## Forecast the Retail Food & Beverages Sales: Seasonal ARIMA model in R

#Read the data from the working directory, create your own working directly to read the dataset.

data1 <- read.csv ("C:/Users/Desktop/data/

final\_sales.csv",header=TRUE,sep=",")

#Convert data into time series data

data <- ts(data1[,2],start = c(1992,1),frequency = 12)

data

#Perform exploratory data analysis to know about the time series data

#Displays the start date of the time series data

start(data)

[1] 1992 1

# displays the end date of the time series data

end(data)

[1] 2017 9

#Displays the frequency of the time series data whether monthly, quaterly, weekly.

frequency(data)

[1] 12

#Displays the data type it is time series data

class(data)

[1] "ts"

#Displays descriptive statistics of the time series data

Summary(data)

Min 1st Qu Median Mean 3rd Qu Max.

28235 35168 42036 43133 50171 64663

|  |  |
| --- | --- |
| |  | | --- | |  | |

#Checking the missing values present in the time series data

is.na(data)

[1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[11] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

[21] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

#Plotting the time series data

plot(data, xlab='Years', ylab = 'Sales')

#Install astsa package

install.packages("astsa")

library(astsa)

#To see acf and pacf in original data

acf2(data,max.lag = 24)

#Seasonally differenced retail sales

datadiff12 <- diff(data,12)

#Plot seasonally differenced retail sales

plot.ts(datadiff12)

#Trend and seasonally differenced retail sales

diff1and12=diff(datadiff12,1)

#Plot Trend and seasonally differenced retail sales

plot(diff1and12)

#To see acf and pacf of trend and seasonally differenced retail sales

acf2(diff1and12,max.lag = 36)

#Install forecast package

install.packages("forecast")

library(forecast)

#Building seasonal  model

model1<- arima(data,order=c(2,1,1),seasonal=list

(order=c(2,1,2),period=12))

summary(model1)

# Portmanteau or Box-Ljung test to check whether residuals are white noise

Acf(residuals(model1))

Box.test(residuals(model1)lag=24,fitdf =1,type="Ljung")

#Rebuilding model with different non seasonal terms

model2 <- arima(data,order=c(6,1,1),seasonal= list

(order=c(2,1,2),period=12))

summary(model2)

# Portmanteau and Box-Ljung test on model2 to check whether residual are white noise

Acf(residuals(model2))

Box.test(residuals(model2),lag=24,fitdf = 1,type="Ljung")

#Forecast for the next 30 month

Pred <- forecast(model2,h=30)

Pred

#Creating the plot for forecast retail sales

plot(Pred,ylab="sales (million in dollars)",xlab="Year")