BF_Final_.R

Sun Dec 09 01:58:07 2018

```
library(fpp)
```

```
## Loading required package: forecast
```

```
## Loading required package: fma
```

```
## Loading required package: expsmooth
```

```
## Loading required package: lmtest
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
## Loading required package: tseries
```

```
library(fpp2)
```

```
## Loading required package: ggplot2
```

```
##
## Attaching package: 'fpp2'
```

```
## The following objects are masked from 'package:fpp':
##
##     ausair, ausbeer, austa, austourists, debitcards, departures,
##     elecequip, euretail, guinearice, oil, sunspotarea, usmelec
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

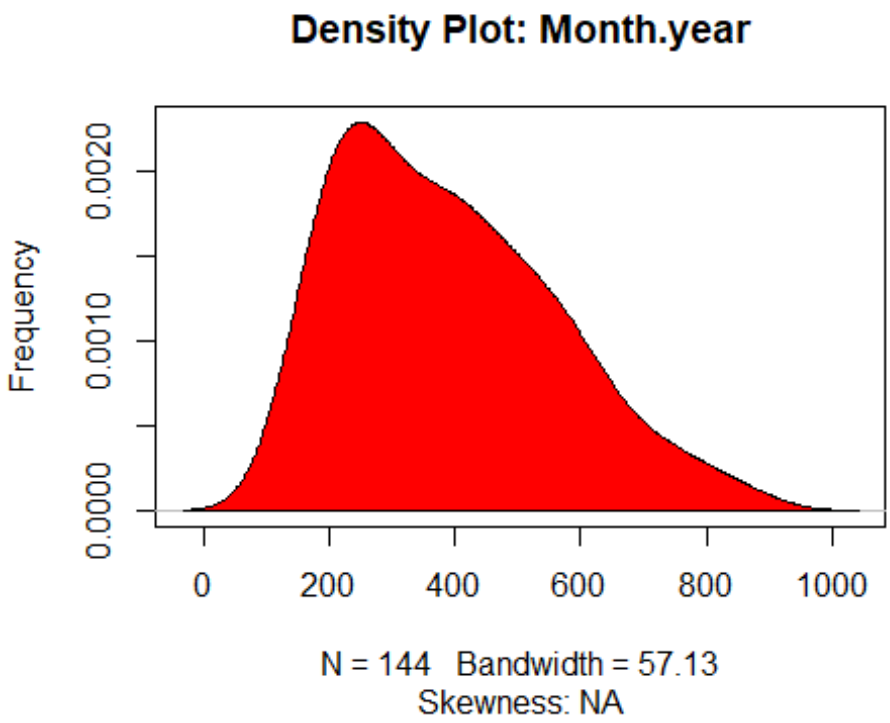
```r
library(forecast)

data<- read.csv("C:\\Users\\ThisPC\\Downloads\\Tractor sales (1).csv")

#1 Create Training and Test data -
set.seed(100)  # setting seed to reproduce results of random sampling
trainingRowIndex <- sample(1:nrow(data), 0.8*nrow(data))  # row indices for t
raining data
trainingData <- data[trainingRowIndex, ]  # model training data
testData  <- data[-trainingRowIndex, ]   # test data

#2 Develop the model on the training data and use it to predict the distance
on test data
# Build the model on training data -
lmMod <- lm(Number.of.tractors.sold ~ Month.year, data=data)  # build the mod
el
salesPred <- predict(lmMod, testData)
summary(lmMod)

##
## Call:
## lm(formula = Number.of.tractors.sold ~ Month.year, data = data)
##
## Residuals:
## ALL 144 residuals are 0: no residual degrees of freedom!
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)             141         NA      NA       NA
## Month.year2003-10         9         NA      NA       NA
## Month.year2003-11        -3         NA      NA       NA
## Month.year2003-12        24         NA      NA       NA
## Month.year2003-2         16         NA      NA       NA
## Month.year2003-3         44         NA      NA       NA
## Month.year2003-4         58         NA      NA       NA
## Month.year2003-5         62         NA      NA       NA
## Month.year2003-6         48         NA      NA       NA
## Month.year2003-7         66         NA      NA       NA
## Month.year2003-8         66         NA      NA       NA
## Month.year2003-9         30         NA      NA       NA
## Month.year2004-1          4         NA      NA       NA
## Month.year2004-10        27         NA      NA       NA
## Month.year2004-11        11         NA      NA       NA
## Month.year2004-12        55         NA      NA       NA
## Month.year2004-2         27         NA      NA       NA
## Month.year2004-3         56         NA      NA       NA
## Month.year2004-4         67         NA      NA       NA
## Month.year2004-5         69         NA      NA       NA
## Month.year2004-6         68         NA      NA       NA
## Month.year2004-7         97         NA      NA       NA
```

```
## Month.year2004-8        97         NA       NA       NA
## Month.year2004-9        58         NA       NA       NA
## Month.year2005-1        42         NA       NA       NA
## Month.year2005-10       63         NA       NA       NA
## Month.year2005-11       53         NA       NA       NA
## Month.year2005-12       91         NA       NA       NA
## Month.year2005-2        59         NA       NA       NA
## Month.year2005-3       108         NA       NA       NA
## Month.year2005-4       110         NA       NA       NA
## Month.year2005-5       148         NA       NA       NA
## Month.year2005-6       108         NA       NA       NA
## Month.year2005-7       138         NA       NA       NA
## Month.year2005-8       138         NA       NA       NA
## Month.year2005-9        91         NA       NA       NA
## Month.year2006-1        74         NA       NA       NA
## Month.year2006-10      100         NA       NA       NA
## Month.year2006-11       88         NA       NA       NA
## Month.year2006-12      131         NA       NA       NA
## Month.year2006-2        98         NA       NA       NA
## Month.year2006-3       129         NA       NA       NA
## Month.year2006-4       138         NA       NA       NA
## Month.year2006-5       166         NA       NA       NA
## Month.year2006-6       164         NA       NA       NA
## Month.year2006-7       181         NA       NA       NA
## Month.year2006-8       198         NA       NA       NA
## Month.year2006-9       122         NA       NA       NA
## Month.year2007-1       106         NA       NA       NA
## Month.year2007-10      125         NA       NA       NA
## Month.year2007-11       98         NA       NA       NA
## Month.year2007-12      140         NA       NA       NA
## Month.year2007-2       120         NA       NA       NA
## Month.year2007-3       189         NA       NA       NA
## Month.year2007-4       221         NA       NA       NA
## Month.year2007-5       244         NA       NA       NA
## Month.year2007-6       199         NA       NA       NA
## Month.year2007-7       229         NA       NA       NA
## Month.year2007-8       240         NA       NA       NA
## Month.year2007-9       158         NA       NA       NA
## Month.year2008-1       116         NA       NA       NA
## Month.year2008-10      148         NA       NA       NA
## Month.year2008-11      129         NA       NA       NA
## Month.year2008-12      180         NA       NA       NA
## Month.year2008-2       109         NA       NA       NA
## Month.year2008-3       188         NA       NA       NA
## Month.year2008-4       209         NA       NA       NA
## Month.year2008-5       252         NA       NA       NA
## Month.year2008-6       229         NA       NA       NA
## Month.year2008-7       282         NA       NA       NA
## Month.year2008-8       269         NA       NA       NA
## Month.year2008-9       185         NA       NA       NA
```

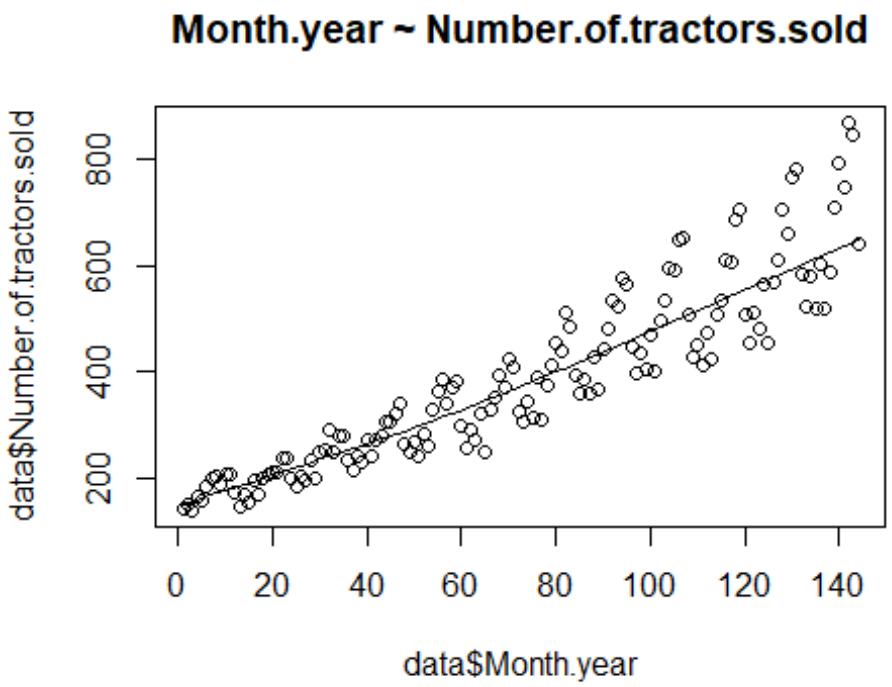
```
## Month.year2009-1       164          NA       NA       NA
## Month.year2009-10      204          NA       NA       NA
## Month.year2009-11      174          NA       NA       NA
## Month.year2009-12      248          NA       NA       NA
## Month.year2009-2       169          NA       NA       NA
## Month.year2009-3       233          NA       NA       NA
## Month.year2009-4       273          NA       NA       NA
## Month.year2009-5       313          NA       NA       NA
## Month.year2009-6       300          NA       NA       NA
## Month.year2009-7       369          NA       NA       NA
## Month.year2009-8       345          NA       NA       NA
## Month.year2009-9       252          NA       NA       NA
## Month.year2010-1       217          NA       NA       NA
## Month.year2010-10      245          NA       NA       NA
## Month.year2010-11      219          NA       NA       NA
## Month.year2010-12      287          NA       NA       NA
## Month.year2010-2       227          NA       NA       NA
## Month.year2010-3       303          NA       NA       NA
## Month.year2010-4       341          NA       NA       NA
## Month.year2010-5       393          NA       NA       NA
## Month.year2010-6       383          NA       NA       NA
## Month.year2010-7       437          NA       NA       NA
## Month.year2010-8       426          NA       NA       NA
## Month.year2010-9       306          NA       NA       NA
## Month.year2011-1       256          NA       NA       NA
## Month.year2011-10      296          NA       NA       NA
## Month.year2011-11      265          NA       NA       NA
## Month.year2011-12      329          NA       NA       NA
## Month.year2011-2       259          NA       NA       NA
## Month.year2011-3       357          NA       NA       NA
## Month.year2011-4       395          NA       NA       NA
## Month.year2011-5       455          NA       NA       NA
## Month.year2011-6       450          NA       NA       NA
## Month.year2011-7       510          NA       NA       NA
## Month.year2011-8       513          NA       NA       NA
## Month.year2011-9       368          NA       NA       NA
## Month.year2012-1       287          NA       NA       NA
## Month.year2012-10      311          NA       NA       NA
## Month.year2012-11      271          NA       NA       NA
## Month.year2012-12      331          NA       NA       NA
## Month.year2012-2       282          NA       NA       NA
## Month.year2012-3       366          NA       NA       NA
## Month.year2012-4       395          NA       NA       NA
## Month.year2012-5       469          NA       NA       NA
## Month.year2012-6       468          NA       NA       NA
## Month.year2012-7       546          NA       NA       NA
## Month.year2012-8       566          NA       NA       NA
## Month.year2012-9       368          NA       NA       NA
## Month.year2013-1       313          NA       NA       NA
## Month.year2013-10      372          NA       NA       NA
```

```
## Month.year2013-11      340         NA      NA       NA
## Month.year2013-12      426         NA      NA       NA
## Month.year2013-2       314         NA      NA       NA
## Month.year2013-3       427         NA      NA       NA
## Month.year2013-4       469         NA      NA       NA
## Month.year2013-5       565         NA      NA       NA
## Month.year2013-6       520         NA      NA       NA
## Month.year2013-7       626         NA      NA       NA
## Month.year2013-8       642         NA      NA       NA
## Month.year2013-9       442         NA      NA       NA
## Month.year2014-1       384         NA      NA       NA
## Month.year2014-10      440         NA      NA       NA
## Month.year2014-11      378         NA      NA       NA
## Month.year2014-12      464         NA      NA       NA
## Month.year2014-2       379         NA      NA       NA
## Month.year2014-3       446         NA      NA       NA
## Month.year2014-4       569         NA      NA       NA
## Month.year2014-5       652         NA      NA       NA
## Month.year2014-6       608         NA      NA       NA
## Month.year2014-7       730         NA      NA       NA
## Month.year2014-8       707         NA      NA       NA
## Month.year2014-9       499         NA      NA       NA
##
## Residual standard error: NaN on 0 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:    NaN
## F-statistic:   NaN on 143 and 0 DF,  p-value: NA
```

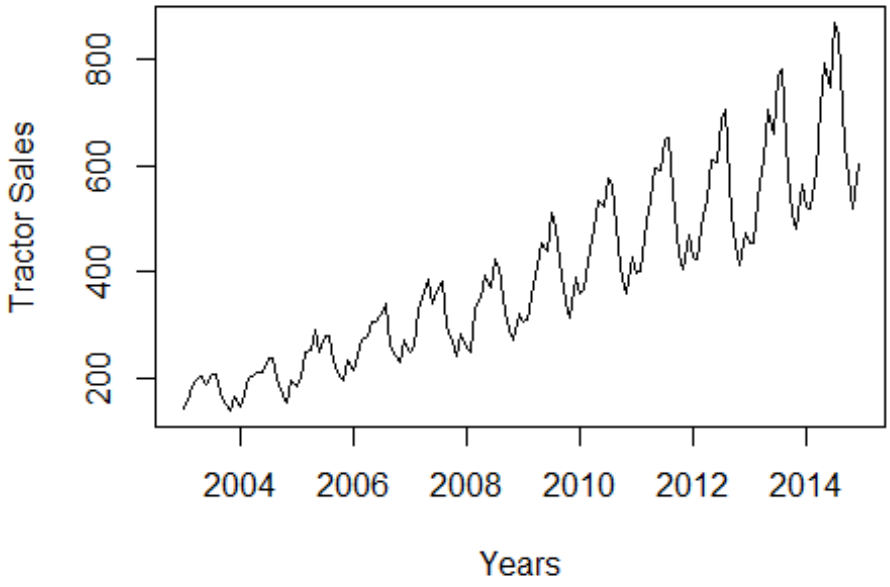
```
#3 Density plot to check if the response variable is close to normality

library(e1071)
par(mfrow=c(1, 1))  # divide graph area in 2 columns
plot(density(data$Number.of.tractors.sold), main="Density Plot: Month.year",
ylab="Frequency", sub=paste("Skewness:", round(e1071::skewness(data$Month.yea
r), 2)))  # density plot for 'Month.year'

## Warning in mean.default(x): argument is not numeric or logical: returning
## NA

## Warning in Ops.factor(x, mean(x)): '-' not meaningful for factors

polygon(density(data$Number.of.tractors.sold), col="red")
```



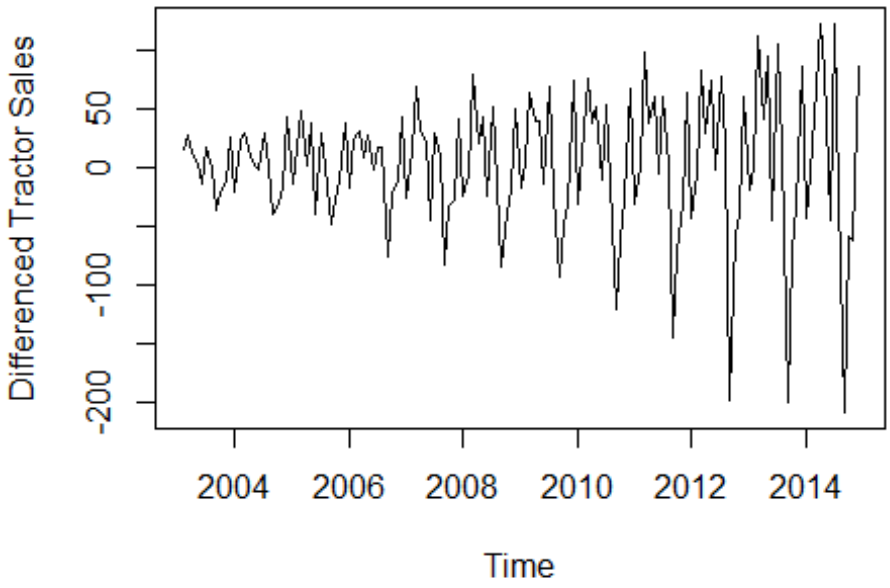

```r
#4 Scatter plot to visualize the linear relationship between the predictor an
d response

scatter.smooth(x=data$Month.year, y=data$Number.of.tractors.sold, main="Month
.year ~ Number.of.tractors.sold")
```



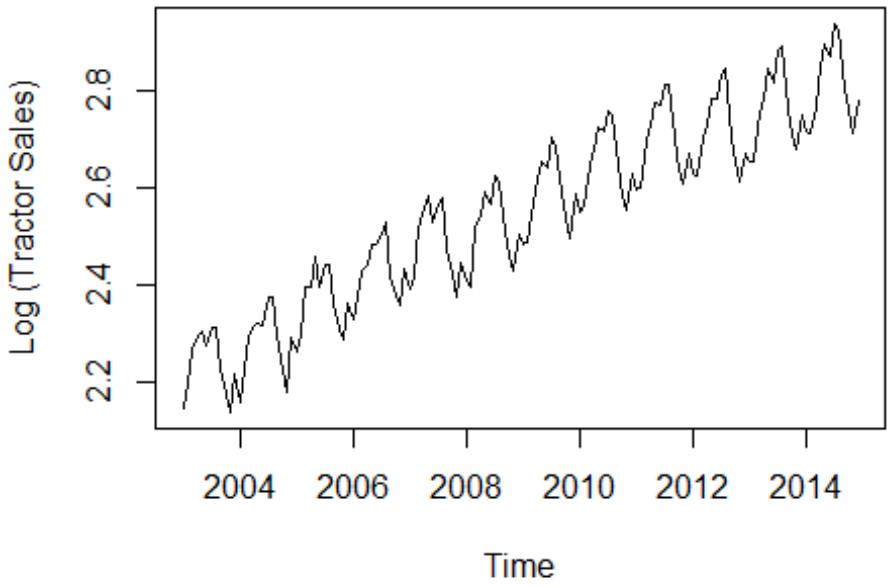

```r
#5 Reading and Plotting the tractor sales data as time series

data<-read.csv("C:\\Users\\APEKSHA\\Downloads\\Tractor sales (1).csv")
data = ts(data[,2],start = c(2003,1),frequency = 12)
plot(data, xlab='Years', ylab = 'Tractor Sales')
```



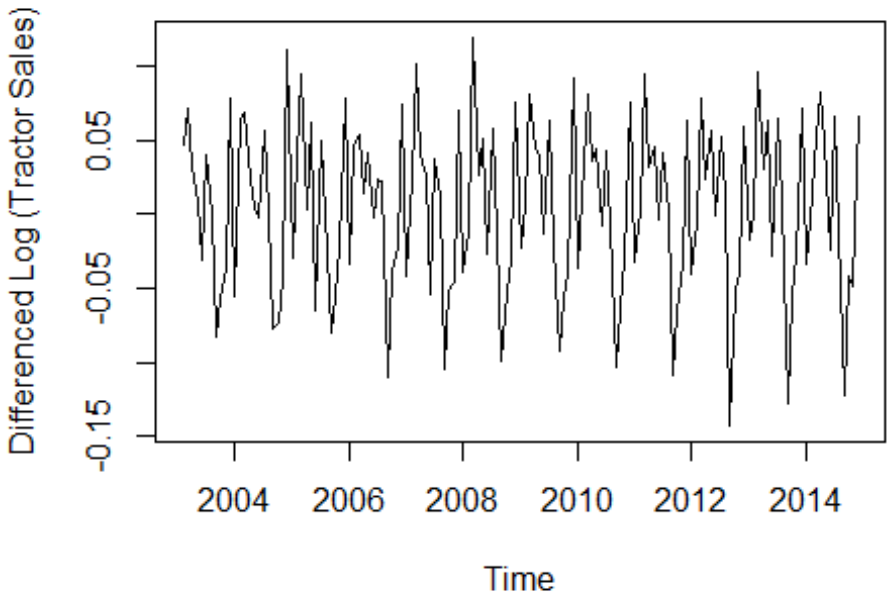

```r
#6 Differencing the  data to make data stationary on mean (remove trend)

plot(diff(data),ylab='Differenced Tractor Sales')
```



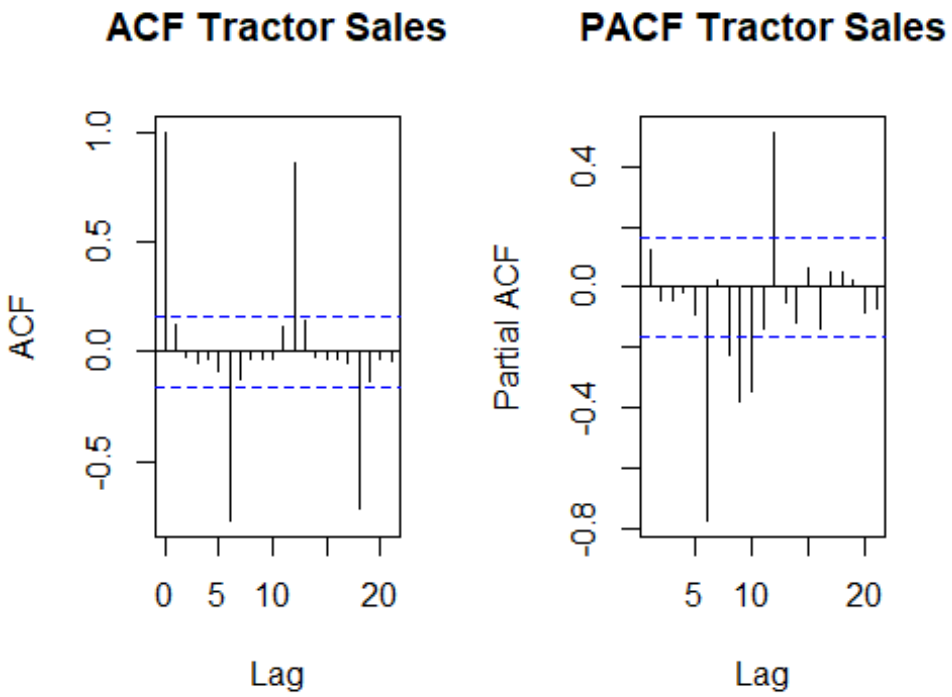
```r
#7 log transforming the data to make data stationary on variance

plot(log10(data),ylab='Log (Tractor Sales)')
```



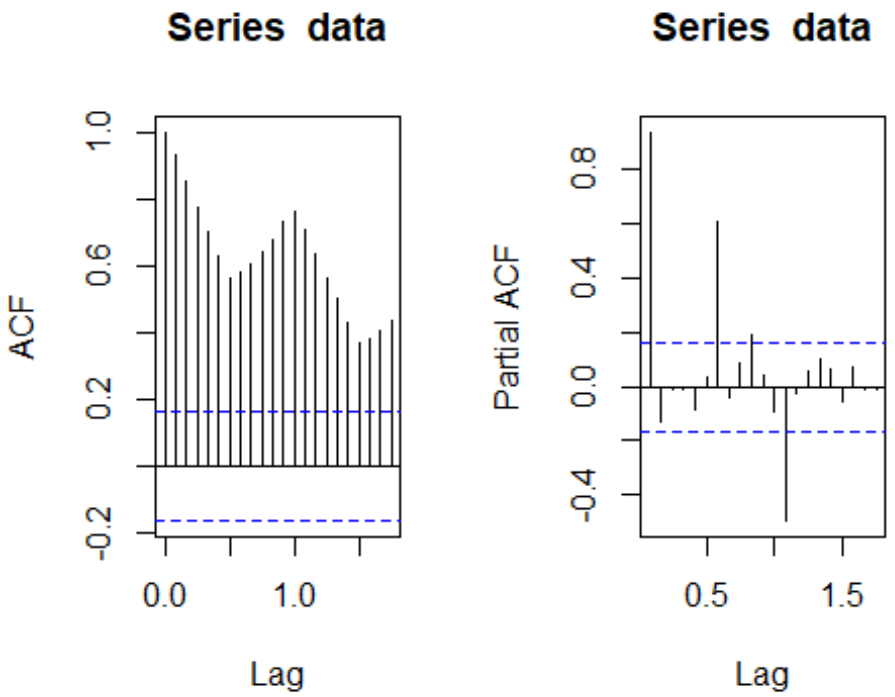

```r
#8 Differencing the log transform data to make data stationary on both mean a
nd variance.

plot(diff(log10(data)),ylab='Differenced Log (Tractor Sales)')
```




```r
#9 Plotting ACF and PACF to identify potential AR and MA model

par(mfrow = c(1,2))
acf(ts(diff(log10(data))),main='ACF Tractor Sales')
pacf(ts(diff(log10(data))),main='PACF Tractor Sales')
```




```r
#10 Both acf() and pacf() generates plots by default

acfRes <- acf(data) # autocorrelation
pacfRes <- pacf(data)  # partial autocorrelation
```

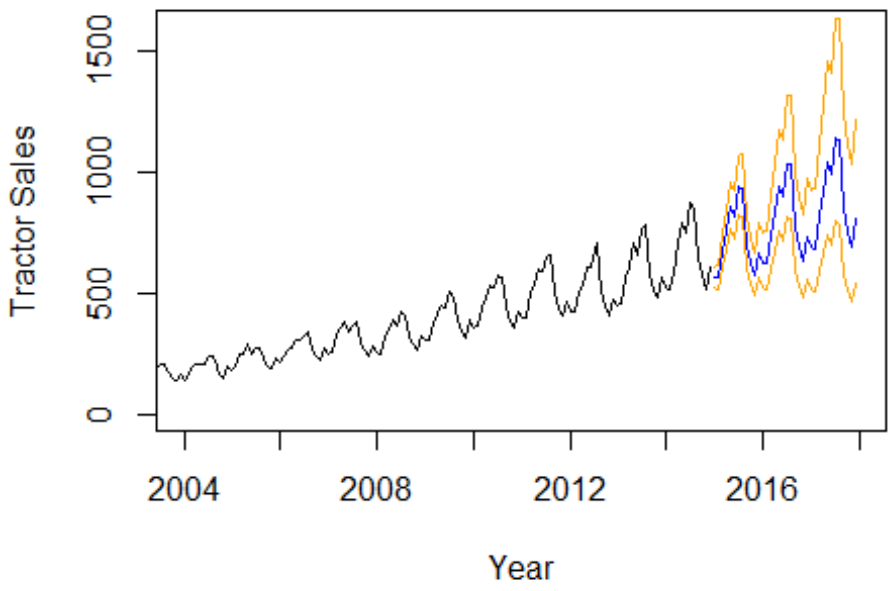



```
##ARIMA

#11 Identification of the best fit ARIMA model

library(tseries)
ARIMAfit = auto.arima(log10(data), approximation=FALSE,trace=FALSE)
summary(ARIMAfit)

## Series: log10(data)
## ARIMA(0,1,1)(0,1,1)[12]
##
## Coefficients:
##           ma1     sma1
##       -0.4047  -0.5529
## s.e.   0.0885   0.0734
##
## sigma^2 estimated as 0.0002571:  log likelihood=354.4
## AIC=-702.79   AICc=-702.6   BIC=-694.17
##
## Training set error measures:
##                       ME       RMSE        MAE         MPE      MAPE
## Training set 0.0002410698 0.01517695 0.01135312 0.008335713 0.4462212
##                    MASE       ACF1
## Training set 0.2158968 0.01062604

#12 Forecasting the sales using the best fit ARIMA model
```

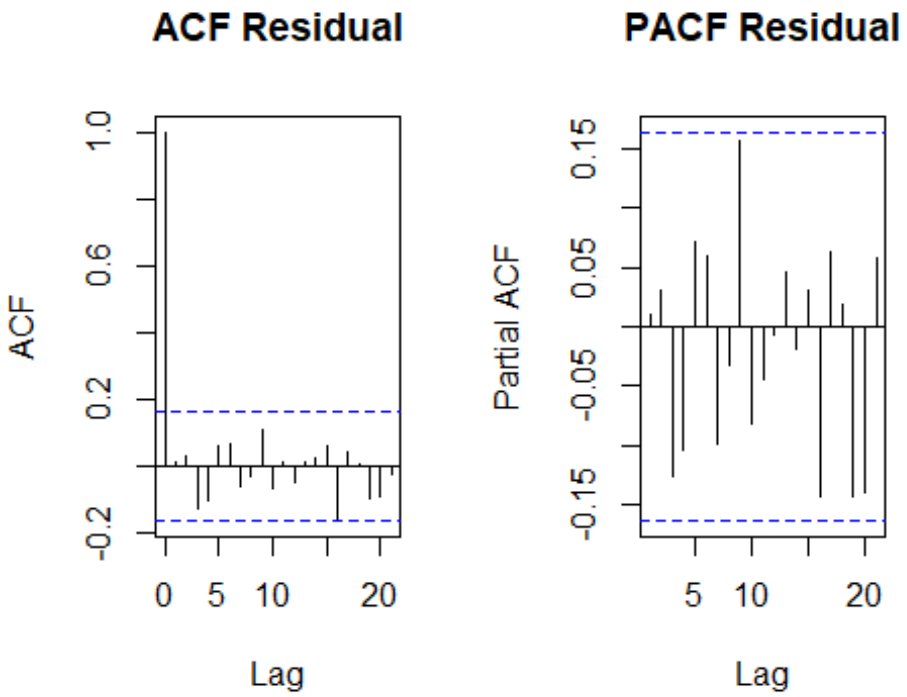
```r
par(mfrow = c(1,1))
pred = predict(ARIMAfit, n.ahead = 36)
pred
```

```
## $pred
##           Jan      Feb      Mar      Apr      May      Jun      Jul
## 2015 2.754168 2.753182 2.826608 2.880192 2.932447 2.912372 2.972538
## 2016 2.796051 2.795065 2.868491 2.922075 2.974330 2.954255 3.014421
## 2017 2.837934 2.836948 2.910374 2.963958 3.016213 2.996138 3.056304
##           Aug      Sep      Oct      Nov      Dec
## 2015 2.970585 2.847264 2.797259 2.757395 2.825125
## 2016 3.012468 2.889147 2.839142 2.799278 2.867008
## 2017 3.054351 2.931030 2.881025 2.841161 2.908891
##
## $se
##             Jan        Feb        Mar        Apr        May        Jun
## 2015 0.01603508 0.01866159 0.02096153 0.02303295 0.02493287 0.02669792
## 2016 0.03923008 0.04159145 0.04382576 0.04595157 0.04798329 0.04993241
## 2017 0.06386474 0.06637555 0.06879478 0.07113179 0.07339441 0.07558934
##             Jul        Aug        Sep        Oct        Nov        Dec
## 2015 0.02835330 0.02991723 0.03140337 0.03282229 0.03418236 0.03549035
## 2016 0.05180825 0.05361850 0.05536960 0.05706700 0.05871534 0.06031866
## 2017 0.07772231 0.07979828 0.08182160 0.08379608 0.08572510 0.08761165
```

```r
plot(data,type='l',xlim=c(2004,2018),ylim=c(1,1600),xlab = 'Year',ylab = 'Tra
ctor Sales')
lines(10^(pred$pred),col='blue')
lines(10^(pred$pred+2*pred$se),col='orange')
lines(10^(pred$pred-2*pred$se),col='orange')
```




```
#13 Plotting ACF and PACF for residuals of ARIMA model to check if there's mo
re information left for extraction.

par(mfrow=c(1,2))
acf(ts(ARIMAfit$residuals),main='ACF Residual')
pacf(ts(ARIMAfit$residuals),main='PACF Residual')
```


ACF Residual
ACF
1.0
0.6
0.2
-0.2
0
5
10
20
Lag
PACF Residual
Partial ACF
0.15
0.05
-0.05
-0.15
5
10
20
Lag