Forecast the Plastic Sales data set. Prepare a document for each model explaining how many dummy variables you have created and RMSE value for each model. Finally which model you will use for Forecasting.

**Ans :**

**R Code :**

## Forecasting Data Driven

########## Plastic Sales Data Set #########

install.packages(c("forecast","fpp","smooth","tseries"))

library(forecast)

library(fpp)

library(smooth)

library(tseries)

library(readxl)

plasticsales<-read.csv('D:\\Data Science\\Excelr\\Assignments\\Assignment\\Forecasting\\PlasticSales.csv')

View(plasticsales)

# Converting data into time series object

sales<-ts(plasticsales$Sales,frequency = 12,start=c(49))

View(sales)

plot(sales)

# dividing entire data into training and testing data

train<-sales[1:48]

test<-sales[49:60] # Considering only 4 Quarters of data for testing because data itself is Quarterly

# seasonal data

# converting time series object

train<-ts(train,frequency = 12)

test<-ts(test,frequency = 12)

# Plotting time series data

plot(train) # Visualization shows that it has level, trend, seasonality => Additive seasonality

#### USING HoltWinters function ################

# Optimum values

# with alpha = 0.2 which is default value

# Assuming time series data has only level parameter

hw\_a<-HoltWinters(train,alpha = 0.2,beta = F,gamma = F)

hwa\_pred<-data.frame(predict(hw\_a,n.ahead=12))

# By looking at the plot the forecasted values are not showing any characters of train data

plot(forecast(hw\_a,h=12))

hwa\_mape<-MAPE(hwa\_pred$fit,test)\*100

# with alpha = 0.2, beta = 0.1

# Assuming time series data has level and trend parameter

hw\_ab<-HoltWinters(train,alpha = 0.2,beta = 0.1,gamma = F)

hwab\_pred<-data.frame(predict(hw\_ab,n.ahead = 12))

# by looking at the plot the forecasted values are still missing some characters exhibited by train data

plot(forecast(hw\_ab,h=12))

hwab\_mape<-MAPE(hwab\_pred$fit,test)\*100

# with alpha = 0.2, beta = 0.1, gamma = 0.1

# Assuming time series data has level,trend and seasonality

hw\_abg<-HoltWinters(train,alpha = 0.2,beta = 0.1,gamma = 0.1)

hwabg\_pred<-data.frame(predict(hw\_abg,n.ahead = 12))

# by looking at the plot the characters of forecasted values are closely following historical data

plot(forecast(hw\_abg,h=12))

hwabg\_mape<-MAPE(hwabg\_pred$fit,test)\*100

# With out optimum values

hw\_na<-HoltWinters(train,beta = F,gamma = F)

hwna\_pred<-data.frame(predict(hw\_na,n.ahead = 12))

hwna\_pred

plot(forecast(hw\_na,h=12))

hwna\_mape<-MAPE(hwna\_pred$fit,test)\*100

hw\_nab<-HoltWinters(train,gamma=F)

hwnab\_pred<-data.frame(predict(hw\_nab,n.ahead=12))

hwnab\_pred

plot(forecast(hw\_nab,h=12))

hwnab\_mape<-MAPE(hwnab\_pred$fit,test)\*100

hw\_nabg<-HoltWinters(train)

hwnabg\_pred<-data.frame(predict(hw\_nabg,n.ahead =12))

hwnabg\_pred

plot(forecast(hw\_nabg,h=12))

hwnabg\_mape<-MAPE(hwnabg\_pred$fit,test)\*100

df\_mape<-data.frame(c("hwa\_mape","hwab\_mape","hwabg\_mape","hwna\_mape","hwnab\_mape","hwnabg\_mape"),c(hwa\_mape,hwab\_mape,hwabg\_mape,hwna\_mape,hwnab\_mape,hwnabg\_mape))

colnames(df\_mape)<-c("MAPE","VALUES")

View(df\_mape)

# Based on the MAPE value who choose holts winter exponential tecnique which assumes the time series

# Data level, trend, seasonality characters with default values of alpha, beta and gamma

new\_model <- HoltWinters(sales)

plot(forecast(new\_model,n.ahead=24))

# Forecasted values for the next 4 quarters

forecast\_new <- data.frame(predict(new\_model,n.ahead=12))

forecast\_new

######## ARIMA Model #############

# Converting data into time series object

sales<-ts(plasticsales$Sales,frequency = 12,start=c(49))

View(sales)

plot(sales)

# dividing entire data into training and testing data

train<-sales[1:48]

test<-sales[49:60] # Considering only 4 Quarters of data for testing because data itself is Quarterly

# seasonal data

# converting time series object

train<-ts(train,frequency = 12)

test<-ts(test,frequency = 12)

plot(train)

acf(train)

pacf(train)

# Auto.Arima model on the price agg data

library(forecast)

model\_AA <- auto.arima(train)

model\_AA

pred\_AA <- data.frame(forecast(model\_AA))

acf(model\_AA$residuals)

pacf(model\_AA$residuals)

windows()

plot(forecast(model\_AA,h=36),xaxt="n")

**Results :**

> # With out optimum values

> hw\_na<-HoltWinters(train,beta = F,gamma = F)

> hwna\_pred<-data.frame(predict(hw\_na,n.ahead = 12))

> hwna\_pred

fit

1 1209.013

2 1209.013

3 1209.013

4 1209.013

5 1209.013

6 1209.013

7 1209.013

8 1209.013

9 1209.013

10 1209.013

11 1209.013

12 1209.013

> plot(forecast(hw\_na,h=12))

> hwna\_mape<-MAPE(hwna\_pred$fit,test)\*100

> hw\_nab<-HoltWinters(train,gamma=F)

> hwnab\_pred<-data.frame(predict(hw\_nab,n.ahead=12))

> hwnab\_pred

fit

1 1206.496

2 1203.993

3 1201.489

4 1198.986

5 1196.482

6 1193.979

7 1191.475

8 1188.972

9 1186.468

10 1183.965

11 1181.461

12 1178.958

> plot(forecast(hw\_nab,h=12))

> hwnab\_mape<-MAPE(hwnab\_pred$fit,test)\*100

> hw\_nabg<-HoltWinters(train)

> hwnabg\_pred<-data.frame(predict(hw\_nabg,n.ahead =12))

> hwnabg\_pred

fit

1 1174.434

2 1112.193

3 1196.874

4 1355.929

5 1513.492

6 1640.174

7 1667.731

8 1735.475

9 1739.271

10 1677.499

11 1467.639

12 1302.111

> plot(forecast(hw\_nabg,h=12))

> hwnabg\_mape<-MAPE(hwnabg\_pred$fit,test)\*100

> df\_mape<-data.frame(c("hwa\_mape","hwab\_mape","hwabg\_mape","hwna\_mape","hwnab\_mape","hwnabg\_mape"),c(hwa\_mape,hwab\_mape,hwabg\_mape,hwna\_mape,hwnab\_mape,hwnabg\_mape))

> colnames(df\_mape)<-c("MAPE","VALUES")

> View(df\_mape)

| **MAPE** | | **VALUES** | |
| --- | --- | --- | --- |
|  |  | |  |
| **1** | hwa\_mape | | 16.233286 |
| **2** | hwab\_mape | | 16.496414 |
| **3** | hwabg\_mape | | 10.136251 |
| **4** | hwna\_mape | | 19.430276 |
| **5** | hwnab\_mape | | 20.058912 |
| **6** | hwnabg\_mape | | 9.838578 |

> new\_model <- HoltWinters(sales)

> plot(forecast(new\_model,n.ahead=24))

> # Forecasted values for the next 4 quarters

> forecast\_new <- data.frame(predict(new\_model,n.ahead=12))

> forecast\_new

fit

1 948.3095

2 914.0795

3 993.7539

4 1149.1910

5 1310.3413

6 1439.5473

7 1453.8645

8 1515.9441

9 1515.4919

10 1456.7702

11 1248.4452

12 1104.7552

> # Converting data into time series object

> sales<-ts(plasticsales$Sales,frequency = 12,start=c(49))

> View(sales)

> plot(sales)

> # dividing entire data into training and testing data

> train<-sales[1:48]

> test<-sales[49:60] # Considering only 4 Quarters of data for testing because data itself is Quarterly

> # seasonal data

> # converting time series object

> train<-ts(train,frequency = 12)

> test<-ts(test,frequency = 12)

> plot(train)

> acf(train)

> pacf(train)

> # Auto.Arima model on the price agg data

> library(forecast)

> model\_AA <- auto.arima(train)

> model\_AA

Series: train

ARIMA(1,0,0)(0,1,1)[12] with drift

Coefficients:

ar1 sma1 drift

0.7769 -0.5223 8.6311

s.e. 0.1237 0.3394 1.4211

sigma^2 estimated as 1410: log likelihood=-182.35

AIC=372.7 AICc=373.99 BIC=379.03

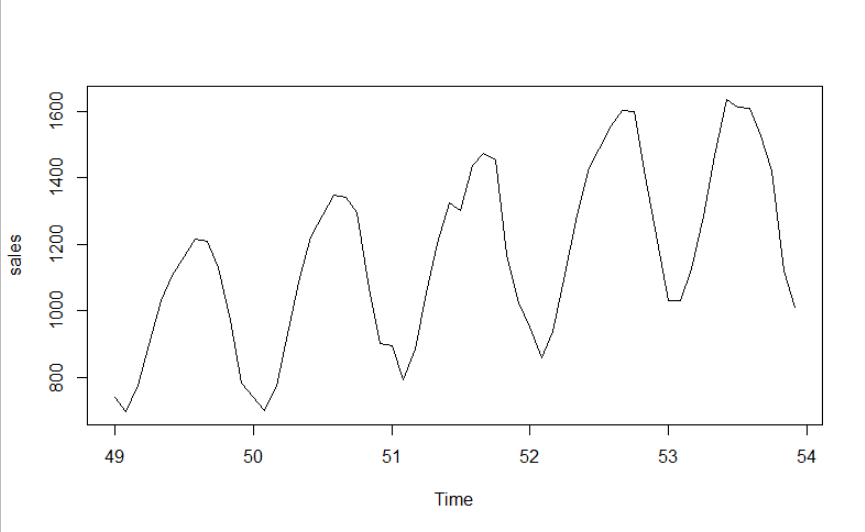
> pred\_AA <- data.frame(forecast(model\_AA))

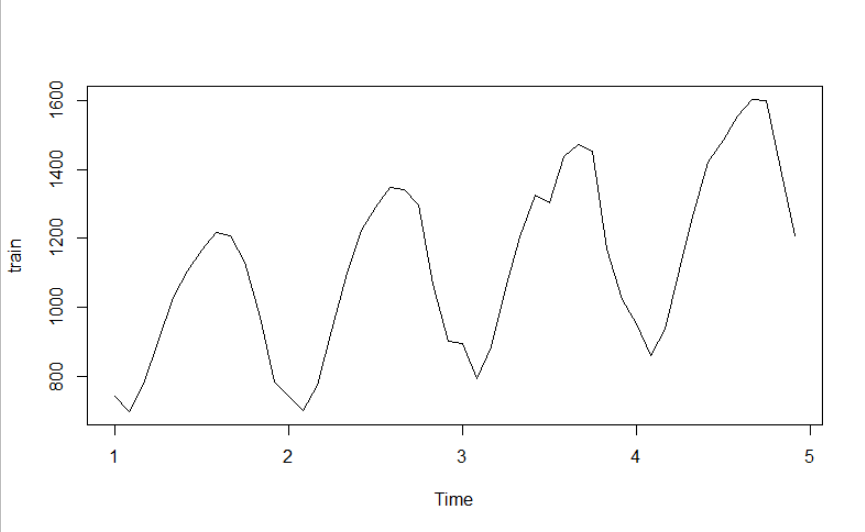
> acf(model\_AA$residuals)

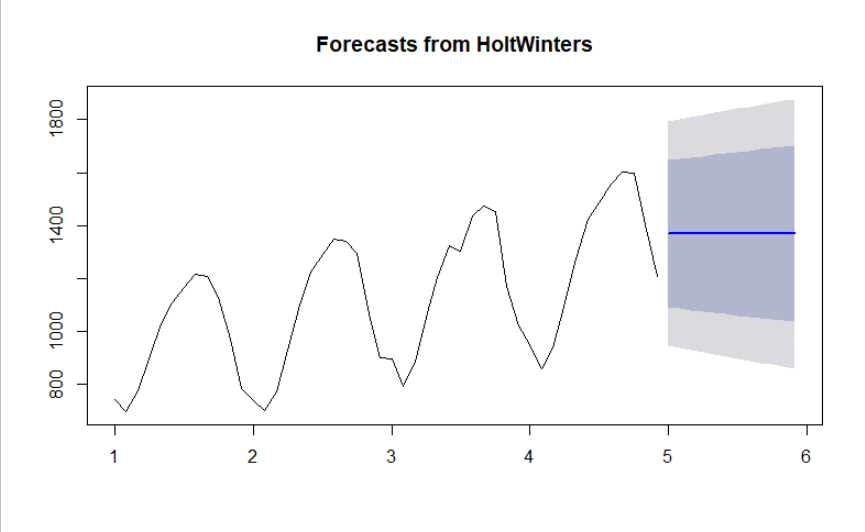
> pacf(model\_AA$residuals)

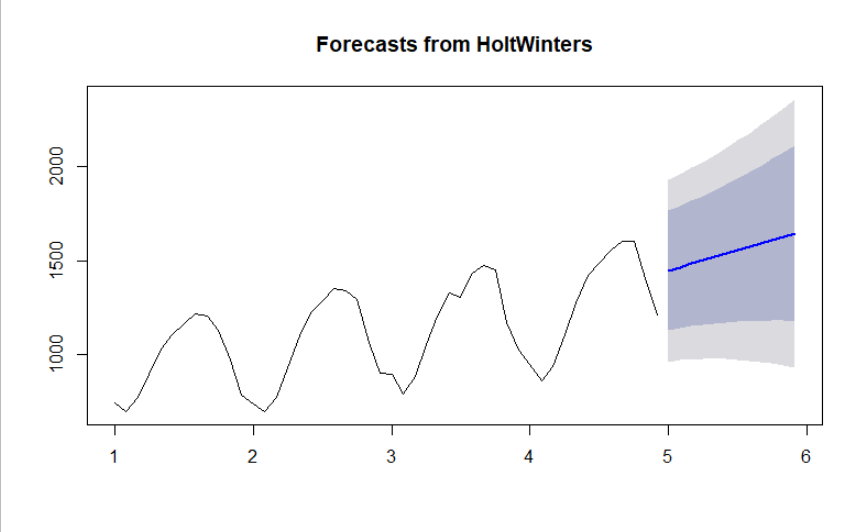
> plot(forecast(model\_AA,h=36),xaxt="n")

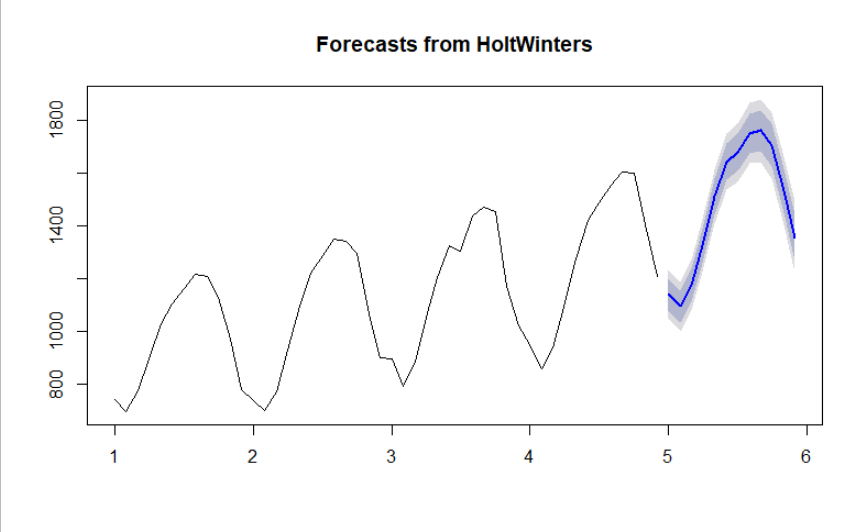
**Plots :**

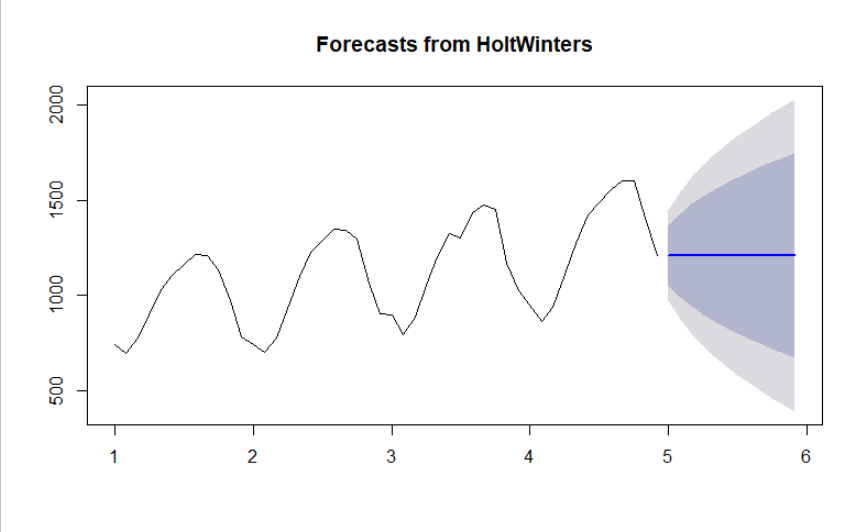


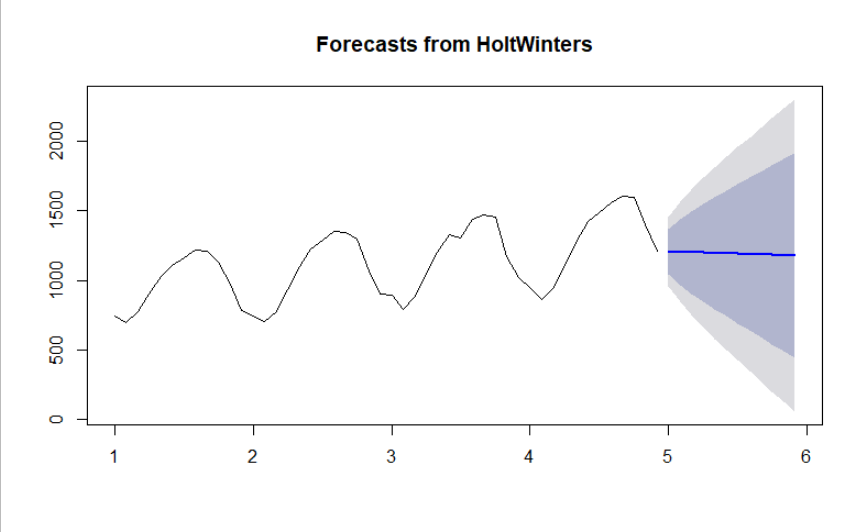


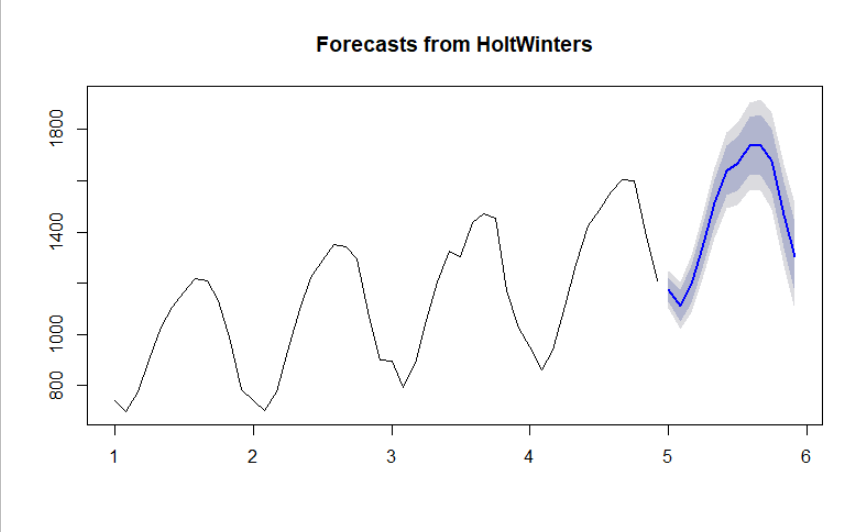


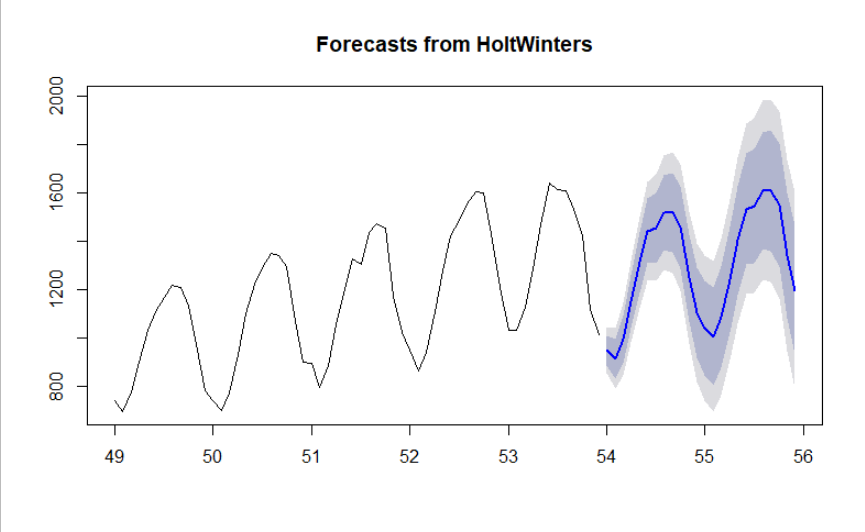


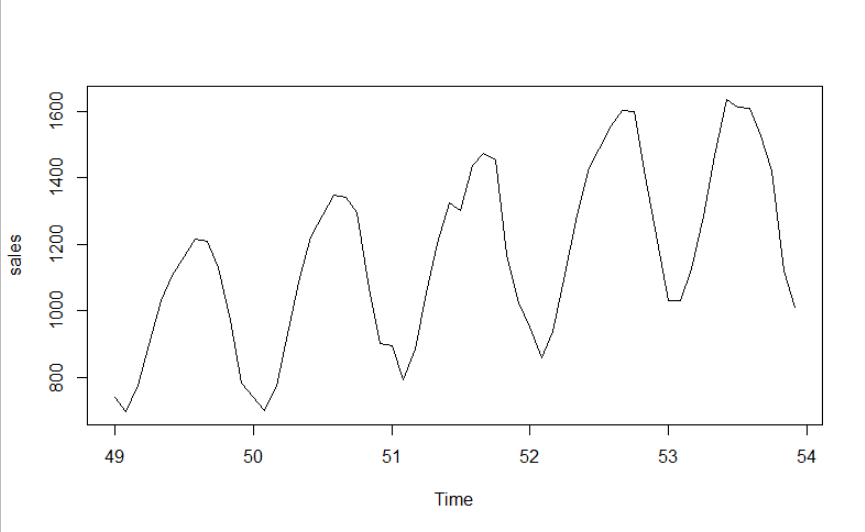


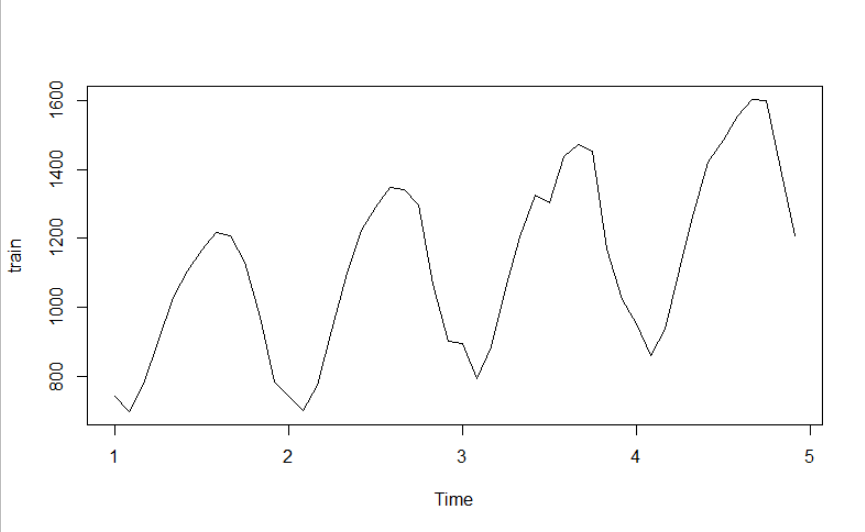


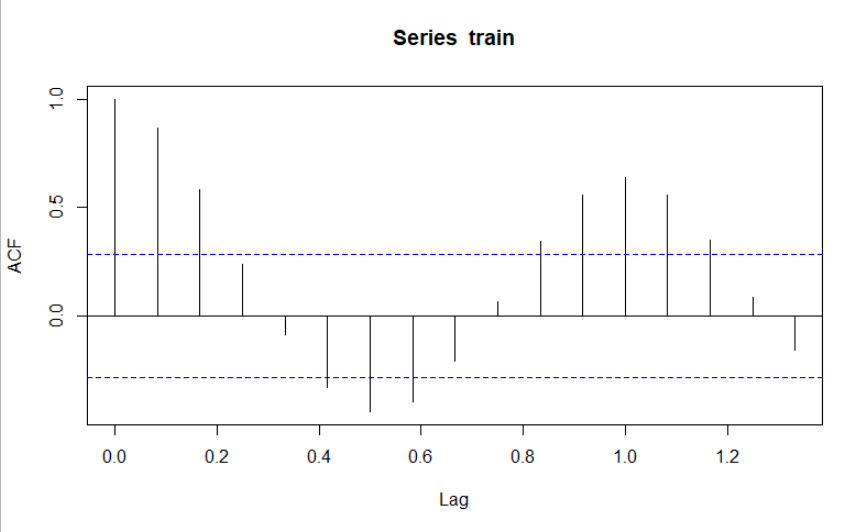


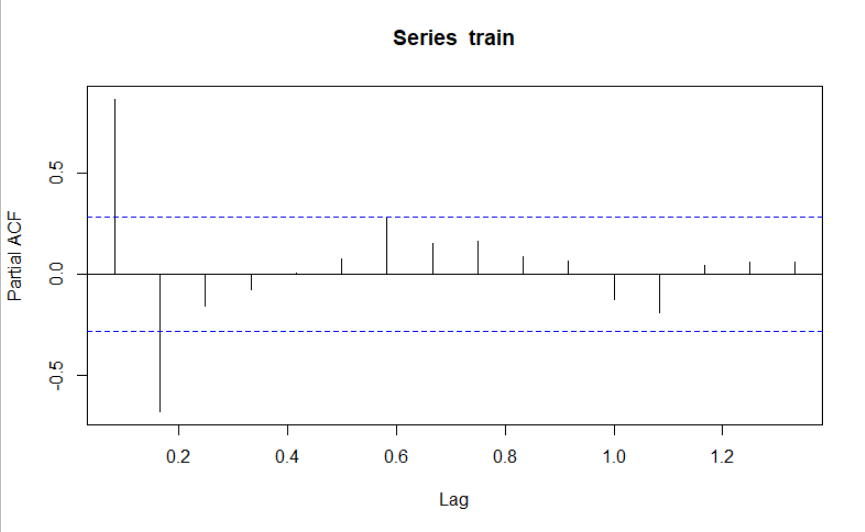


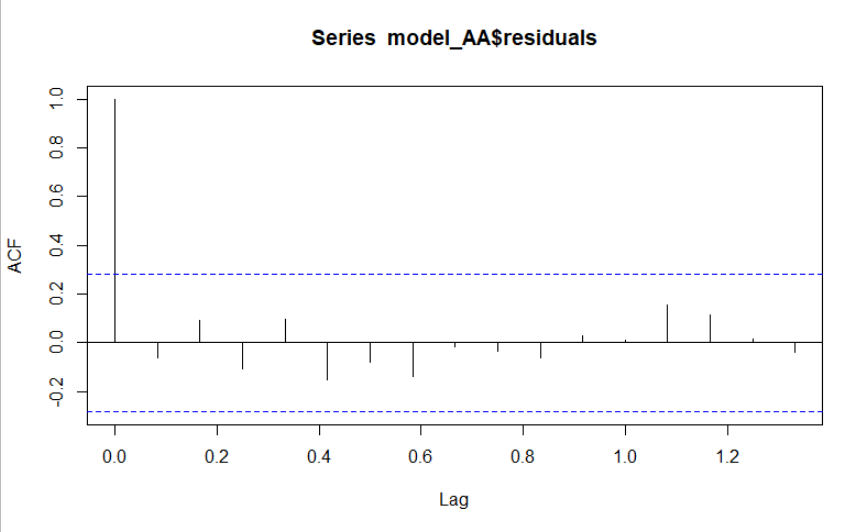


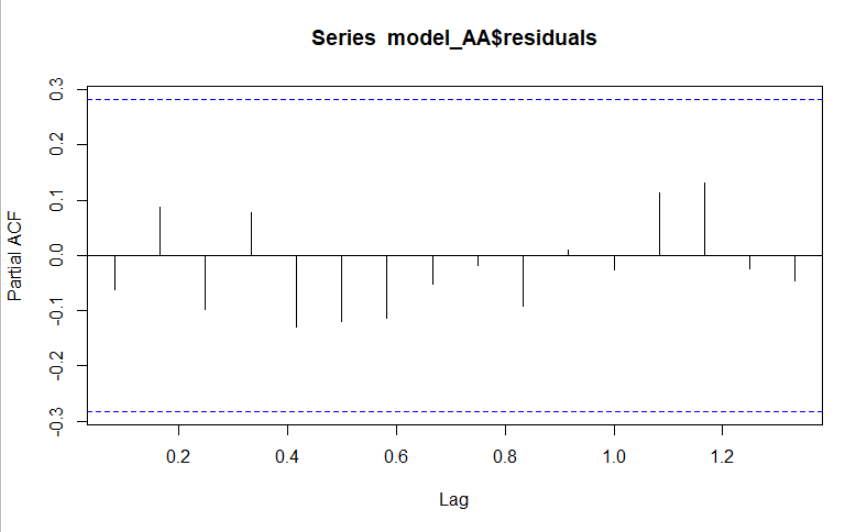


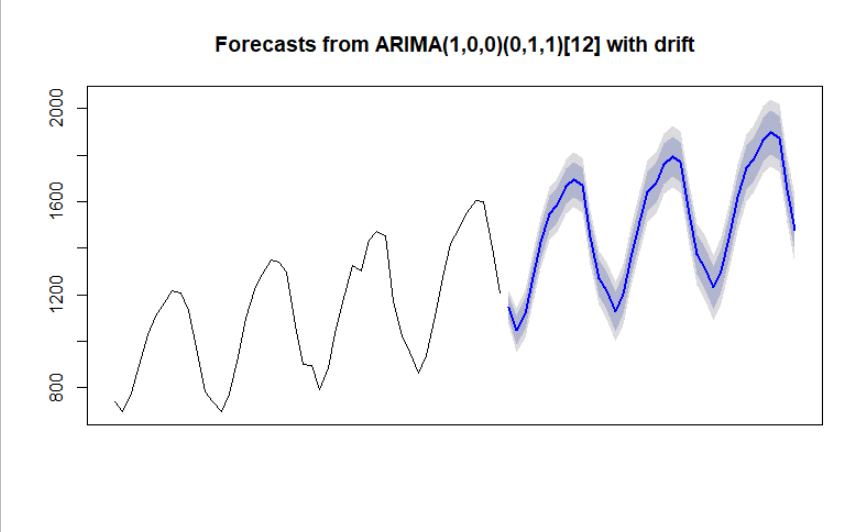












**Inference :**

The MAPE values of various models are as below :

| **MAPE** | | **VALUES** | |
| --- | --- | --- | --- |
|  |  | |  |
| **1** | hwa\_mape | | 16.233286 |
| **2** | hwab\_mape | | 16.496414 |
| **3** | hwabg\_mape | | 10.136251 |
| **4** | hwna\_mape | | 19.430276 |
| **5** | hwnab\_mape | | 20.058912 |
| **6** | hwnabg\_mape | | 9.838578 |