**IDS – 476: Business Forecasting Using Time-series**

**Final Project Report**

**Group Members:**

**Adusumelli, Jaideep – 672039068  
Kondathasula, Uday – 656467843  
Chowdhury, Gaurav – 651395167  
Kompella, Kaushik – 663144176  
Sirothia, Shubham – 650889404  
Thomas, Shobin – 657777256**

**Objective**

To analyze the monthly airline passenger revenue miles\* to isolate the seasonality from the trend and analyze the trend of the airline revenue, so that we could forecast the airline revenue for the next 2 years.

\*(The passenger revenue miles is a measure used in the airline industry which depicts the total number of revenue paying passengers multiplied by the total number of miles travelled by each one of them.)

**Data**

For this project, we used the monthly passenger revenue miles’ data of the spirit airlines from January of 2003 to the December of 2016. We obtained this data from the Bureau of Transportation statistics, United States of America.

**Software or Programming language used: R**

**Preliminary analysis**

The data that we downloaded consisted of two variables a) Monthly Revenue, and b) Date. First we converted all the dates into the format of Year and month.

Then we plotted the below graph to check the trend displayed by the actual data.

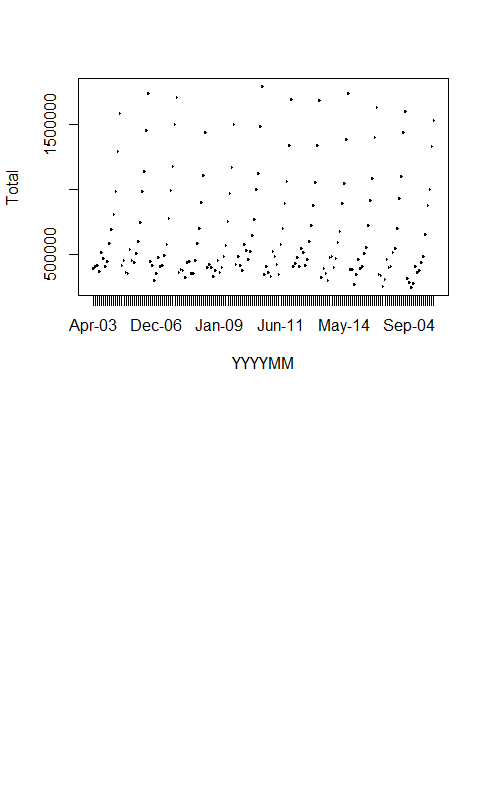


Fig 1. Plot of actual data

Then we converted the data into time series using the code:

> rpm = read.csv("data.csv")

> rpm.ts = ts(as.numeric(rpm$Total), start=c(2003,01),freq=12)

> plot(rpm.ts,ylab="Total")

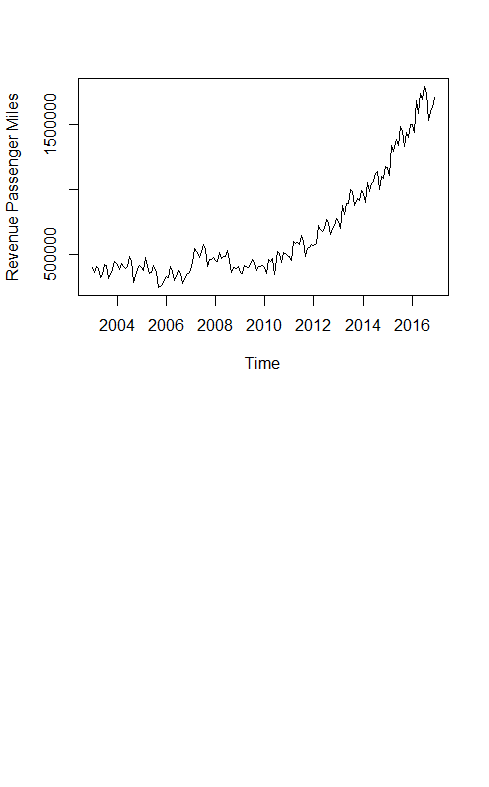


Fig 2. Time series plot of data

**Decomposition of data**

We then decomposed the data into 3 parts a) Trend, b) Seasonality, and c) random.

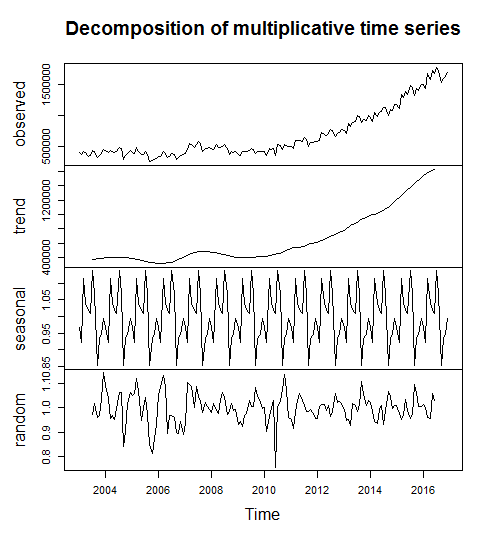


Fig 3. Decomposed graphs of time series data

As we can see in figure two that our data had an upward increasing trend, hence we had to use the multiplicative decomposition.

Xt = Mt \* St \* Ztj

Mt: Trend in the data

St: Seasonality in the data

Zt: Randomness in the data

Then we plotted the auto decomposition factor graph for the decomposed data.

> rpm.acf=acf(as.numeric(Rpm.decom$random), na.action = na.omit, lag.max = 40)

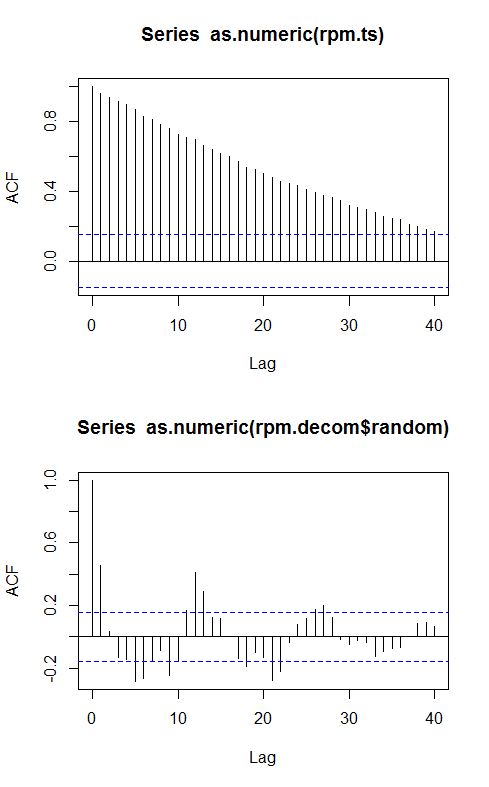


Fig 4. ACF plot for the random variable in decomposed data.

As we can see that the ACF plot showed very uneven plot, with no significant trends. Hence, further modelling needed to be done.

Then we fitted a linear model along with a seasonal parameter. The ACF plot for the residuals from the linear model were as follows.

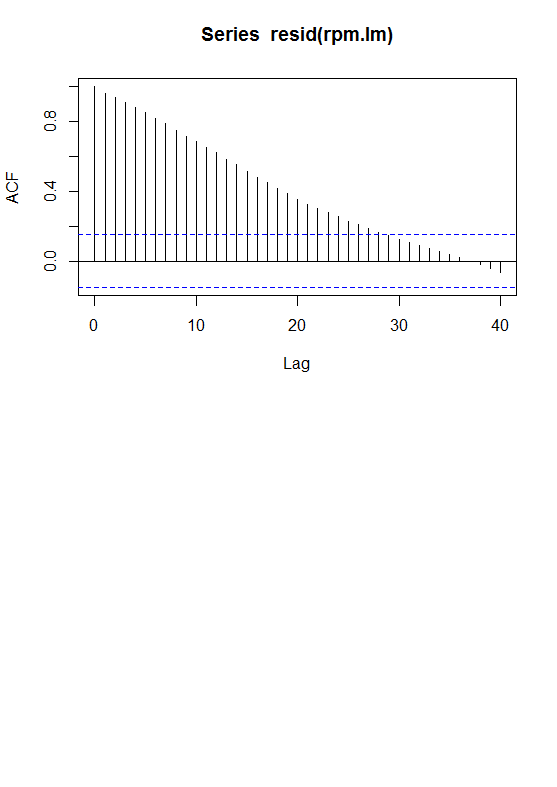


Fig 5. ACF plot for the linear model with seasonality.

The ACF plot didn’t die instead it had went on to negative side of the graph.

So, we tried fitting a SARIMA model. For this we plotted the PACF and ACF curves below, which show us that that the ACF plot dies off at 1 and PACF also dies of at 1, hence the parameters used by us were p=1, q=1, d=1 with seasonality on the AR(1) model.

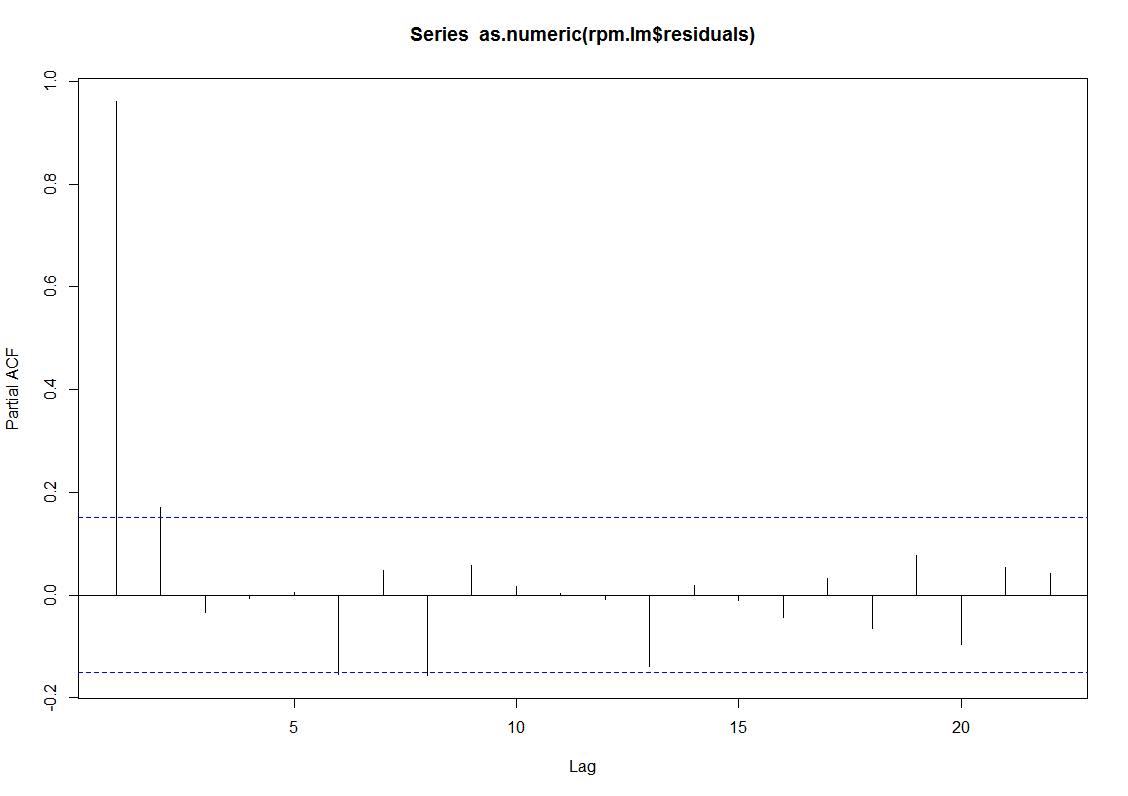


Fig 6. PACF plot for the SARIMA model.

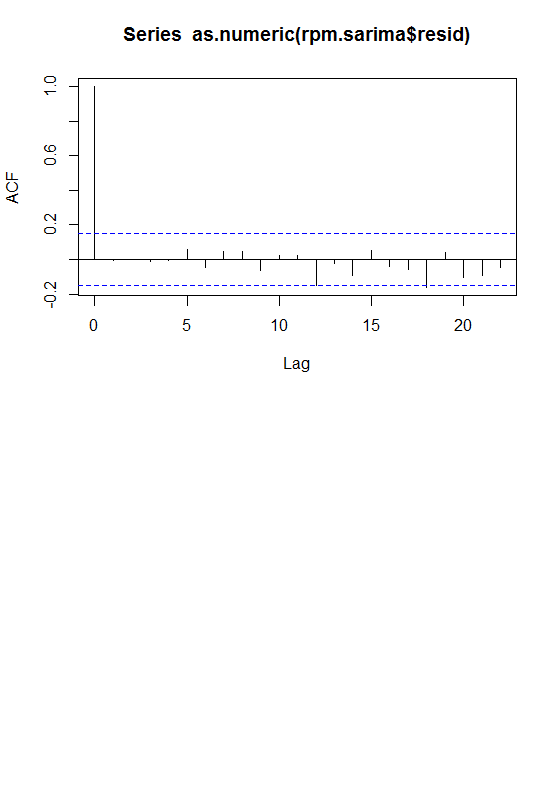


Fig 7. ACF plot for the SARIMA model.

**Forecasting**

After the model building we then applied it on our data which gave us a forecast of the next twelve months in the form of a .txt file which is attached along with this document.

We got the following graph as out put.

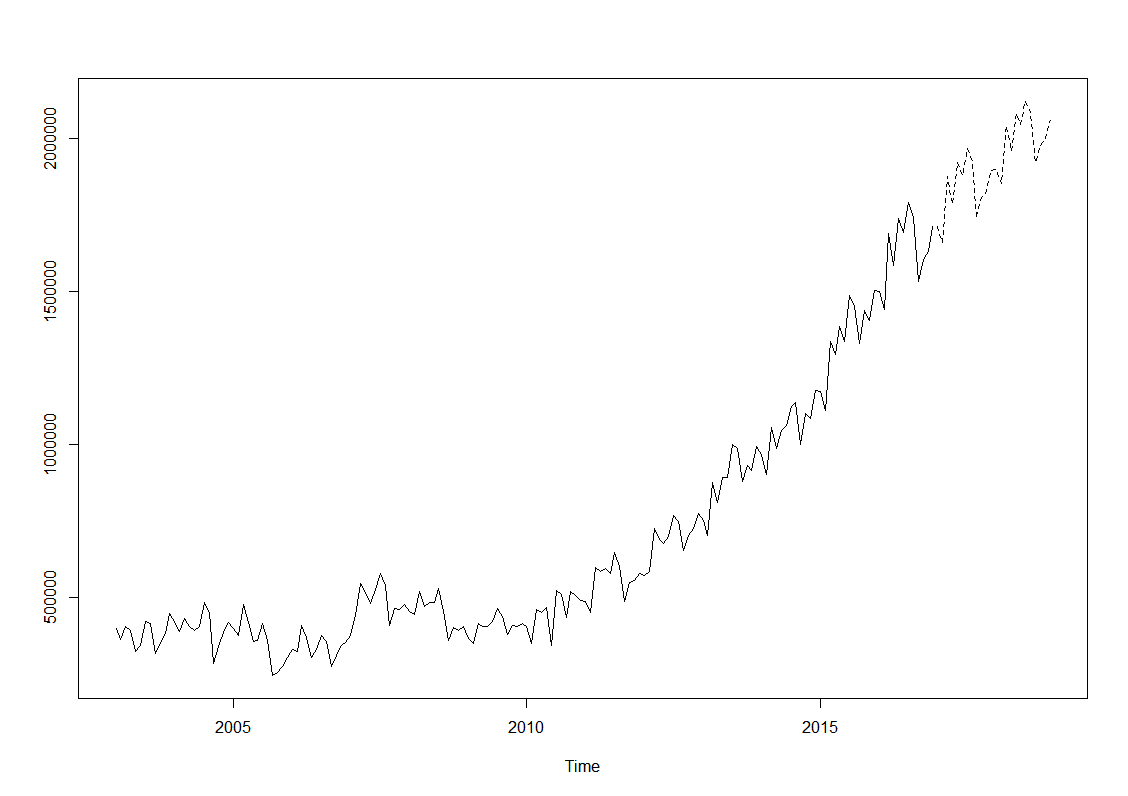


Fig 8. Forecasting output of the SARIMA model with the dotted line being the output.