Forecast the Airlines Passengers 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

########## Airlines Data Set #########

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

library(forecast)

library(fpp)

library(smooth)

library(tseries)

library(readxl)

airlines<-read\_xlsx("D:\\Data Science\\Excelr\\Assignments\\Assignment\\Forecasting\\Airlines+Data.xlsx") # Aviation.csv

View(airlines)

# Converting data into time series object

air<-ts(airlines$Passengers,frequency = 12,start=c(95))

View(air)

plot(air)

# dividing entire data into training and testing data

train<-air[1:84]

test<-air[85:96] # 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(air)

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

air<-ts(airlines$Passengers,frequency = 12,start=c(95))

View(air)

plot(air)

# dividing entire data into training and testing data

train<-air[1:84]

test<-air[85:96] # 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 :**

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

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

> hwnab\_pred

fit

1 281.6136

2 285.2273

3 288.8409

4 292.4546

5 296.0682

6 299.6819

7 303.2955

8 306.9091

9 310.5228

10 314.1364

11 317.7501

12 321.3637

> 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

1 284.4054

2 274.4373

3 311.4902

4 314.4740

5 317.5493

6 360.0018

7 402.8916

8 382.4633

9 344.6450

10 306.3237

11 271.6805

12 311.1774

> 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 | | 17.236298 |
| **2** | hwab\_mape | | 11.554964 |
| **3** | hwabg\_mape | | 6.552500 |
| **4** | hwna\_mape | | 18.555987 |
| **5** | hwnab\_mape | | 13.104290 |
| **6** | hwnabg\_mape | | 1.730844 |

> new\_model <- HoltWinters(air)

> 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

1 321.4296

2 313.9403

3 352.8969

4 349.6872

5 354.9324

6 408.0821

7 445.7063

8 433.7851

9 383.4245

10 336.6096

11 303.4973

12 341.0933

> # Auto.Arima model on the price agg data

> library(forecast)

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

> model\_AA

Series: train

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

Coefficients:

ma1 sar1

-0.2591 -0.2603

s.e. 0.1296 0.1179

sigma^2 estimated as 96.77: log likelihood=-262.5

AIC=531 AICc=531.36 BIC=537.79

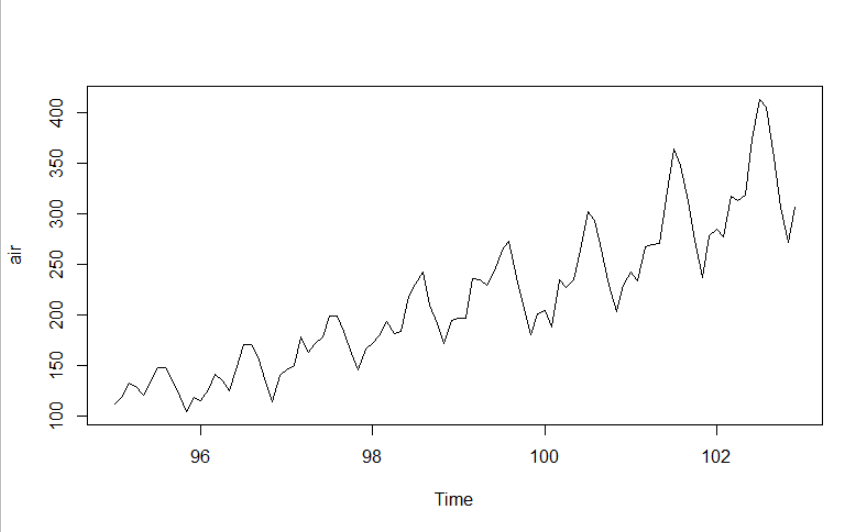
> pred\_AA <- data.frame(forecast(model\_AA))

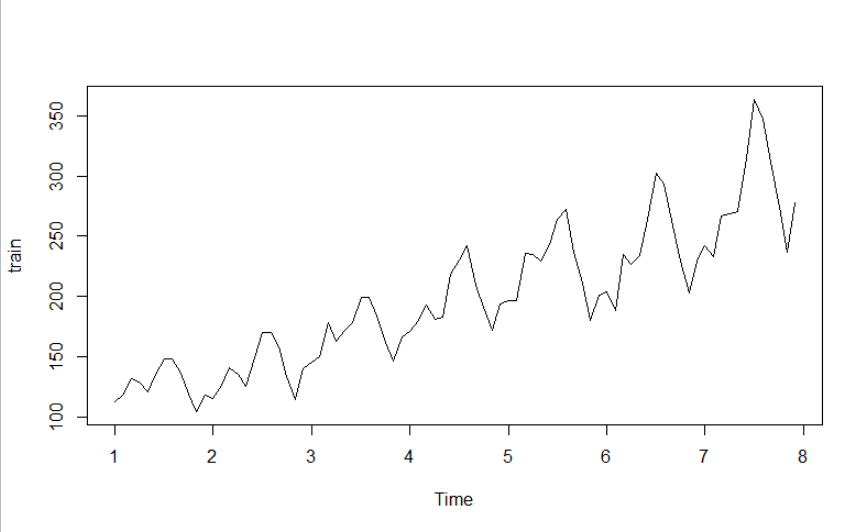
> acf(model\_AA$residuals)

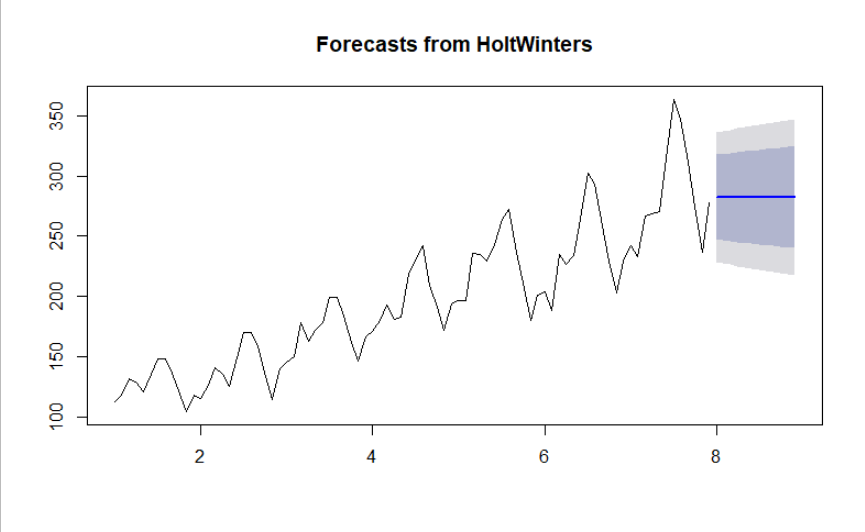
> pacf(model\_AA$residuals)

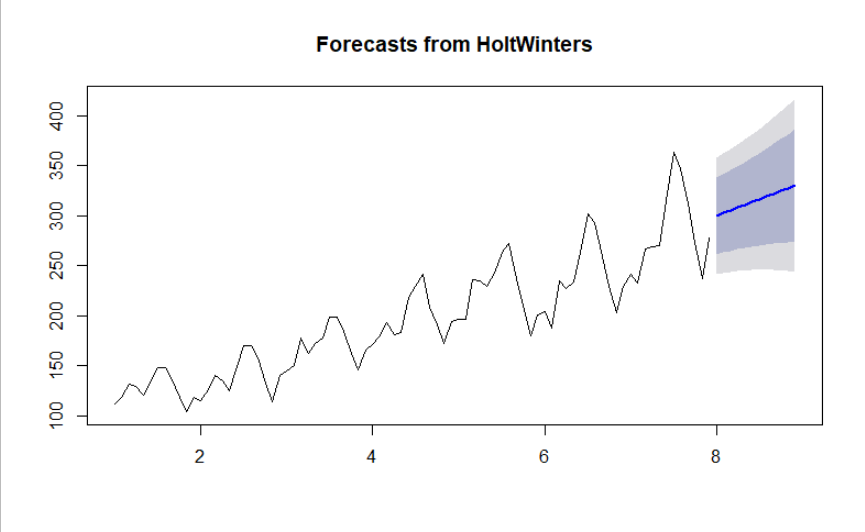
> plot(forecast(model\_AA,h=36),xaxt="n")

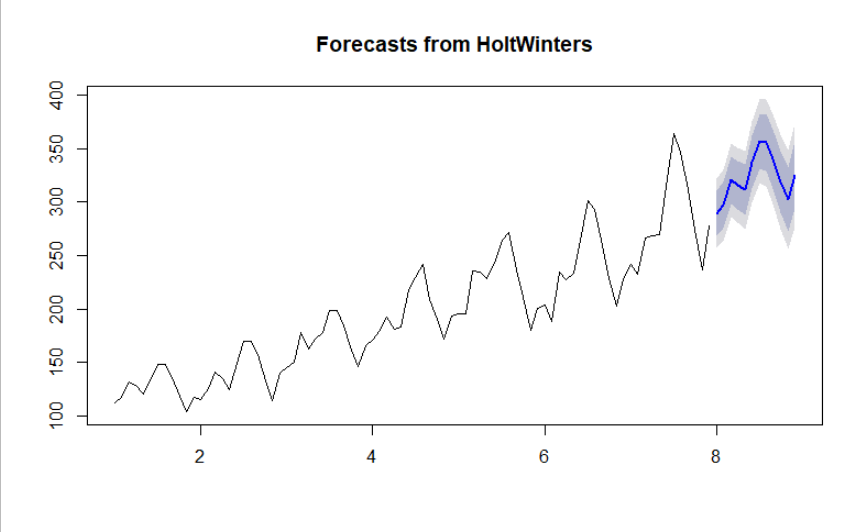
**Plots :**

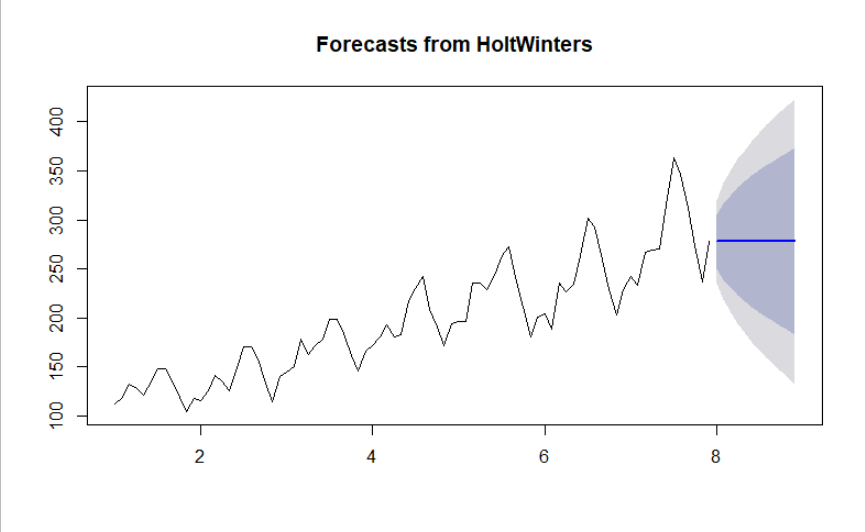


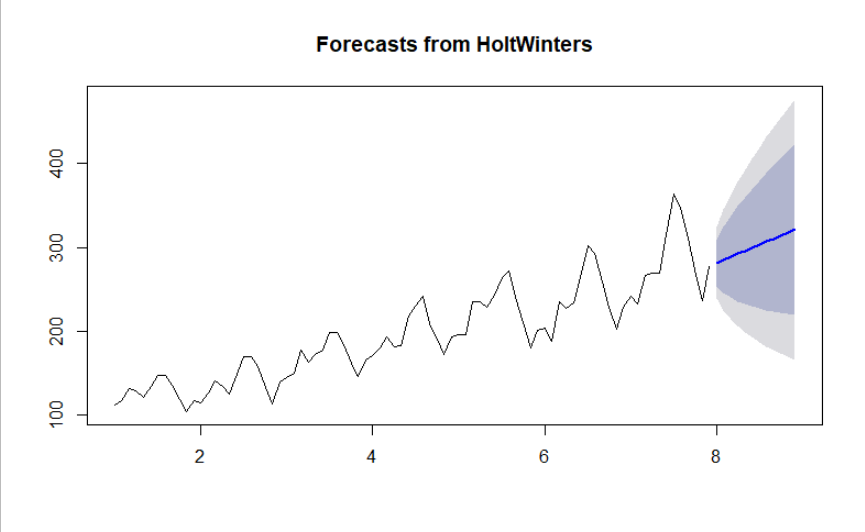


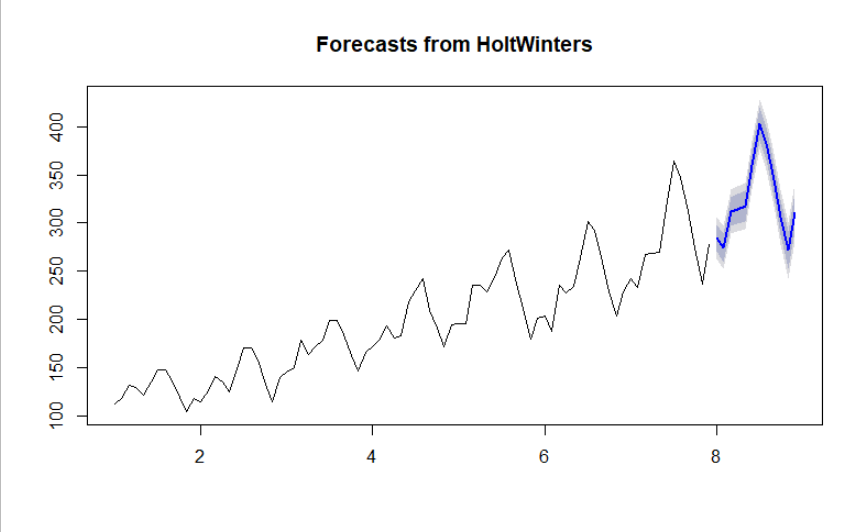


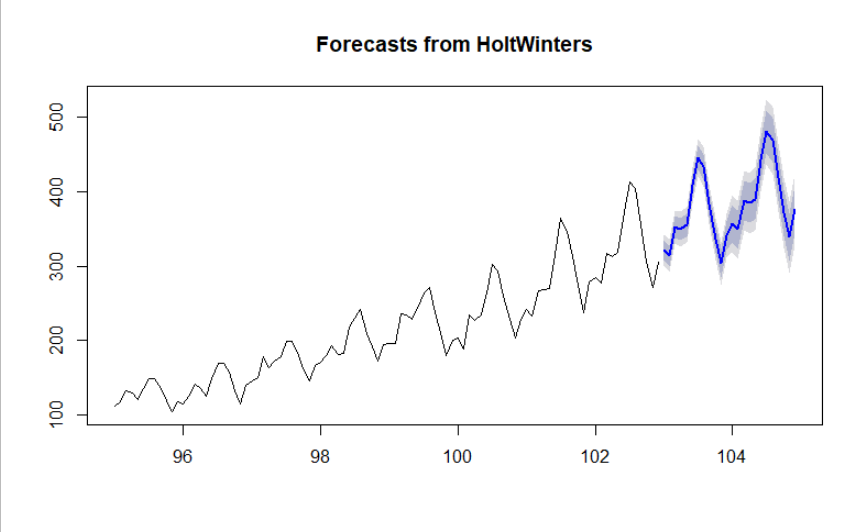


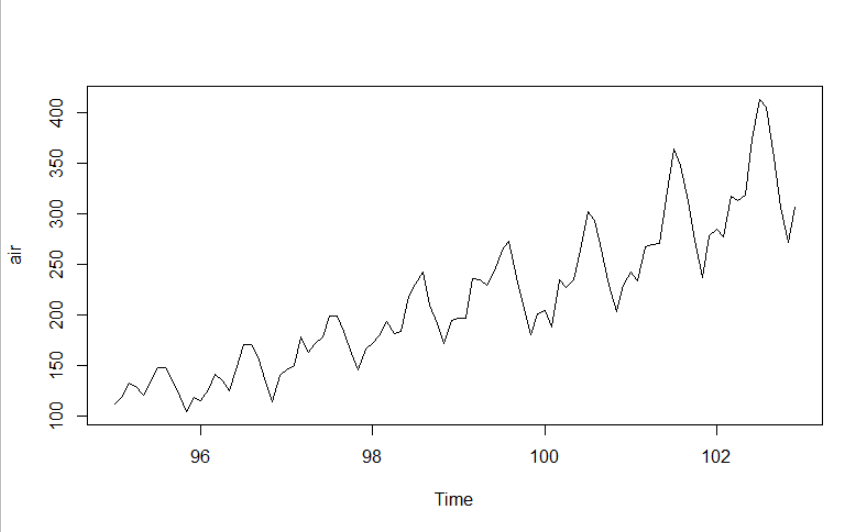


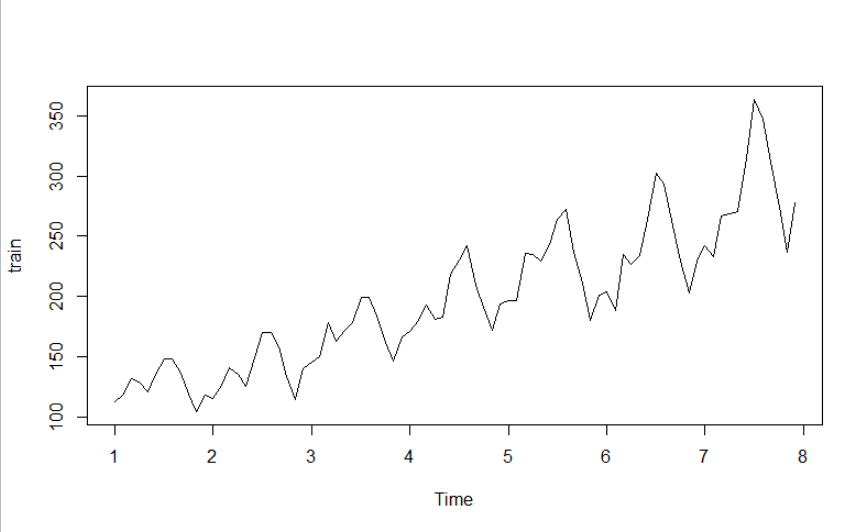


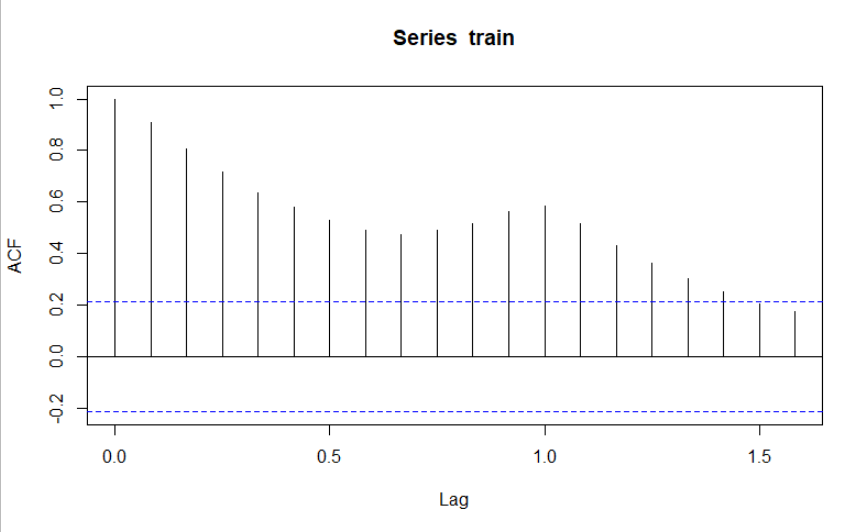


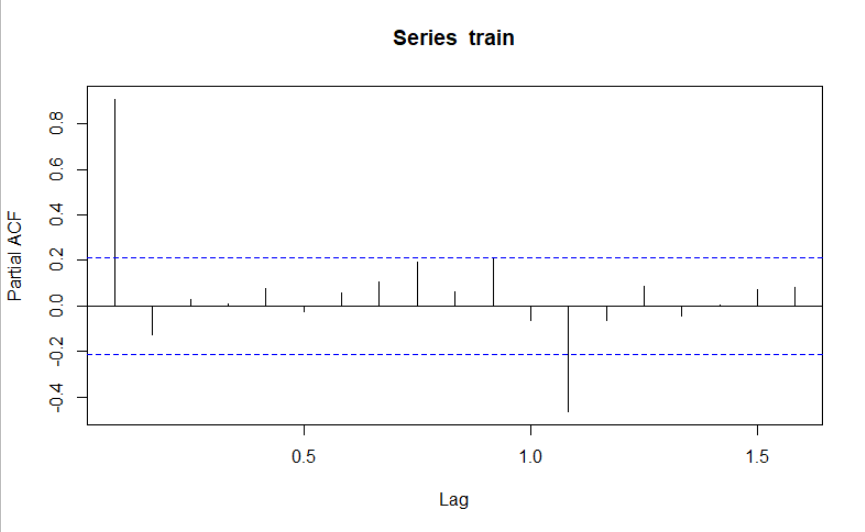


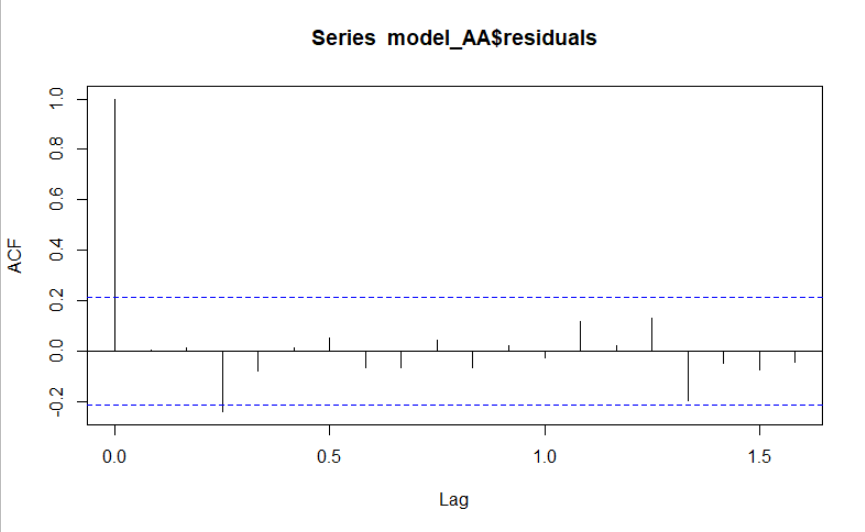


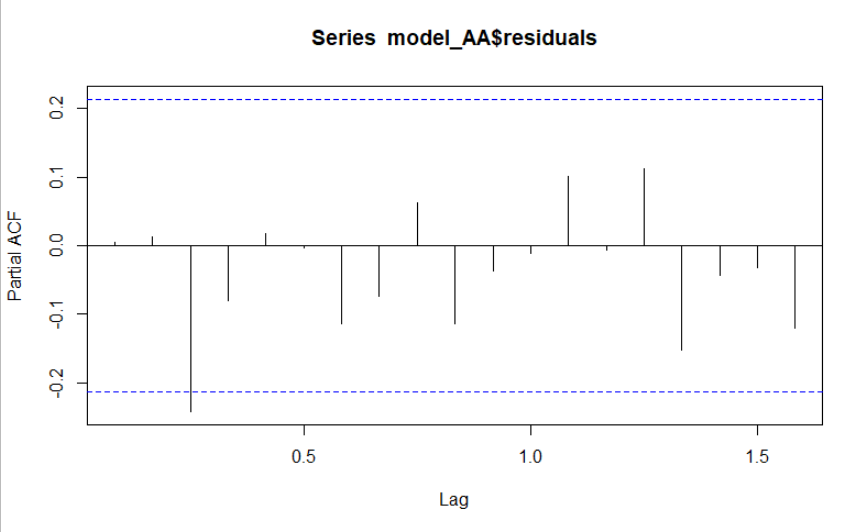


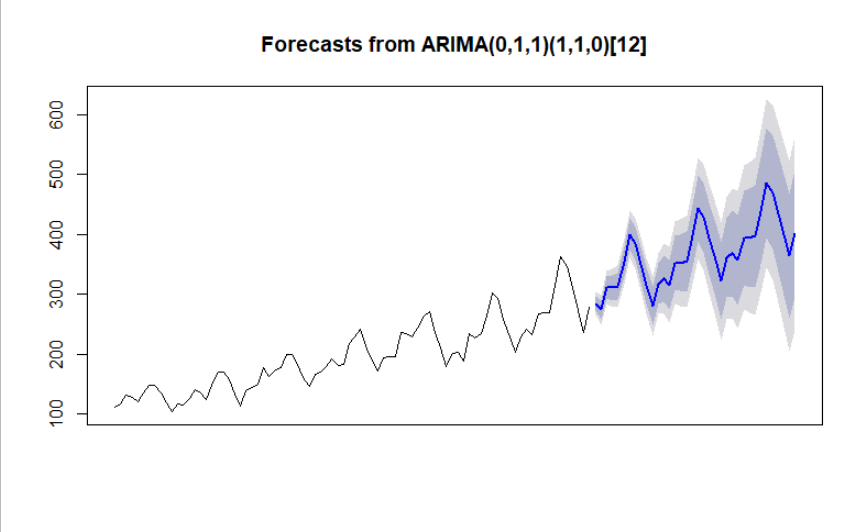












**Inference :**

The MAPE values of various models are as below :

| **MAPE** | | **VALUES** | |
| --- | --- | --- | --- |
|  |  | |  |
| **1** | hwa\_mape | | 17.236298 |
| **2** | hwab\_mape | | 11.554964 |
| **3** | hwabg\_mape | | 6.552500 |
| **4** | hwna\_mape | | 18.555987 |
| **5** | hwnab\_mape | | 13.104290 |
| **6** | hwnabg\_mape | | 1.730844 |