read.csv("/Users/ST3233/Assignment/aircon.csv", skip=3,header =F)
> aircon = ts(as.numeric(aircon_data[,2]), start=c(2004,1), frequency=12)
```

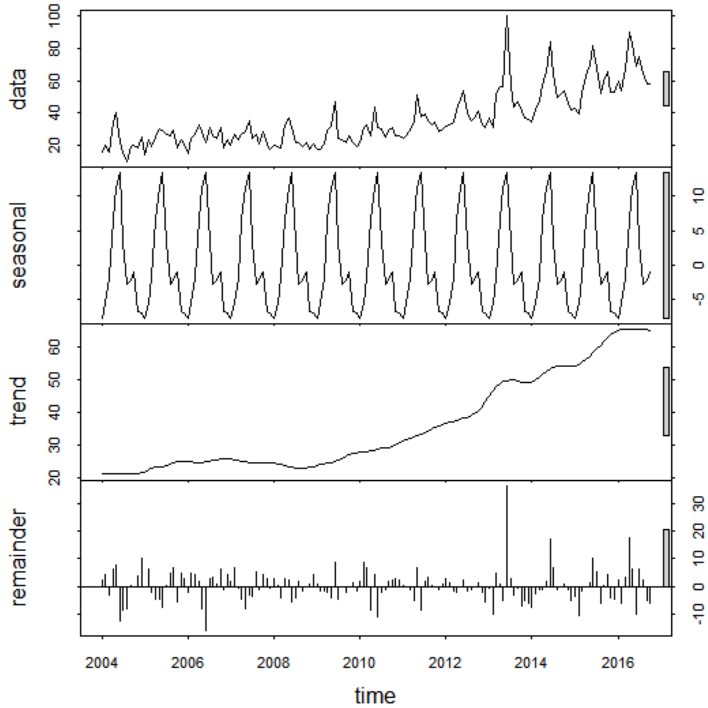
```
> aircon
     Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
2004  16  20  16  34  40  22  16  10  19  20  18  25
2005  14  23  19  25  30  29  27  26  29  18  23  21
2006  15  24  27  33  28  22  31  26  24  31  18  23
2007  20  27  23  27  28  35  24  27  21  28  21  17
2008  20  19  18  33  37  31  22  22  19  22  17  21
2009  17  17  21  29  32  47  24  23  22  26  22  19
2010  22  31  33  26  44  31  30  25  29  31  26  26
2011  24  26  29  34  51  38  39  35  33  34  28  30
2012  32  33  34  42  49  54  40  35  37  41  33  31
2013  37  31  51  56  56 100  56  44  47  41  37  36
2014  34  42  48  57  66  84  64  50  52  54  46  42
2015  43  39  52  63  70  82  68  52  60  66  53  53
2016  60  54  67  90  83  69  75  65  58  58
```

#Plot of the time series from 2004 to 2016

```
> ts.plot(aircon, main = "Number of queries containing the word 'Aircon'", lwd=3)
```




```
> aircon.decomp=stl(aircon,s.window="periodic")
> plot(aircon.decomp)
```

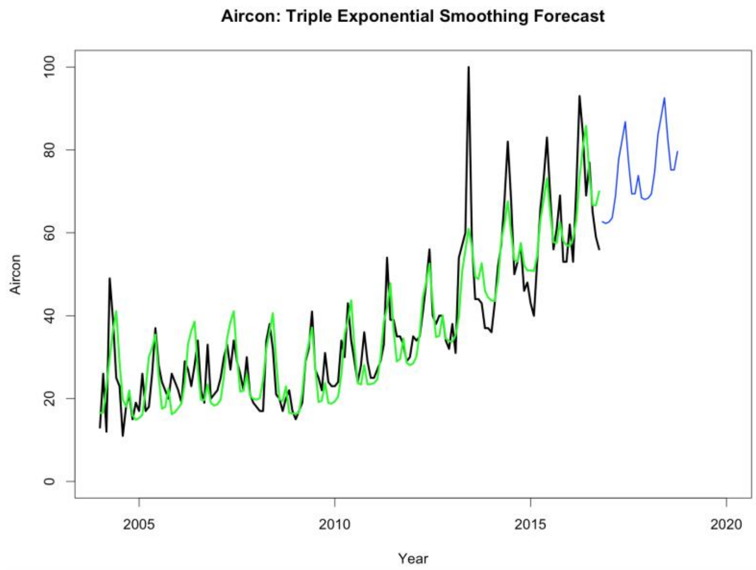


#Using Holt-Winter's algorithm, we can fit a triple exponential smoothing model to get the forecast for the next few months from Nov 2016 to June 2017

> fit_triple = hw(aircon, initial="optim", seasonal="additive",h=24)

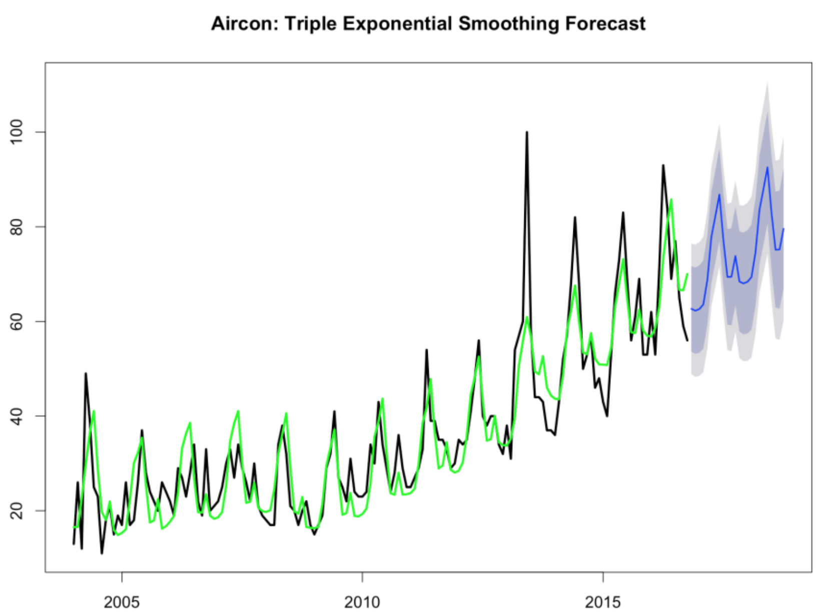
plot forecasts and fitted value

>plot(aircon, ylab="Aircon", xlim=c(2004, 2018),ylim=c(0,100),xlab="Year", main="Aircon: Triple Exponential Smoothing Forecast", type="l",lw=3)
> lines(fitted(fit_triple), col="green", type="l", lwd=3)
> lines(fit_triple$mean, col="blue", type="l",lwd=2)



plot the confidence interval

> plot(fit_triple, main="Aircon: Triple Exponential Smoothing Forecast", type="l",lw=3)
> lines(fitted(fit_triple), col="green", type="l", lwd=3)



Base on the forecasts using triple exponential smoothing, if the aircon retailer sold 51 devices in June 2016, he can expect a rise in his sales in June 2017.

Exercise 4 (Yule-Walker)

(i)

```
> q4data<-read.table("/Users/yanjia/Desktop/asg_1_MA3.dat")
> coeff= arima(q4data, order=c(3,0,0))
> coeff

Call:
arima(x = q4data, order = c(3, 0, 0))

Coefficients:
         ar1     ar2      ar3  intercept
      0.3373  0.2608  -0.0727    -0.0360
s.e.  0.0100  0.0102   0.0100     0.0213

sigma^2 estimated as 1.023:  log likelihood = -14301.43,  aic = 28612.86
```

$$\hat{p}\ (1) = 0.3373$$

$$\hat{p}\ (2) = 0.2608$$

$$\hat{p}\ (3) = -0.0727$$

(ii)

$$X_k = \alpha X_{k-1} + \beta X_{k-2} + \gamma X_{k-3} + W_k$$

$$\hat{p}\,(0) = 1$$

$$\hat{p}\,(1) = \alpha\hat{p}\,(0) + \beta\hat{p}\,(1) + \gamma\hat{p}\,(2) \Rightarrow \hat{p}\,(1) = \alpha + \beta\hat{p}\,(1) + \gamma\hat{p}\,(2)$$

$$\hat{p}\,(2) = \alpha\hat{p}\,(1) + \beta\hat{p}\,(0) + \gamma\hat{p}\,(1) \Rightarrow \hat{p}\,(2) = \alpha\hat{p}\,(1) + \beta + \gamma\hat{p}\,(1)$$

$$\hat{p}\,(3) = \alpha\hat{p}\,(2) + \beta\hat{p}\,(1) + \gamma\hat{p}\,(0) \Rightarrow \hat{p}\,(3) = \alpha\hat{p}\,(2) + \beta\hat{p}\,(1) + \gamma$$

(iii) Forming matrices to solve the above 3 equations to find $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$

$$\begin{array}{ccc} \hat{\alpha} & \hat{\beta} & \hat{\gamma} \\ \end{array}$$
$$\begin{pmatrix} 1 & 0.3373 & 0.2608 \\ 0.3373 & 1 & 0.3373 \\ 0.2608 & 0.3373 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 1.000 & 0.337 & 0.261 \\ 0.337 & 1.000 & 0.337 \\ 0.261 & 0.337 & 1.000 \end{pmatrix}$$

$$B = \begin{pmatrix} 0.337 \\ 0.261 \\ -0.073 \end{pmatrix}$$

$$A^{-1} = \begin{pmatrix} 1.160 & -0.326 & -0.193 \\ -0.326 & 1.220 & -0.326 \\ -0.193 & -0.326 & 1.160 \end{pmatrix}$$

$$X = A^{-1} \times B = \begin{pmatrix} 1.160 & -0.326 & -0.193 \\ -0.326 & 1.220 & -0.326 \\ -0.193 & -0.326 & 1.160 \end{pmatrix} \times \begin{pmatrix} 0.337 \\ 0.261 \\ -0.073 \end{pmatrix} = \begin{pmatrix} 0.320 \\ 0.232 \\ -0.234 \end{pmatrix}$$

Using Rstudio

```
> A<-matrix(c(1, 0.3373, 0.2608, 0.3373, 1, 0.3373, 0.2608, 0.3373, 1), nrow=3, ncol=3)
> B<-matrix(c(0.3373, 0.2608, -0.0727),nrow=3)
> solve(A,B)
           [,1]
[1,]  0.3202321
[2,]  0.2318561
[3,] -0.2344216
```

$\widehat{\alpha} = 0.3202321$ $\hat{\beta} = 0.2318561$ $\widehat{\gamma}$ = -0.2344216

R CODE

EXERCISE 1

data=read.table("D:/NUS/ST3233 Applied Time Series Analysis/datasets/electricity_load.dat")
electricitytimeseries=ts(data$V1,frequency=168,start=c(0))

```
plot(electricitytimeseries)
electricity.decomp=stl(electricitytimeseries,s.window="periodic",robust=TRUE)
plot(electricity.decomp)
trend=electricity.decomp$time.series[,"trend"]
plot(electricitytimeseries)
lines(trend,col="red")
library(fpp)
seasonplot(electricitytimeseries-trend,s=168,main="seasonal plot of detrended electricity")
electricitylog=log(electricitytimeseries)
electricitysmoothing=hw(electricitylog,h=336,initial="simple",seasonal="additive")
plot(electricitylog,main="holt-winters")
lines(fitted(electricitysmoothing),col="green")
lines(electricitysmoothing$mean,col="red")
plot(electricitysmoothing)
```

EXERCISE 2

```
set.seed(123)
T = 10^3
MA = arima.sim(list(ma=c(-5/6,1/6)),T)
plot(MA, lwd=2, type="l", main="MA(2) Trajectory")

arima(MA, order = c(0,0,2))

set.seed(123)
T = 40
num_experiments = 1000
beta_list = rep(0, num_experiments)
for (i in 1:num_experiments){
  MA2 = arima.sim(list(ma=c(-5/6,1/6)),T)
  beta_list[i] = arima (MA2, order = c(0,0,2))$coef[2]
 }

hist(beta_list, col="steelblue",
     xlim = c(-1,1), xlab = "Estimates of beta",
     main = "Histogram of the estimates of beta")
mean(beta_list)
var(beta_list)

set.seed(123)
T = 10^4
MA = arima.sim(list(ma=c(3)),T)
arima(MA, order = c(0,0,1))
```

EXERCISE 3

```
aircon_data = read.csv("/Users/limweiqi/Desktop/aircon.csv", skip=3,header =F)
aircon = ts(as.numeric(aircon_data[,2]), start=c(2004,1), frequency=12)
ts.plot(aircon, main = "Number of queries containing the word 'Aircon'", lwd=3)
aircon.decomp=stl(aircon,s.window="periodic")
plot(aircon.decomp)
fit_triple = hw(aircon, initial="optim", seasonal="additive",h=24)

plot(aircon, ylab="Aircon", xlim=c(2004, 2018),ylim=c(0,100),xlab="Year", main="Aircon: Triple Exponential Smoothing Forecast", type="l",lw=3)
lines(fitted(fit_triple), col="green", type="l", lwd=3)
lines(fit_triple$mean, col="blue", type="l",lwd=2)
plot(fit_triple, main="Aircon: Triple Exponential Smoothing Forecast", type="l",lw=3)
lines(fitted(fit_triple), col="green", type="l", lwd=3)
```