TIME SERIES Assignment

The attached data shows monthly demand of two different types of consumable items in a certain store from January 2002 to September 2017. The ultimate objective of this exercise is to predict sales for the period October 2017 to December 2018.

1. Read the data as time series objects in R. Plot the data. What are the major features you notice in the series? How do the two series differ?

Reading the Data Set In Time Series:

RCODE:

getwd()

data <- read\_excel("~/Desktop/grate lakes/Demand.xlsx")

head(data)

summary(data)

str(data)

data$`Item A`=as.integer(data$`Item A`)

data$`Item B`=as.integer(data$`Item B`)

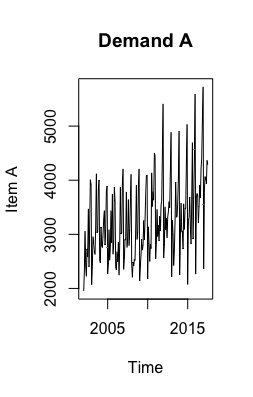
str(data)

#visualisation

par(mfrow=c(2,2))

demandA <- ts(data[,3], start=c(2002,1), frequency=12)

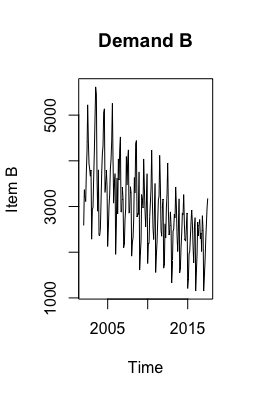
plot(demandA)



Conclusion= From the plot of demand A we conclude that it is showing the characteristic of both additive and multiplicative(from the year 2015 onwards).It is also showing kind of trend with variation.

demandB <- ts(data[,4], start=c(2002,1), frequency=12)

plot(demandB)



CONCLUSION: Conclusion= From the plot of demand B we conclude that it is showing the characteristic of multiplicative. It is also showing seasonality and negative trend.

1. Any missing value in the series? How do you tackle missing values in the series?

Conclusion: Yes there was missing values in Demand A and demand B. In Demand A we replace the missing values with average of December values.

In Demand B we replace the missing values with the adjacent values of demand B.

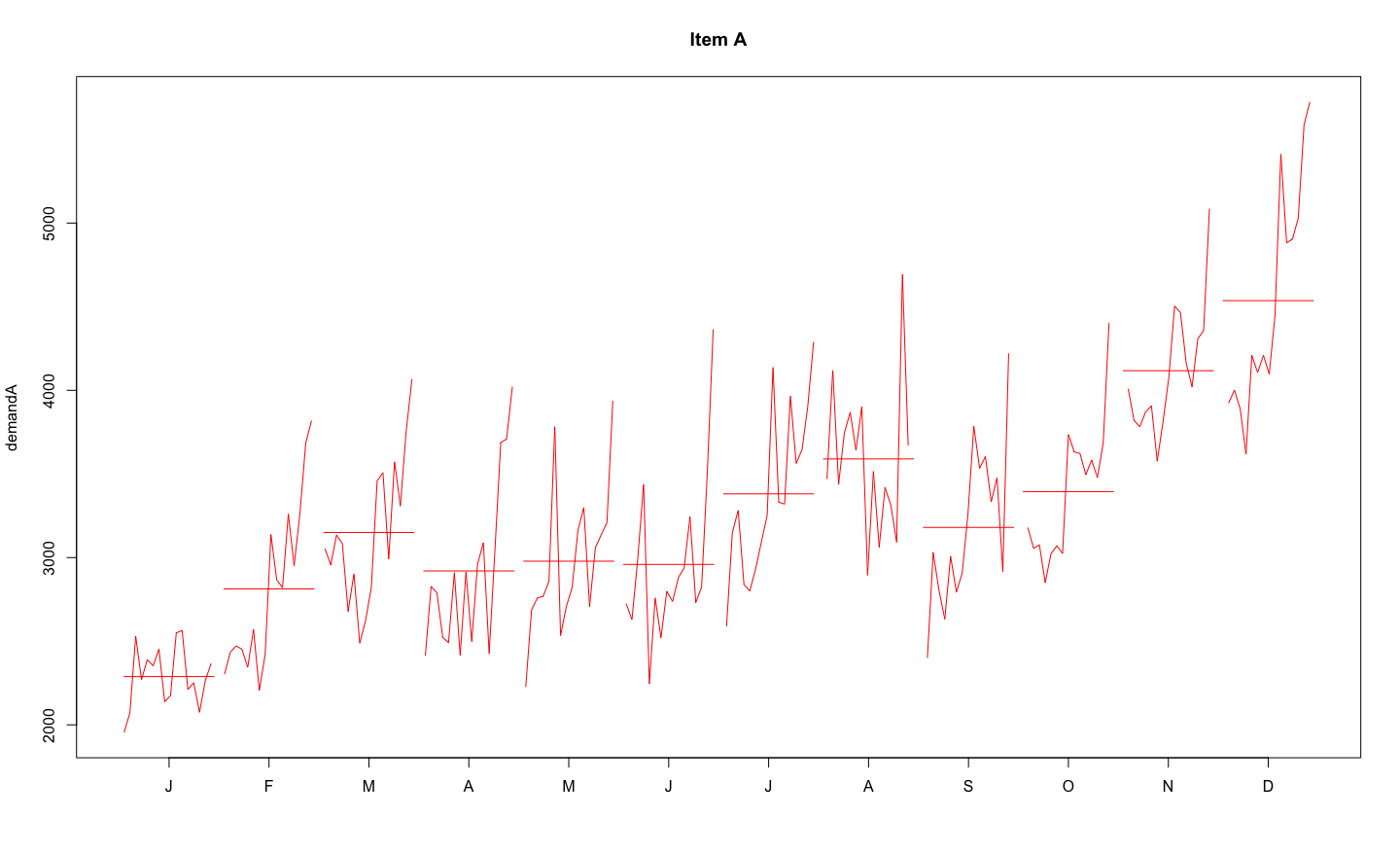
1. Before a formal extraction of time series components is done, can you check for seasonal changes in the data for the two series separately? Particularly whether there are more variability in a season compared to the others, whether seasonal variations are changing across years etc. Compare the behavior of the two series.

MONTH PLOT:

demandA <- ts(data[,3], start=c(2002,1), frequency=12)

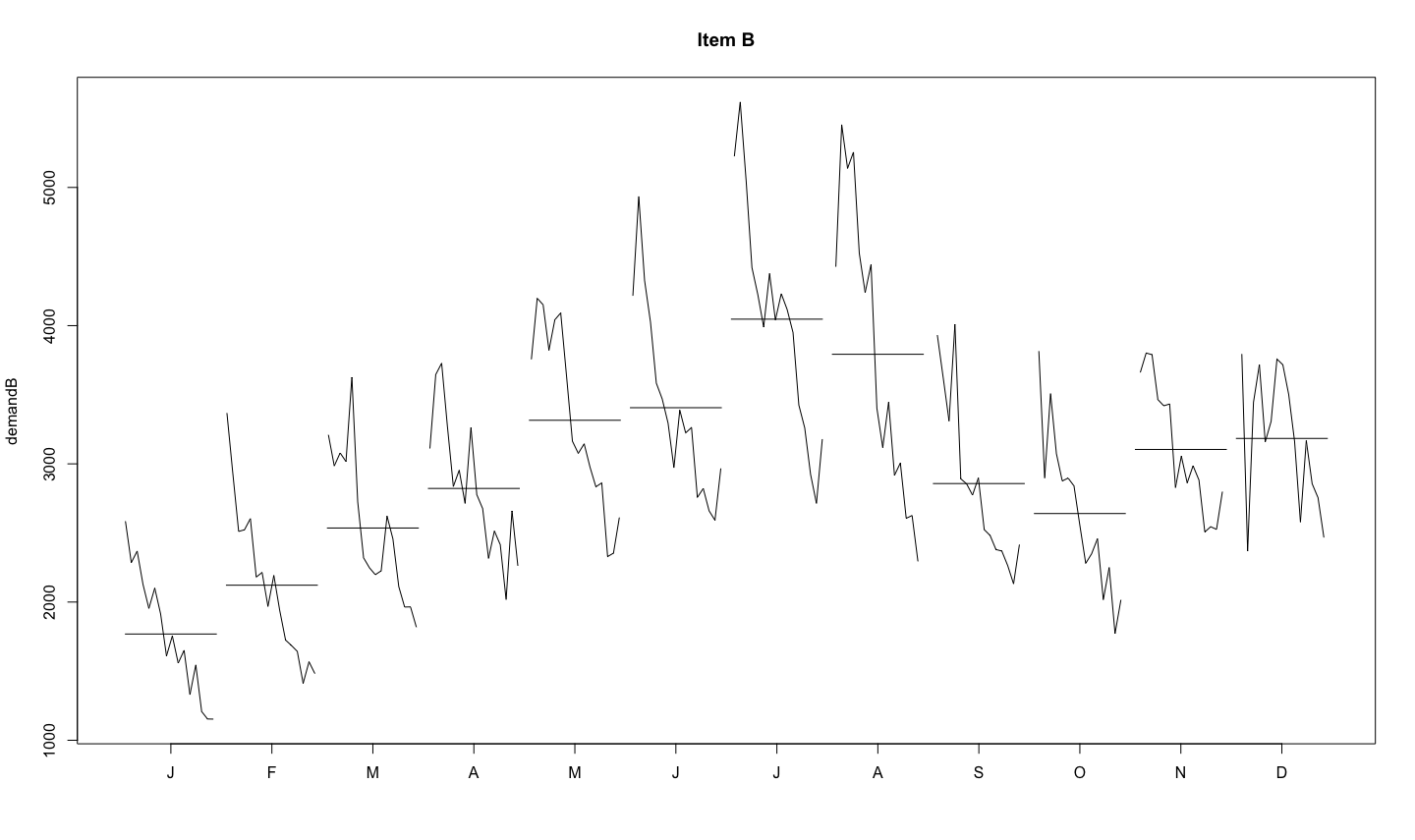
demandB <- ts(data[,4], start=c(2002,1), frequency=12)

monthplot(demandA,main="Item A")



Conclusion: In the month of November and December there is increase in demand because of seasonality effect. There is also overall increase in the demand.

monthplot(demandB,main="Item B")



Conclusion: There is highest demand in the month of July and August because of seasonality effect.

1. Decompose each series to extract trend and seasonality, if there are any. Which seasonality is more appropriate – additive or multiplicative? Explain the seasonal indices. In which month(s) do you see higher sales and which month(s) you see lower sales? Any difference in the nature of demand of the two items?

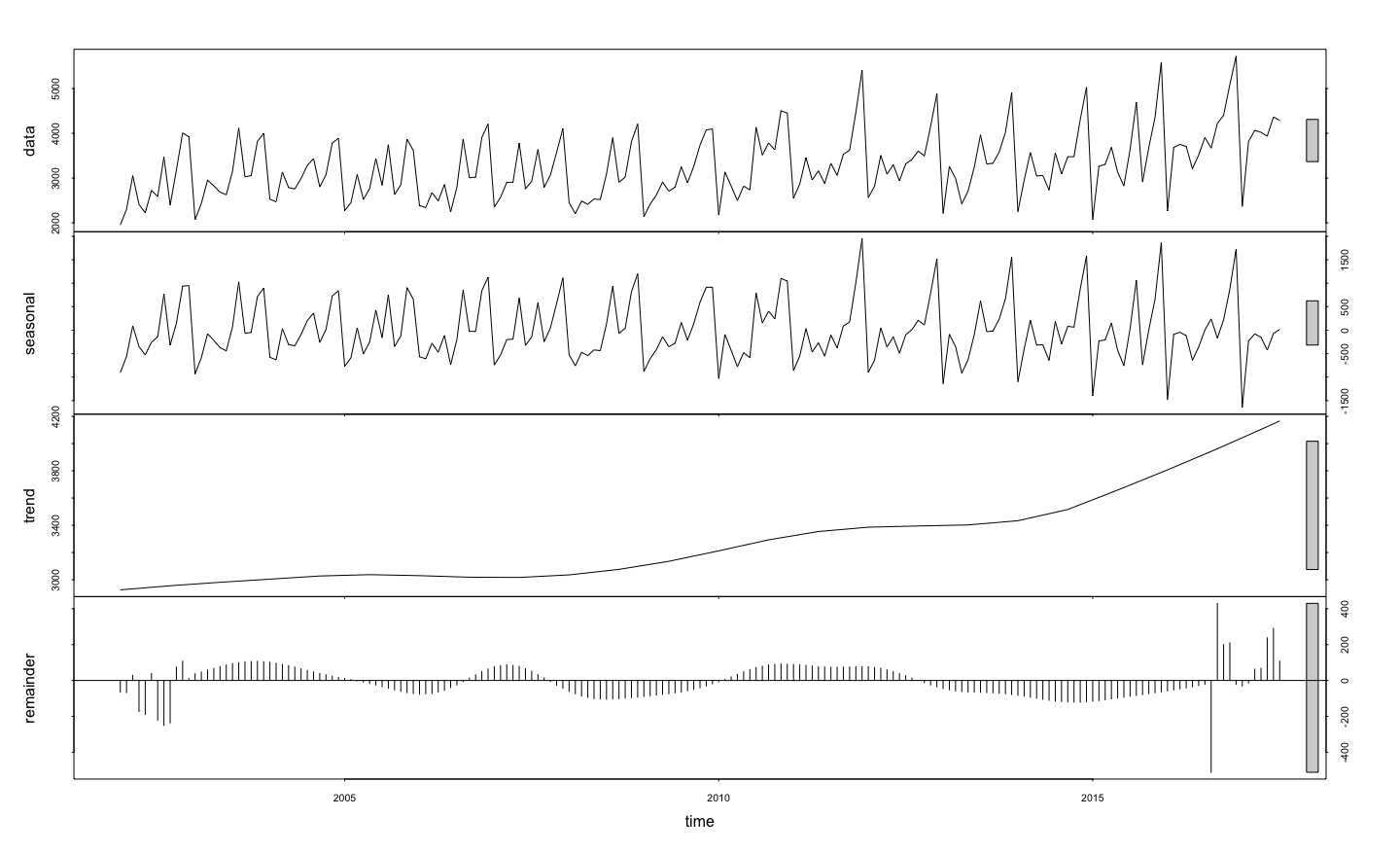
RCODE:

FOR DEMAND A:

decompose\_A<-stl(demandA[,1],s.window=2)

plot(decompose\_A) # additive

DEMAND A = SEASONALITY+TREND+RESIDUAL



Which seasonality is more appropriate – additive or multiplicative:

Ans) Additive seasonality is more appropriate for Demand A as there is more or less constant variation in trend is there. It is also showing the characteristic of multiplicative seasonality (from the year 2015 onwards).

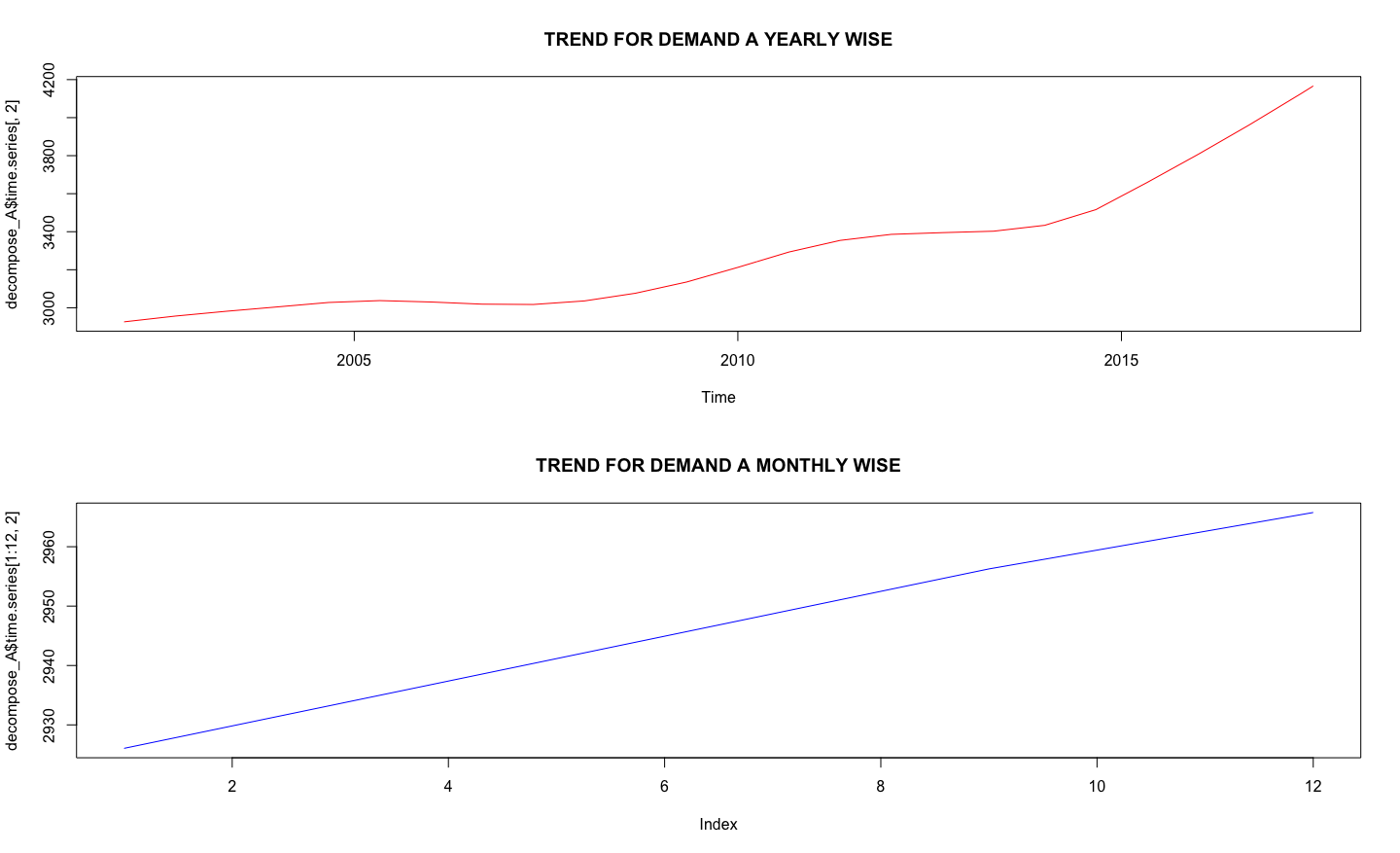
TREND FOR DEMAND A:

RCODE:

decompose\_A<-stl(demandA[,1],s.window=2)

plot(decompose\_A$time.series[,2],main='TREND FOR DEMAND A YEARLY WISE',col='red') ## trend

plot(decompose\_A$time.series[1:12,2],main='TREND FOR DEMAND A MONTHLY WISE',col='BLUE',type='l')



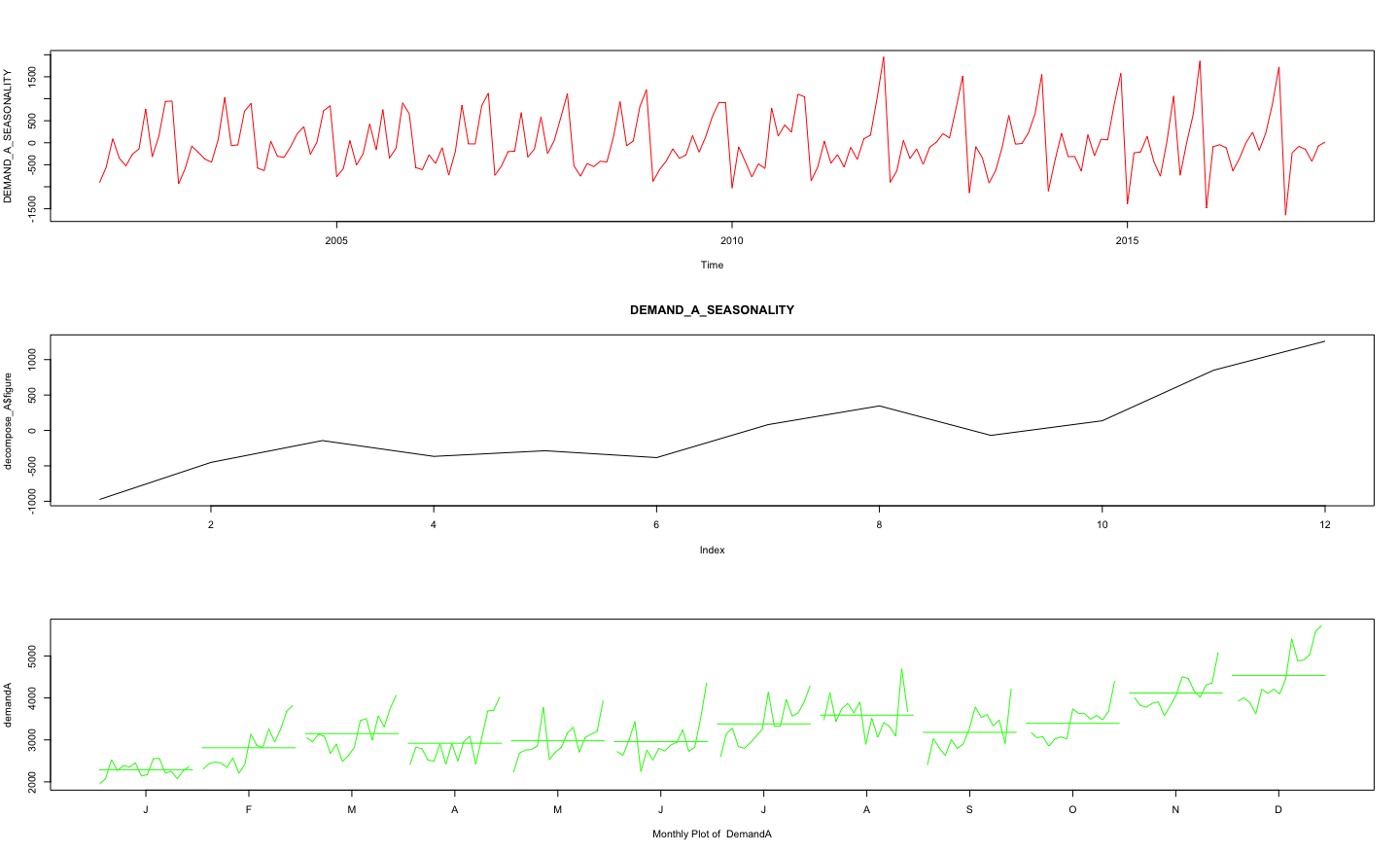
Conclusion: **There is an overall increase in trend for the demand A from the year 2002 to 2017.It means that the demand for the consumable item A increases over the period of time.**

(SEASONALITY + MONTHLY) GRAPH FOR DEAMAMD\_A:

R CODE:

plot(decompose\_A$time.series[,1],ylab='DEMAND\_A\_SEASONALITY',col='red')

monthplot(demandA,xlab="Monthly Plot of DemandA",col='green')



In which month(s) do you see higher sales and which month(s) you see lower sales for Demand A?

Ans) Heigher Sales – DECEMBER (Average Demand Around 4400)

Lower Sales - JANUARY (Average Demand Around 2300)

DEMAND B=(SEASONALITY+TREND+RESIDUAL)

R CODE:

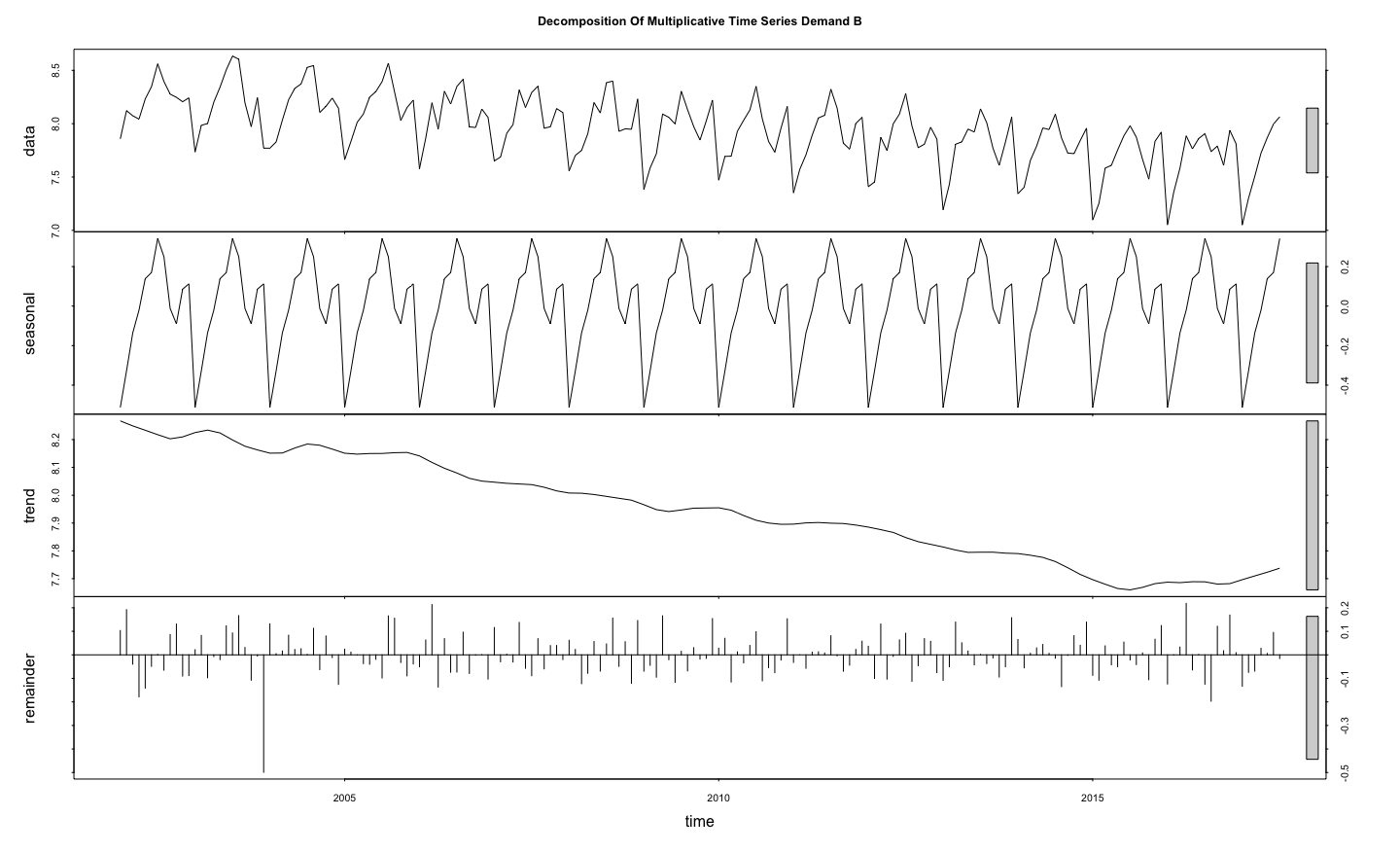
# Analysis of a multiplicative series

logDemandB <- log(demandB)

decompose\_b <- stl(logDemandB[,1], s.window="p")

decompose\_b$time.series[1:12,1]decompose\_b

plot(decompose\_b)



Which seasonality is more appropriate – additive or multiplicative:

Ans) Multiplicative seasonality is more appropriate for Demand B as there is variation in trend over the period of time.

TREND FOR DEMAND B

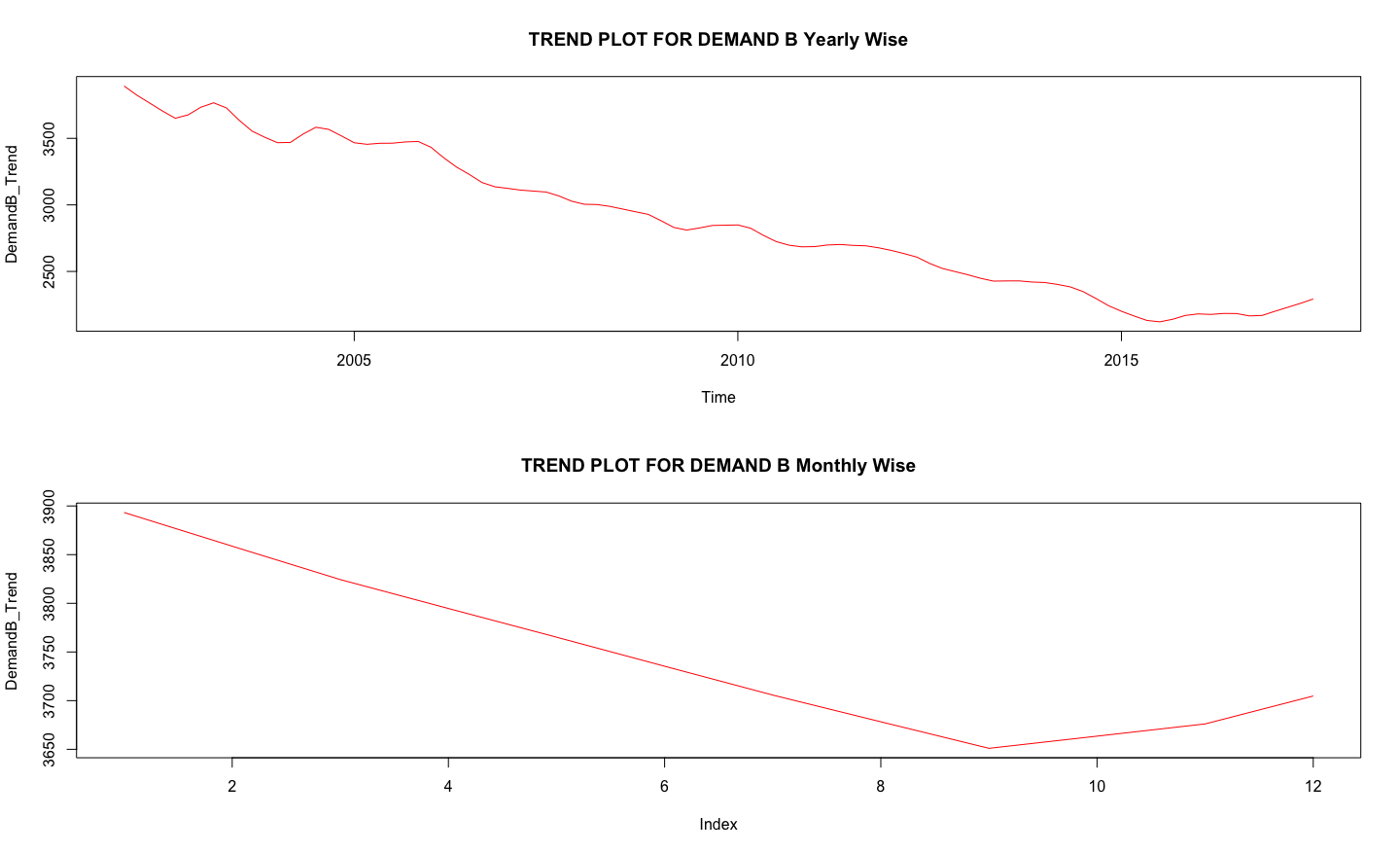
RCODE

DemandB\_Trend<- exp(decompose\_b$time.series[,2])

plot(DemandB\_Trend, type="l",col='red',main = 'TREND PLOT FOR DEMAND B Yearly Wise')

DemandB\_Trend<- exp(decompose\_b$time.series[1:12,2])

plot(DemandB\_Trend, type="l",col='red',main = 'TREND PLOT FOR DEMAND B Monthly Wise')



Conclusion: **There is an overall decrease in trend for the demand B from the year 2002 to 2017 but from the year 2015(August) onwards there is slightly increase in trend .**

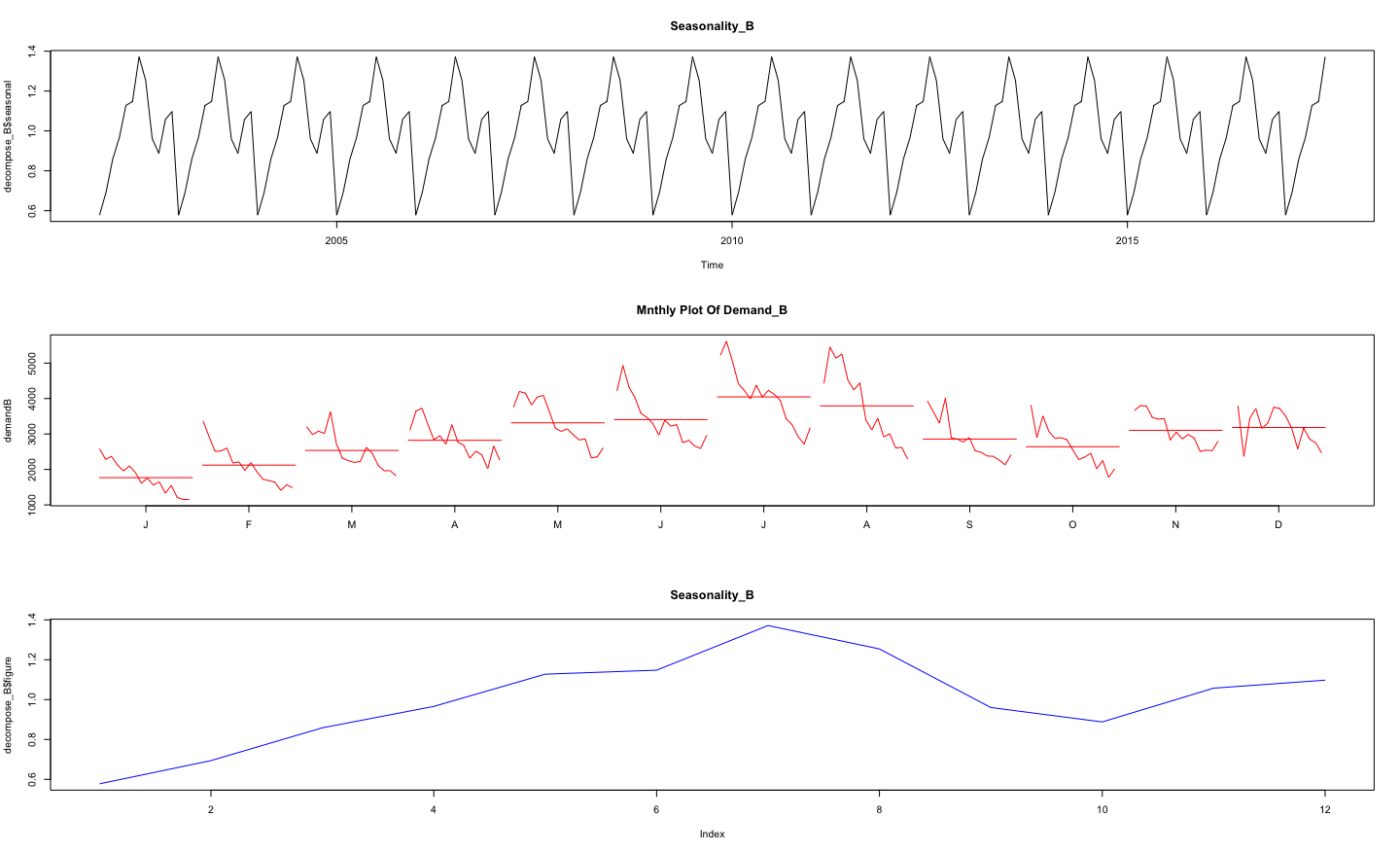
**It means that the demand for the consumable item B decreases over the period of time (2002 to 2015) and then demand started increasing slightly.**

(SEASONALITY + MONTHLY) GRAPH FOR DEAMAMD B

RCODE

DemandB\_Seasonality<- exp(decompose\_b$time.series[,1])

plot(DemandB\_Trend, type="l",col='red',main = 'Seasonality PLOT FOR DEMAND B Monthly Wise')



In which month(s) do you see higher sales and which month(s) you see lower sales for Demand B(consumable item B) ?

Ans) Heigher Sales – JULY (Average Demand Around 4000)

Lower Sales - JANUARY (Average Demand Around 1800)

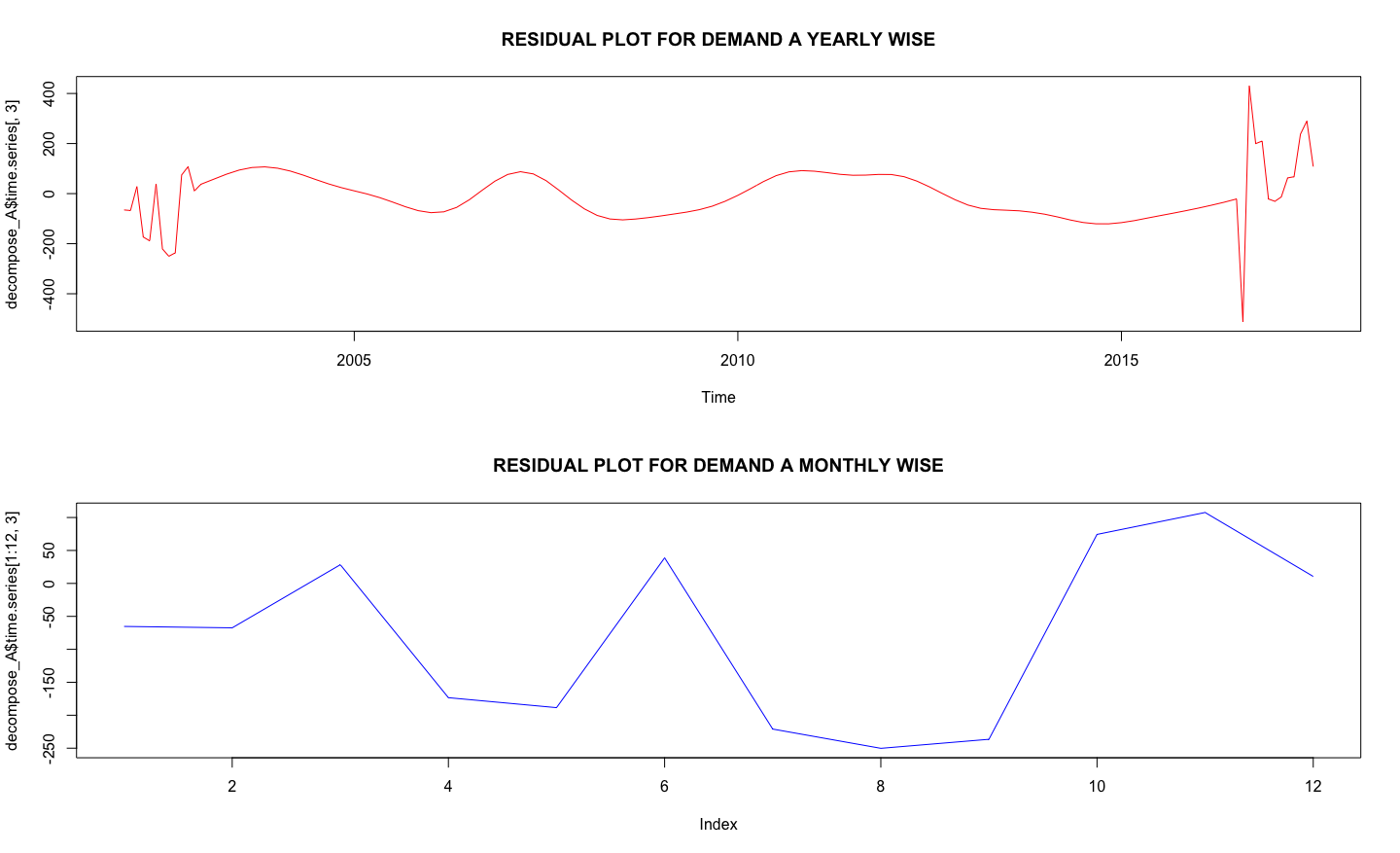
5) Can you extract the residuals for the two decomposition exercises? Comment on the pattern of the residuals?

R CODE:

# RESIDUAL A

plot(decompose\_A$time.series[,3],main='RESIDUAL PLOT FOR DEMAND A YEARLY WISE',col='red')

plot(decompose\_A$time.series[1:12,3],main='RESIDUAL PLOT FOR DEMAND A MONTHLY WISE',col='blue',type='l')



RCODE:

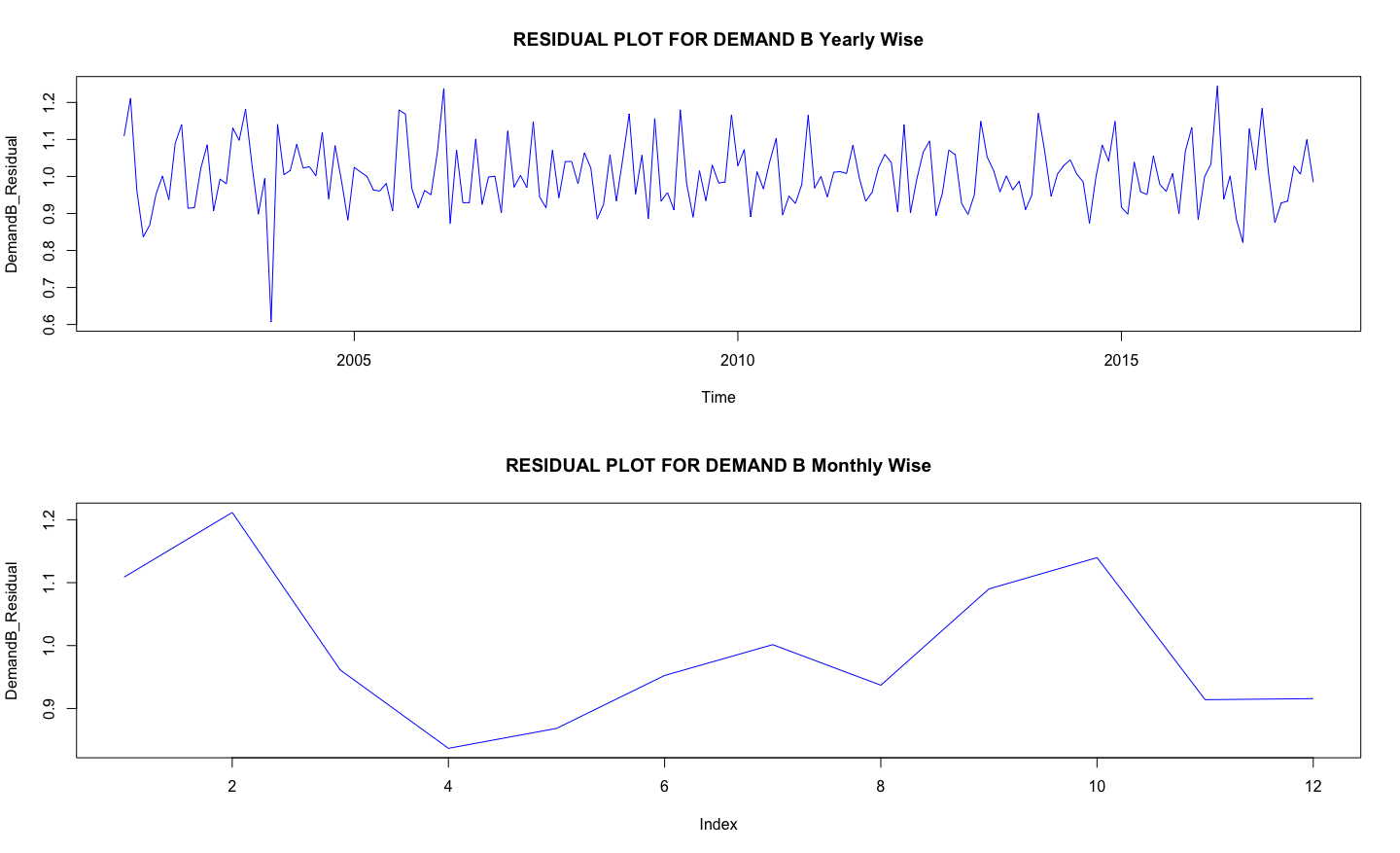
RESIDUAL B:

DemandB\_Residual <- exp(decompose\_b$time.series[,3])

plot(DemandB\_Residual, type="l",col='blue',main = 'RESIDUAL PLOT FOR DEMAND B ')

DemandB\_Residual <- exp(decompose\_b$time.series[1:12,3])

plot(DemandB\_Residual, type="l",col='blue',main = 'RESIDUAL PLOT FOR DEMAND B )



1. Before the final forecast is undertaken one would like to compare a few models. Use the last 21 months as hold-out sample fit a suitable exponential smoothing model to the rest of the data and calculate MAPE. What are the values of α, β and γ? What role do they play in the modeling? For the same hold-out period compare forecast by decomposition and compute MAPE. Which model gives smaller MAPE? Give a comparison for the two demands.

Ans)

RCODE:

# Dividing a time series into train and test

ATrain <- window(demandA, start=c(2002,1), end=c(2015,10), frequency=12)

ATest <- window(demandA, start=c(2015,11), frequency=12)

## Applying holt winters

#As multiplicative

ATrain.fc = hw(ATrain, seasonal = 'm',h=21)

> ATrain.fc

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

Nov 2015 4460.961 3862.696 5059.225 3545.994 5375.927

Dec 2015 4782.958 4137.958 5427.958 3796.516 5769.401

Jan 2016 2591.259 2239.836 2942.681 2053.804 3128.713

Feb 2016 3008.551 2598.168 3418.934 2380.924 3636.177

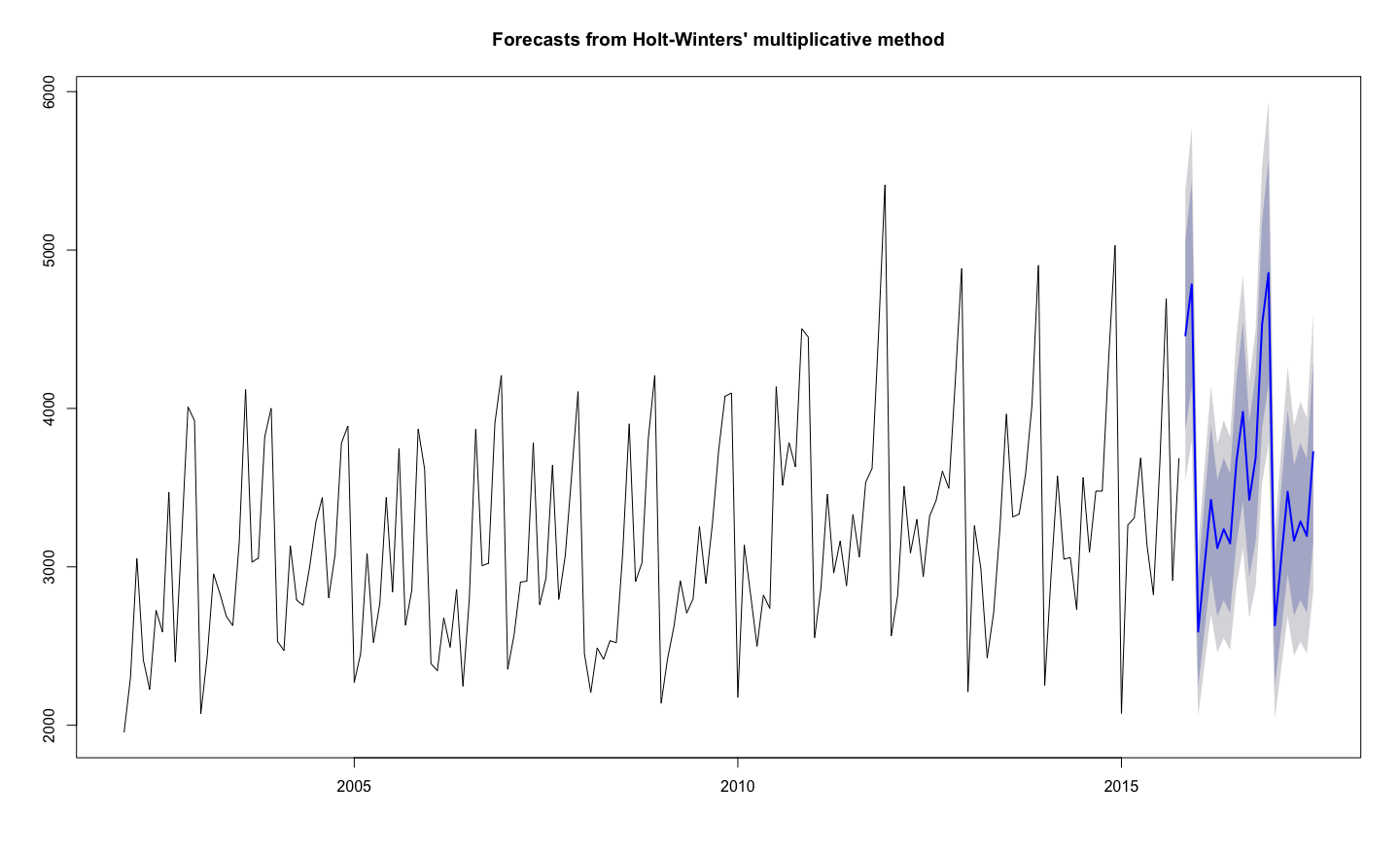
Mar 2016 3422.540 2952.918 3892.163 2704.315 4140.766

Apr 2016 3118.320 2687.849 3548.791 2459.972 3776.669

May 2016 3237.447 2787.769 3687.126 2549.723 3925.172

Jun 2016 3147.717 2707.746 3587.688 2474.840 3820.594

plot(ATrain.fc)



ATrain.fc$model

Call:

hw(y = ATrain, h = 21, seasonal = "m")

Smoothing parameters:

alpha = 0.103

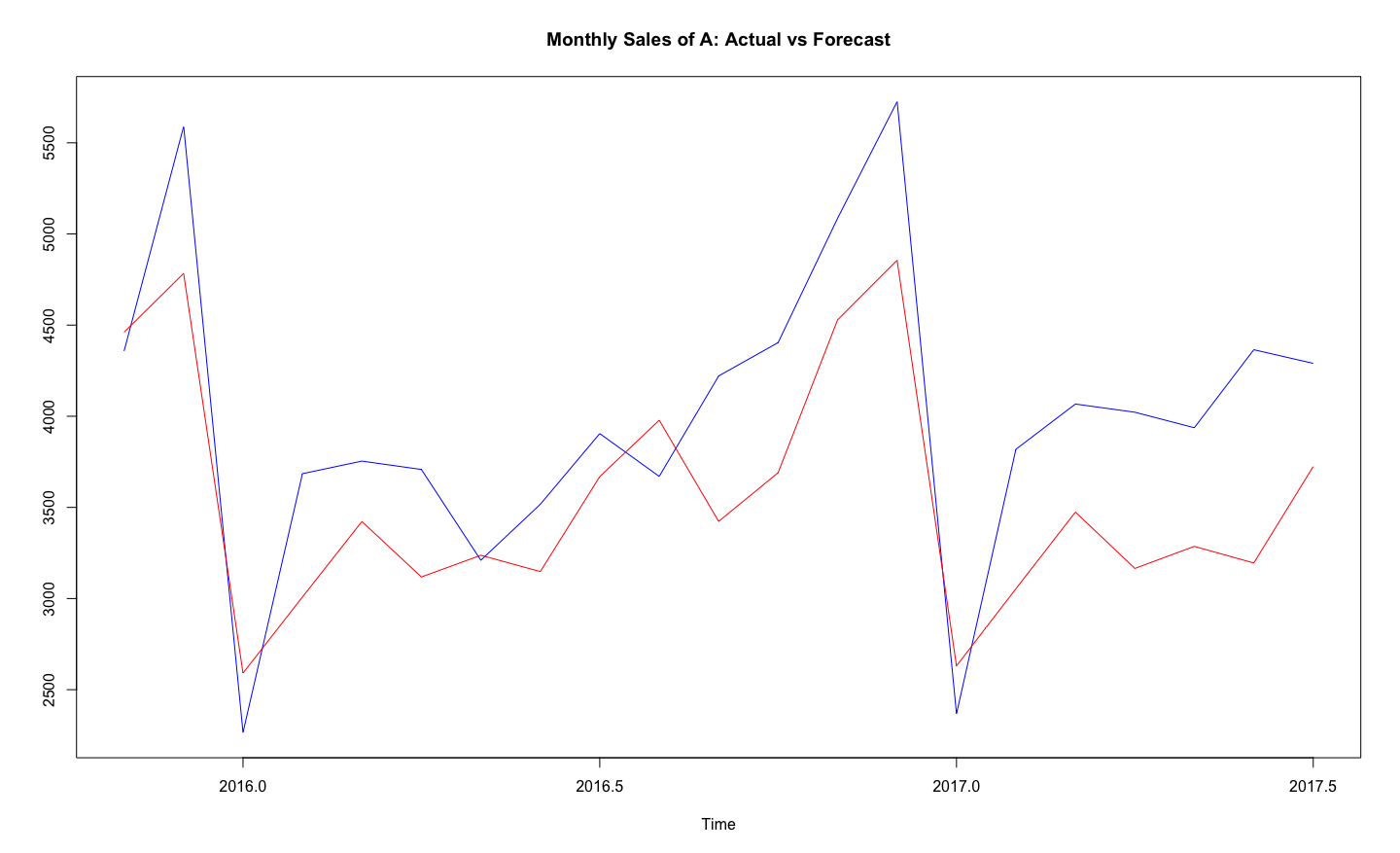
beta = 0.002

gamma = 1e-04

**#Checking MAPE**

Vec<- cbind(ATest,ATrain.fc$mean)

ts.plot(Vec, col=c("blue", "red"), main="Monthly Sales of A: Actual vs Forecast")



MAPE <- mean(abs(Vec[,1]-Vec[,2])/Vec[,1])

MAPE

13.5 %

#Adjusting alpha, beta, gamma value to get a better model

ATrain.fc = hw(ATrain, seasonal = 'm',h=21,alpha=0.04,beta=0.05,gamma=0.4)

Vec<- cbind(ATest,ATrain.fc$mean)

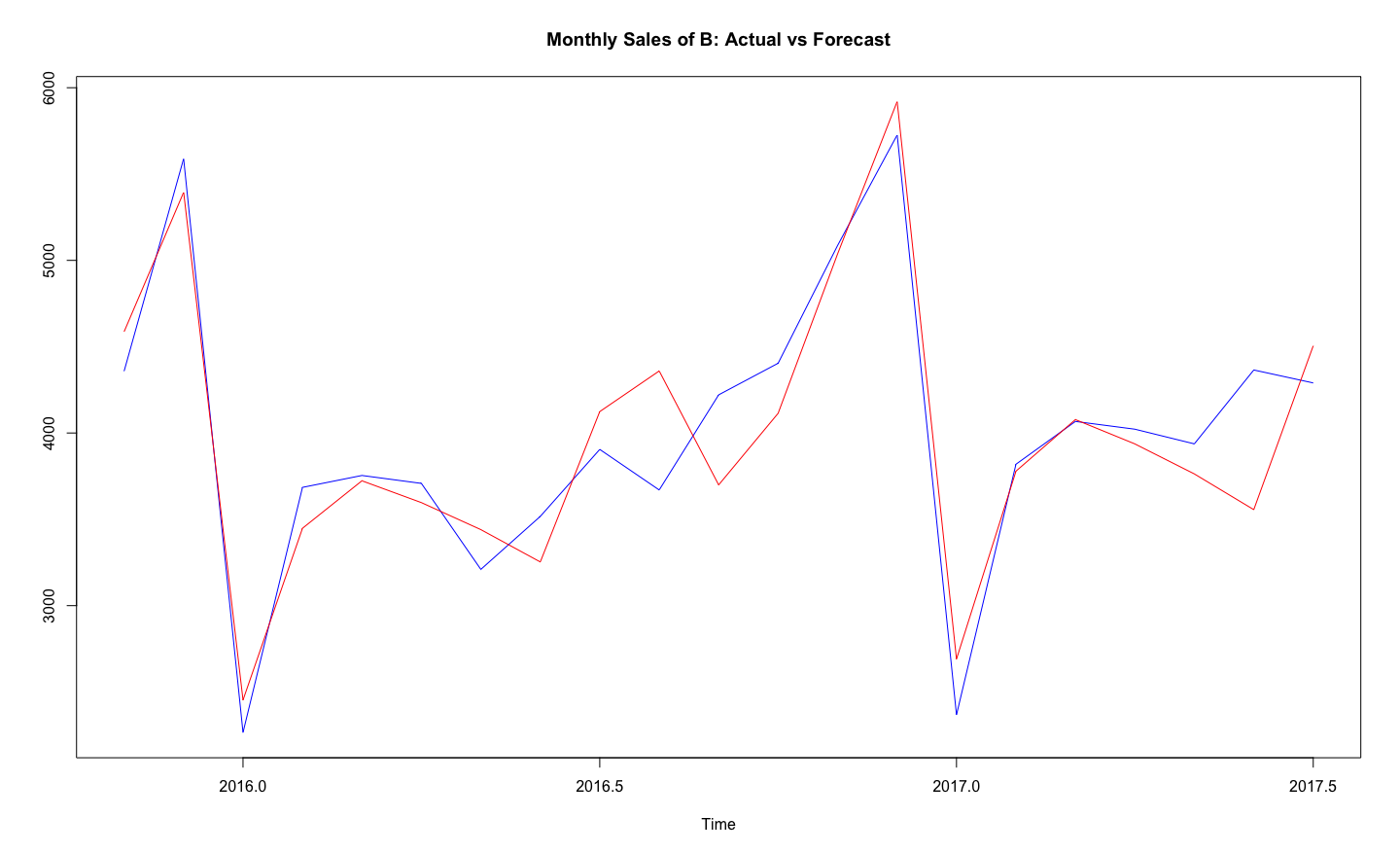
ts.plot(Vec, col=c("blue", "red"), main="Monthly Sales of B: Actual vs Forecast")

Smoothing parameters:

alpha = 0.04

beta = 0.05

gamma = 0.4



MAPE <- mean(abs(Vec1[,1]-Vec1[,2])/Vec1[,1])

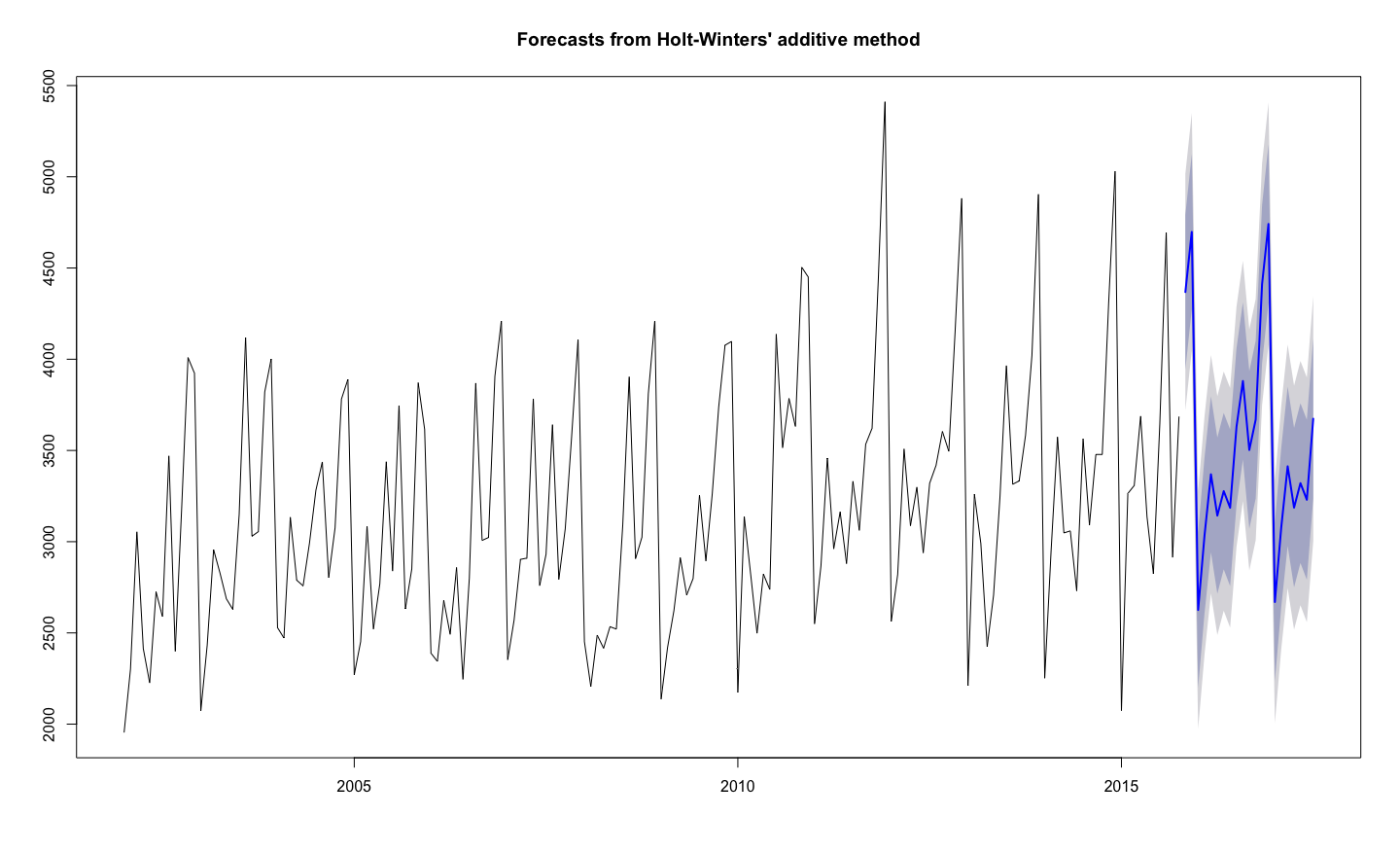
MAPE

**6.38%**

# NOW CONSIDERING A AS AN ADDITIVE

ATrain.fca = hw(ATrain, seasonal = 'a',h=21)

ATrain.fca

plot(

ATrain.fca$model

Smoothing parameters:

alpha = 0.0587

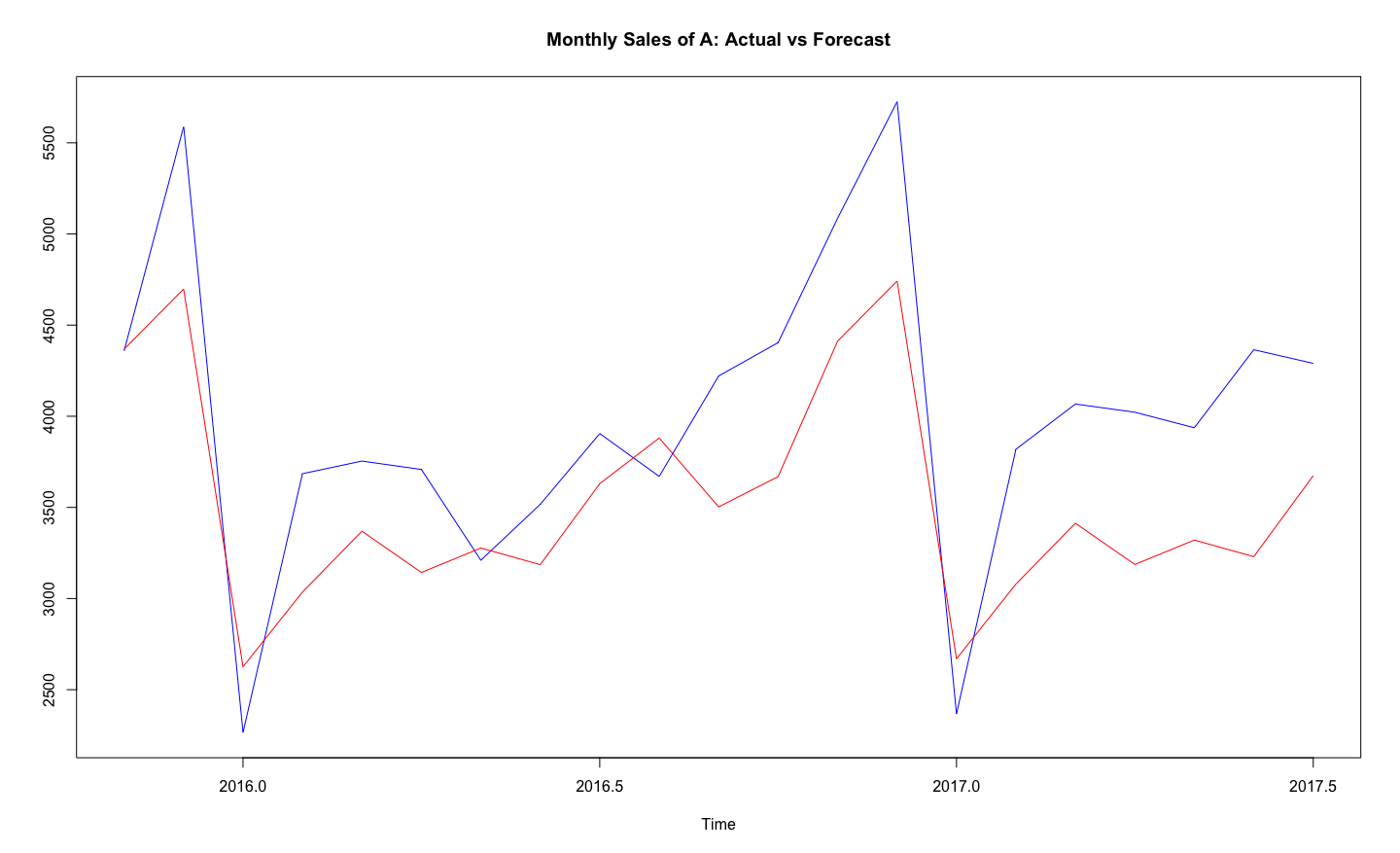
beta = 1e-04

gamma = 0.0064

**#Checking MAPE**

Vec<- cbind(ATest,ATrain.fca$mean)

ts.plot(Vec, col=c("blue", "red"), main="Monthly Sales of A: Actual vs Forecast")



MAPE <- mean(abs(Vec[,1]-Vec[,2])/Vec[,1])

MAPE

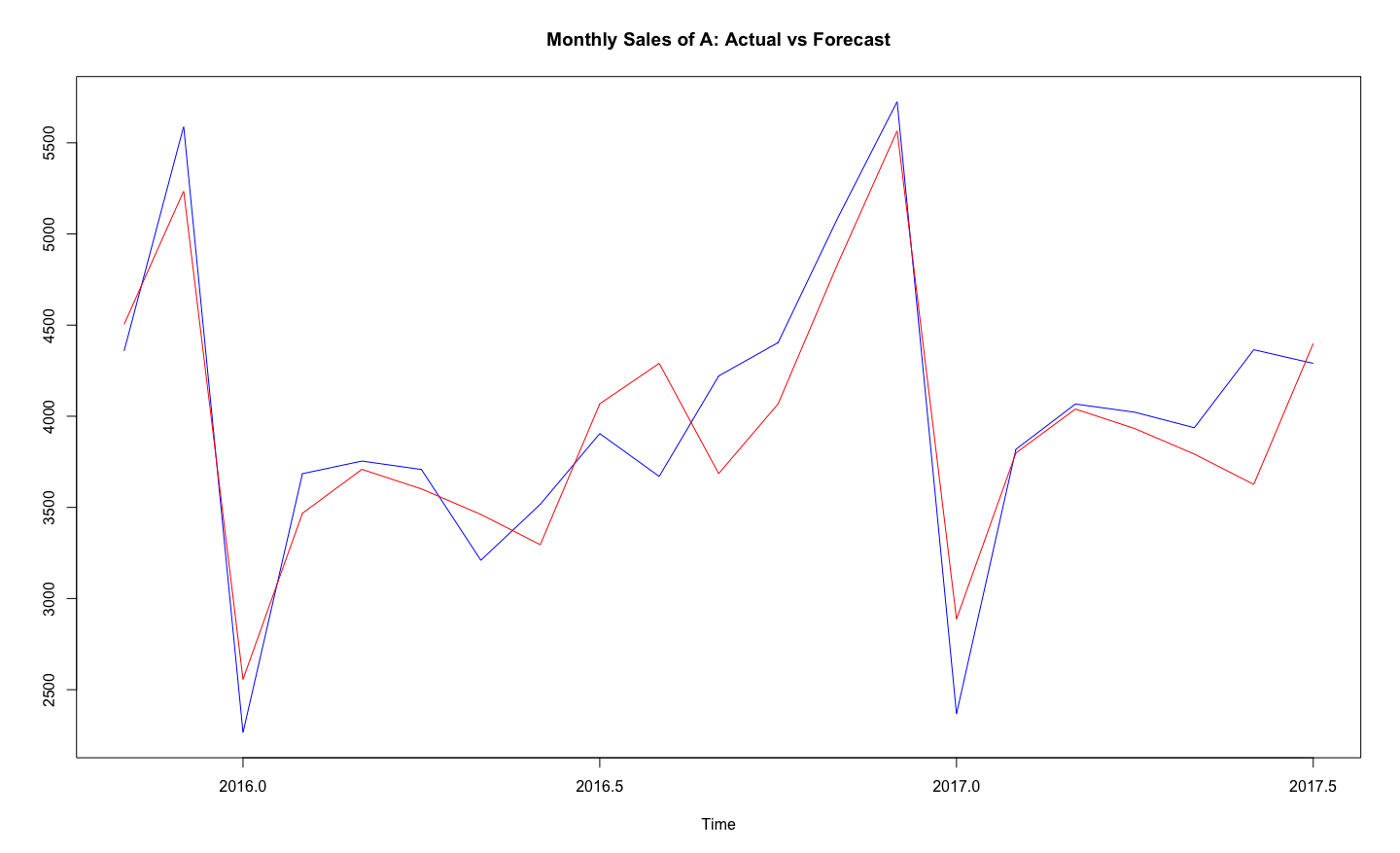
13.74%

#Adjusting alpha, beta, gamma value to get a better model

ATrain.fca1 = hw(ATrain, seasonal = 'a',h=21,alpha=0.055,beta=0.035,gamma=0.4)

Vec<- cbind(ATest,ATrain.fca1$mean)

ts.plot(Vec, col=c("blue", "red"), main="Monthly Sales of A: Actual vs Forecast")



MAPE <- mean(abs(Vec[,1]-Vec[,2])/Vec[,1])

**MAPE**

**6.9%**

**CONCLUSION:**

FROM THE ABOVE MODELLING WE GET THE LESS MAPE VALUE IN THE CASE OF MULTIPLICATIVE AS COMPARE TO ADDTIVE.THEREFORE I AM CONSIDERING MUTTIPLICATIVE MODEL FOR THE FORCASTING OF DEMAND A.

MAPE=6.38%

Smoothing parameters

alpha = 0.04 beta = 0.05 gamma = 0.4

**NOW THE SAME ABOVE PROCEDURE FOR ITEM B:**

## Dividing a time series into train and test

BTrain <- window(demandB, start=c(2002,1), end=c(2015,10), frequency=12)

BTest <- window(demandB, start=c(2015,11), frequency=12)

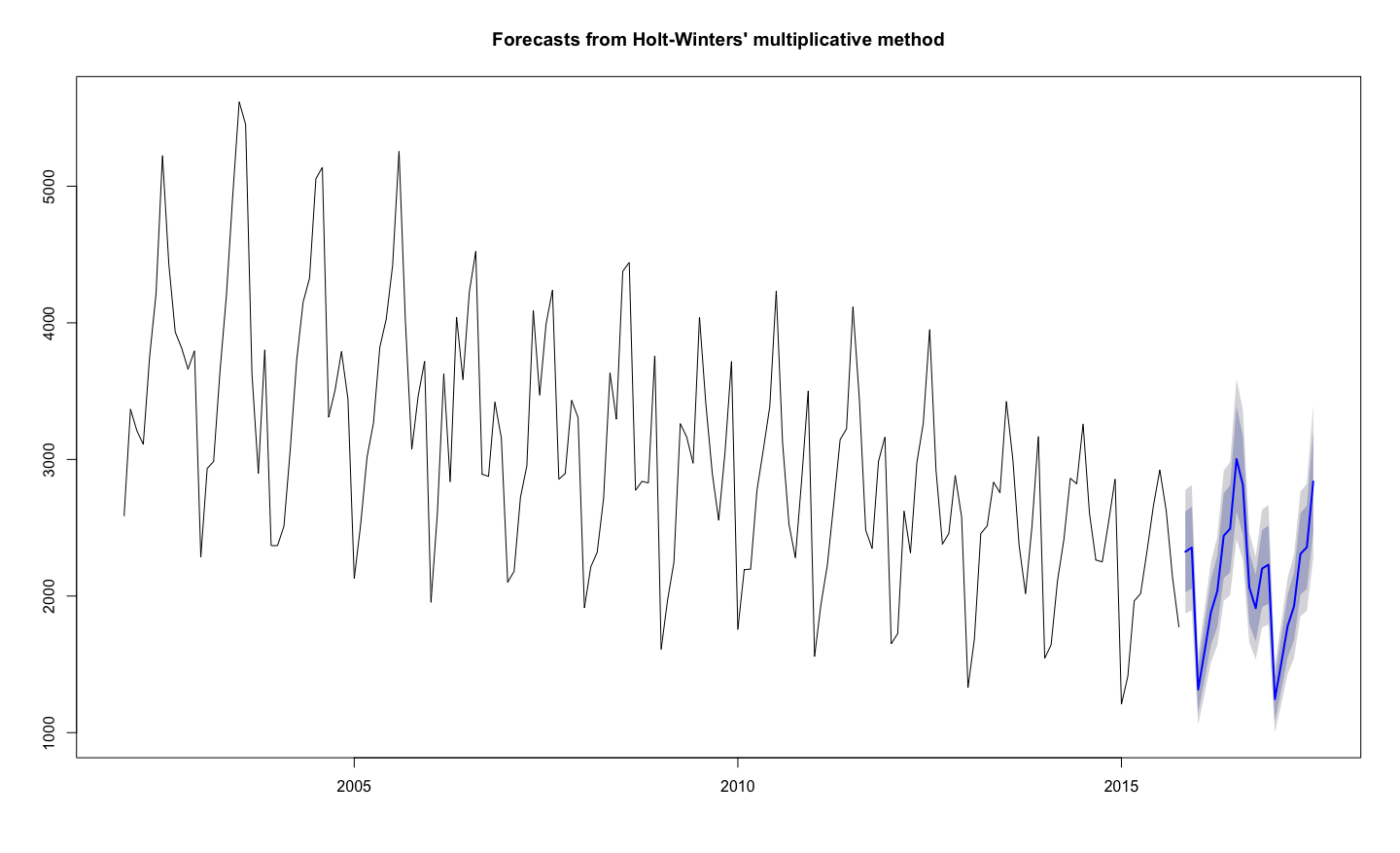
**#Applying holt winters**

**#As multiplicative**

BTrain.fc = hw(BTrain, seasonal = 'm',h=21)

BTrain.fc

plot(BTrain.fc)



BTrain.fc$model

Smoothing parameters:

alpha = 0.0218

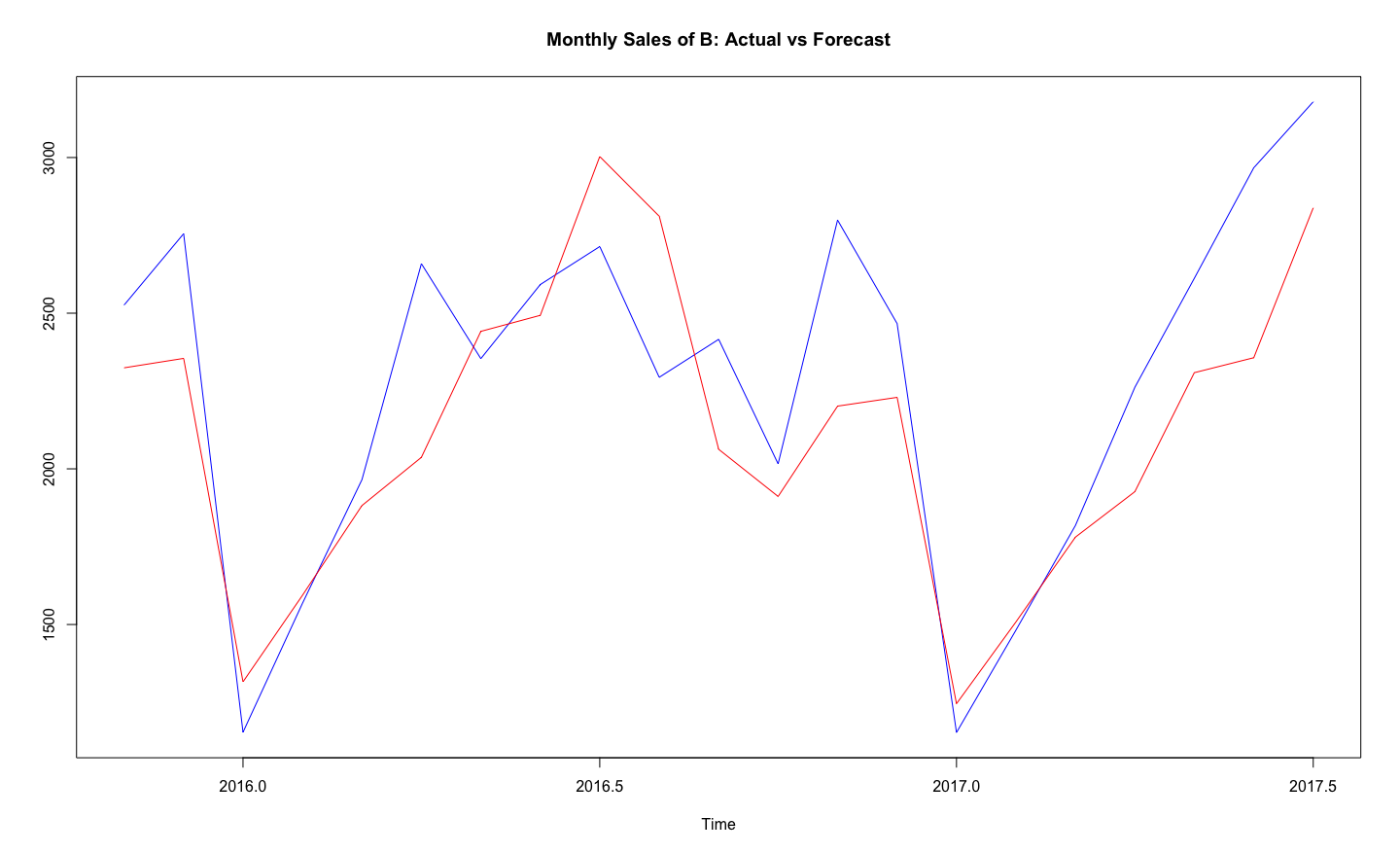
beta = 0.0014

gamma = 1e-04

**#Checking MAPE**

Vec5<- cbind(BTest,BTrain.fc$mean)

ts.plot(Vec5, col=c("blue", "red"), main="Monthly Sales of B: Actual vs Forecast")



MAPE <- mean(abs(Vec5[,1]-Vec5[,2])/Vec5[,1])

MAPE

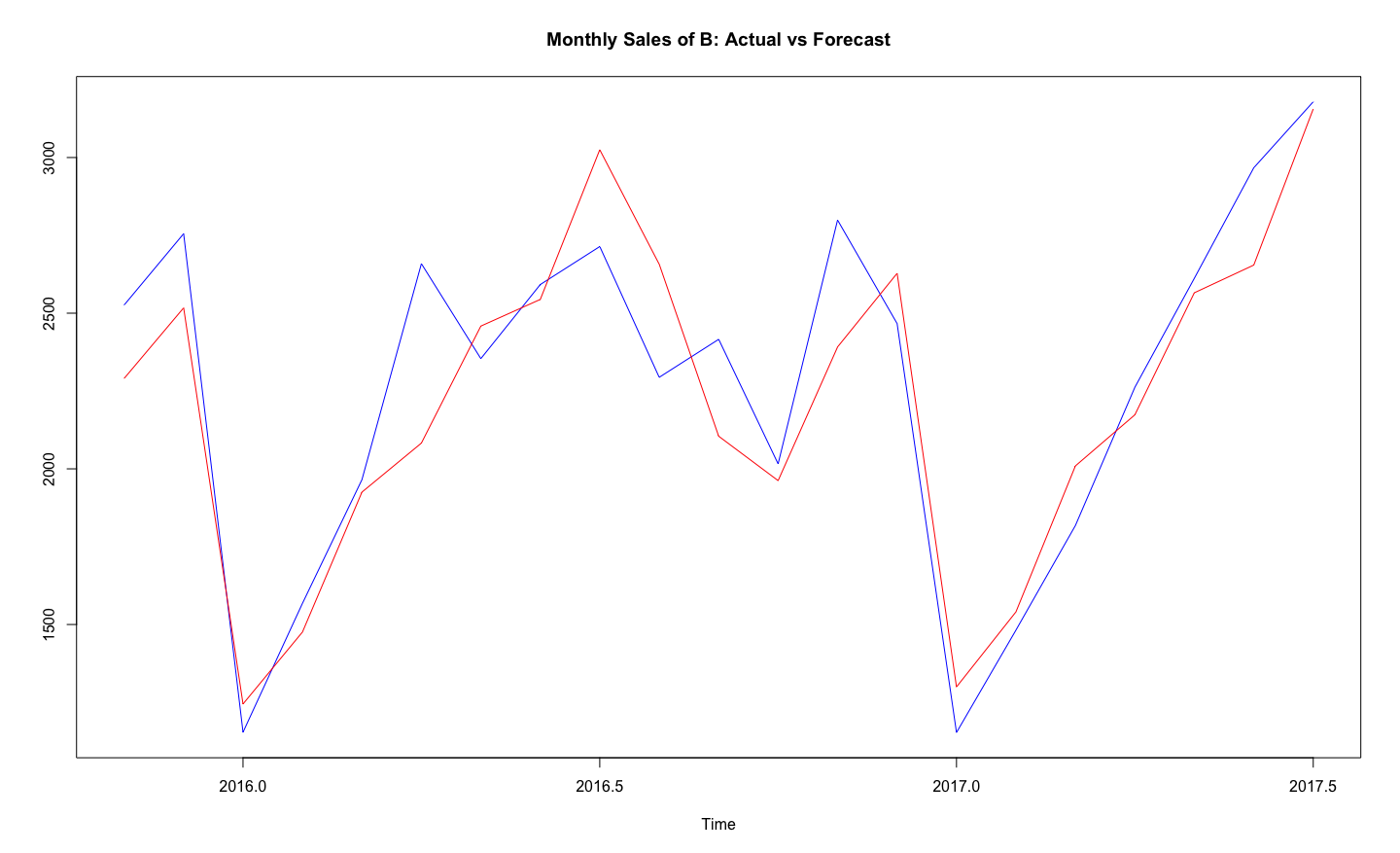
**10.8%**

**##Adjusting alpha, beta, gamma to get a better model**

BTrain.fc1 = hw(BTrain, seasonal = 'm',h=21,alpha=0.2,beta=0.12,gamma=0.18)

Vec<- cbind(BTest,BTrain.fc1$mean)

ts.plot(Vec, col=c("blue", "red"), main="Monthly Sales of B: Actual vs Forecast")



MAPE <- mean(abs(Vec[,1]-Vec[,2])/Vec[,1])

MAPE

8.03%

**CONCLUSION:**

FROM THE ABOVE MODELLING WE GET THE LESS MAPE VALUE OF 8.03% FOR MULTIPLICATIVE .THEREFORE I AM CONSIDERING MUTTIPLICATIVE MODEL FOR THE FORCASTING OF DEMAND B.

MAPE=8.03%

Smoothing parameters

alpha = 0.2 beta = 0.12 gamma = 0.18

SIGNIFICANCE OF THE FOLLOWING PARAMETER IN MODELLING:

Alpha= for the estimate of the level at the current time point .

Value of smoothing parameter for the level.

Beta= for the estimate of the slope b of the trend component at the current time point

Value of smoothing parameter for the trend.

Gamma= Value of smoothing parameter for the seasonal component.

Q 6 )

Use the ‘best’ model obtained from above to forecast demand for the period Oct 2017 to December 2018 for both items. Provide forecasted values as well as their upper and lower confidence limits. If you are the store manager what decisions would you make after looking at the demand of the two items over years?

# Now Forecasting with Original Data

**ITEM A**

demandA.fc = hw(demandA, seasonal = 'm',h=17,alpha=0.05,beta=0.03,gamma=0.4)

demandA.fc

> demandA.fc (FORCASTED VALUE)

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

Aug 2017 4539.850 3869.702 5209.999 3514.947 5564.754

Sep 2017 4385.999 3736.516 5035.483 3392.700 5379.298

Oct 2017 4754.753 4046.589 5462.918 3671.710 5837.797

Nov 2017 5620.941 4776.201 6465.682 4329.023 6912.860

Dec 2017 6652.673 5640.113 7665.233 5104.097 8201.250

Jan 2018 2847.535 2406.844 3288.227 2173.556 3521.515

Feb 2018 4411.108 3714.104 5108.113 3345.132 5477.085

Mar 2018 4686.030 3926.978 5445.083 3525.160 5846.901

Apr 2018 4614.035 3844.921 5383.149 3437.777 5790.293

May 2018 4364.920 3613.561 5116.279 3215.815 5514.024

Jun 2018 4582.749 3765.614 5399.884 3333.049 5832.449

Jul 2018 5042.606 4108.816 5976.396 3614.497 6470.714

Aug 2018 5217.317 4086.538 6348.097 3487.939 6946.695

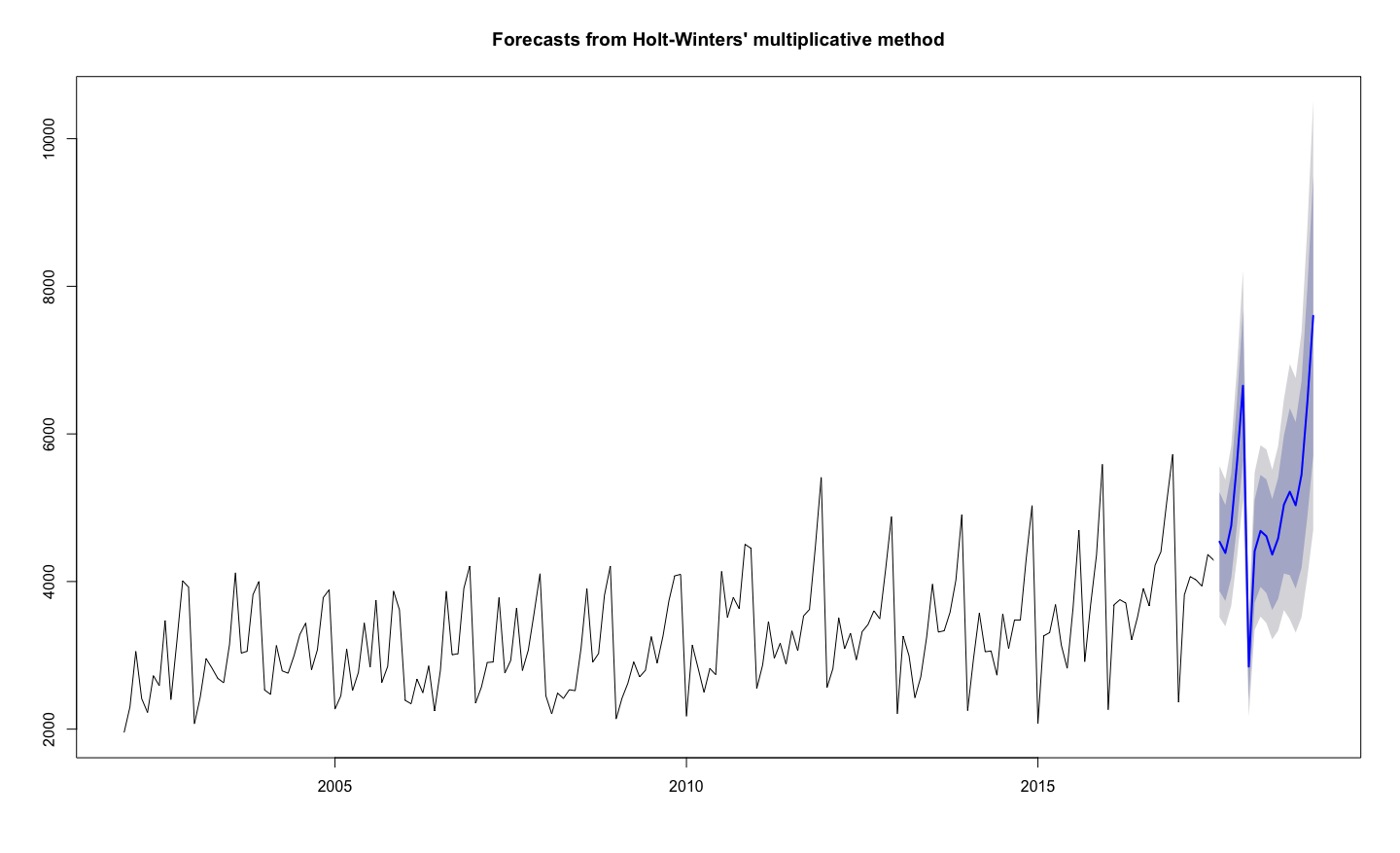
Sep 2018 5032.700 3905.050 6160.350 3308.108 6757.291

Oct 2018 5447.563 4183.753 6711.373 3514.732 7380.394

Nov 2018 6430.427 4883.912 7976.942 4065.236 8795.618

Dec 2018 7599.718 5703.235 9496.202 4699.297 10500.140

plot(demandA.fc)



**CONCLUSION: After looking at the forecasting of the item A over years I conclude that there is an increase in demand for Item A.I would like to keep more Item A in my store seeing the demand.**

**ITEM B:**

# Forecasting with original data

DemandB.fc = hw(demandB, seasonal = 'm',h=17,lpha=0.2,beta=0.12,gamma=0.18)

DemandB.fc

> DemandB.fc

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

Aug 2017 2916.733 2453.5734 3379.892 2208.3917 3625.074

Sep 2017 2552.590 2101.4956 3003.684 1862.7006 3242.479

Oct 2017 2335.788 1867.7699 2803.805 1620.0163 3051.559

Nov 2017 3010.332 2321.2045 3699.460 1956.4023 4064.262

Dec 2017 3153.349 2328.6626 3978.035 1892.1002 4414.597

Jan 2018 1517.197 1065.8906 1968.503 826.9835 2207.410

Feb 2018 1926.708 1279.0432 2574.373 936.1902 2917.226

Mar 2018 2521.243 1570.3158 3472.169 1066.9256 3975.560

Apr 2018 2956.970 1714.5072 4199.432 1056.7874 4857.152

May 2018 3307.055 1769.4294 4844.681 955.4594 5658.651

Jun 2018 3562.007 1740.7514 5383.262 776.6369 6347.377

Jul 2018 4094.058 1805.1260 6382.989 593.4388 7594.677

Aug 2018 3709.424 1411.1787 6007.669 194.5612 7224.286

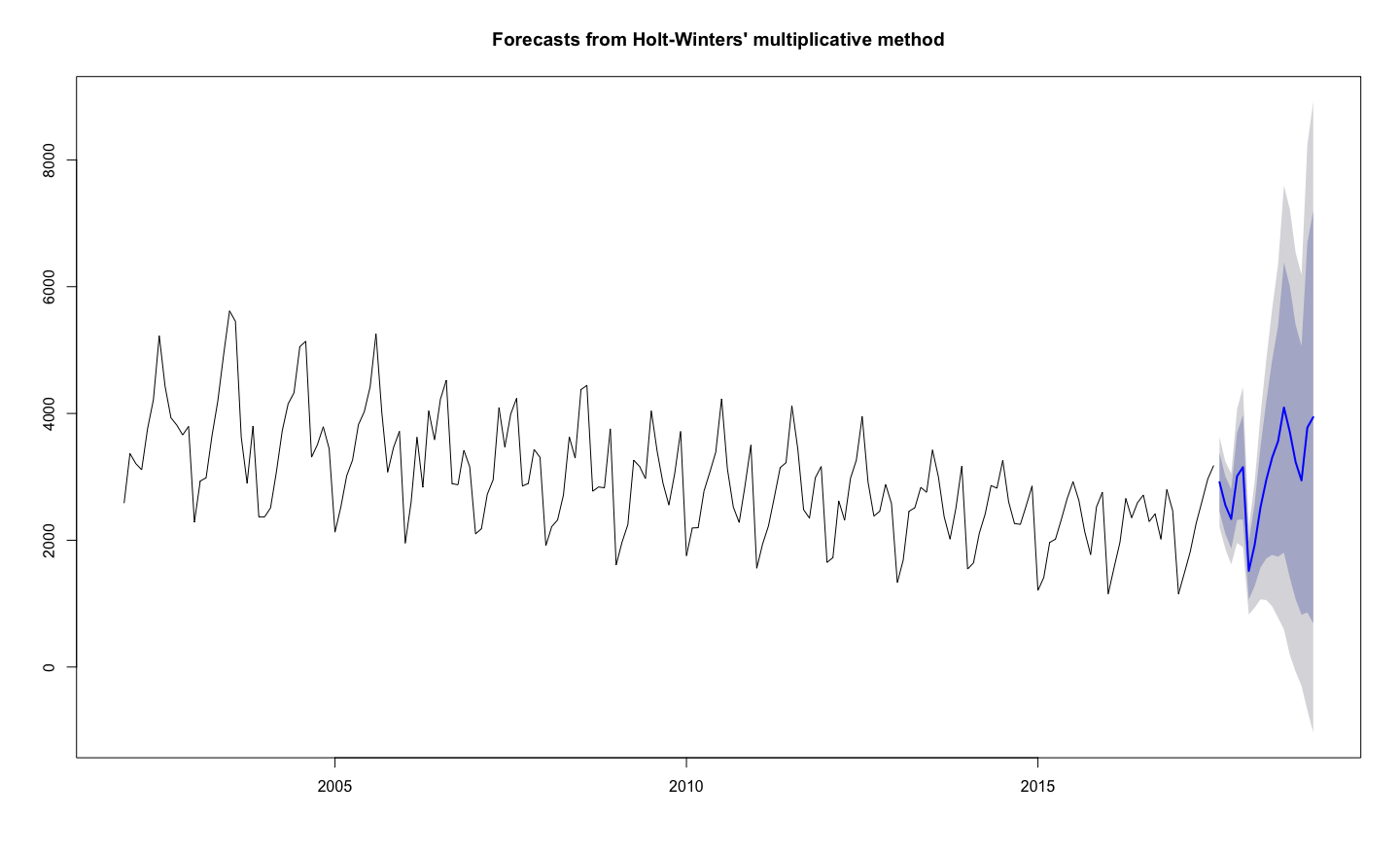
Sep 2018 3231.509 1069.5339 5393.484 -74.9467 6537.965

Oct 2018 2944.070 824.8211 5063.319 -297.0416 6185.182

Nov 2018 3778.256 861.6942 6694.818 -682.2404 8238.752

Dec 2018 3941.657 688.6791 7194.634 -1033.3433 8916.657

plot(DemandB.fc)



CONCLUSION: **After looking at the forecasting of the item B over the years. I conclude that there is decrease in demand for Item B . I would not like to keep Item B in my store seeing the demand.**

