PHD\_Regressin.R

SanjeevaReddy Kota

Mon Nov 20 13:28:24 2017

rm(list=ls(all=TRUE))  
  
setwd("G:/PHD Hackthon")  
  
  
library(readxl)

## Warning: package 'readxl' was built under R version 3.4.2

library(forecast)  
library(lubridate)

## Warning: package 'lubridate' was built under R version 3.4.2

##   
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':  
##   
## date

library(DataCombine)

## Warning: package 'DataCombine' was built under R version 3.4.2

library(imputeTS)

## Warning: package 'imputeTS' was built under R version 3.4.2

library(plyr)

##   
## Attaching package: 'plyr'

## The following object is masked from 'package:lubridate':  
##   
## here

library(dplyr)

## Warning: package 'dplyr' was built under R version 3.4.2

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:plyr':  
##   
## arrange, count, desc, failwith, id, mutate, rename, summarise,  
## summarize

## The following objects are masked from 'package:lubridate':  
##   
## intersect, setdiff, union

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

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

library(TTR)

## Warning: package 'TTR' was built under R version 3.4.2

library(graphics)  
library(data.table)

## Warning: package 'data.table' was built under R version 3.4.2

##   
## Attaching package: 'data.table'

## The following objects are masked from 'package:dplyr':  
##   
## between, first, last

## The following object is masked from 'package:DataCombine':  
##   
## shift

## The following objects are masked from 'package:lubridate':  
##   
## hour, isoweek, mday, minute, month, quarter, second, wday,  
## week, yday, year

library(Quandl)

## Loading required package: xts

## Loading required package: zoo

## Warning: package 'zoo' was built under R version 3.4.2

##   
## Attaching package: 'zoo'

## The following object is masked from 'package:imputeTS':  
##   
## na.locf

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

##   
## Attaching package: 'xts'

## The following objects are masked from 'package:data.table':  
##   
## first, last

## The following objects are masked from 'package:dplyr':  
##   
## first, last

library(DMwR)

## Loading required package: lattice

## Loading required package: grid

##   
## Attaching package: 'DMwR'

## The following object is masked from 'package:plyr':  
##   
## join

library(gtools)

##   
## Attaching package: 'gtools'

## The following object is masked from 'package:imputeTS':  
##   
## na.replace

library(car)

##   
## Attaching package: 'car'

## The following object is masked from 'package:gtools':  
##   
## logit

## The following object is masked from 'package:dplyr':  
##   
## recode

#traindata  
train=read.csv('Train.csv',header = TRUE,sep=',')  
train\_women=train[train$ProductCategory=='WomenClothing',]  
#droping product category  
train\_women$ProductCategory=NULL  
str(train\_women)

## 'data.frame': 84 obs. of 3 variables:  
## $ Year : int 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 ...  
## $ Month : int 1 2 3 4 5 6 7 8 9 10 ...  
## $ Sales.In.ThousandDollars.: int 1755 1729 2256 2662 2732 2220 2164 2371 2421 2579 ...

#Holidays  
EventHolidays=read\_xlsx('Events\_HolidaysData.xlsx',sheet=1,col\_names=TRUE)  
str(EventHolidays)

## Classes 'tbl\_df', 'tbl' and 'data.frame': 150 obs. of 4 variables:  
## $ Year : num 2009 2009 2009 2009 2009 ...  
## $ MonthDate : POSIXct, format: "2001-01-01" "2019-01-01" ...  
## $ Event : chr "New Year's Day" "Martin Luther King Jr. Day" "Valentine's Day" "Presidents' Day" ...  
## $ DayCategory: chr "Federal Holiday" "Federal Holiday" "Event" "Federal Holiday" ...

#MacroEconomics  
MacroEconomics=read\_xlsx('MacroEconomicData.xlsx',sheet=1,col\_names=TRUE)  
str(MacroEconomics)

## Classes 'tbl\_df', 'tbl' and 'data.frame': 96 obs. of 18 variables:  
## $ Year-Month : chr "2009 - Jan" "2009 - Feb" "2009 - Mar" "2009 - Apr" ...  
## $ Monthly Nominal GDP Index (inMillion$) : num 14422 14389 14341 14327 14346 ...  
## $ Monthly Real GDP Index (inMillion$) : num 14407 14366 14352 14352 14368 ...  
## $ CPI : num 233 235 235 236 236 ...  
## $ PartyInPower : chr "Democrats" "Democrats" "Democrats" "Democrats" ...  
## $ unemployment rate : num 7.8 8.3 8.7 9 9.4 9.5 9.5 9.6 9.8 10 ...  
## $ CommercialBankInterestRateonCreditCardPlans : num 12 13 13 13 13.3 ...  
## $ Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan: num 11.4 11.1 11.1 11.1 11.2 ...  
## $ Earnings or wages in dollars per hour : num 22.1 22.2 22.2 22.1 22 ...  
## $ AdvertisingExpenses (in Thousand Dollars) : chr "137" "200" "?" "214" ...  
## $ Cotton Monthly Price - US cents per Pound(lbs) : num 57.7 55.2 51.5 56.8 62 ...  
## $ Change(in%) : num 4.02 -4.32 -6.72 10.25 9.11 ...  
## $ Average upland planted(million acres) : num 9.3 9.3 9.3 9.3 9.3 ...  
## $ Average upland harvested(million acres) : num 7.56 7.56 7.56 7.56 7.4 ...  
## $ yieldperharvested acre : num 799 799 799 787 803 796 796 809 827 799 ...  
## $ Production (in 480-lb netweright in million bales) : num 12.6 12.6 12.6 12.4 12.4 ...  
## $ Mill use (in 480-lb netweright in million bales) : num 4.17 3.87 3.72 3.62 3.52 3.47 3.47 3.47 3.47 3.37 ...  
## $ Exports : num 11.6 11.1 11.6 12.2 12.3 ...

Train\_MacroEconomics=MacroEconomics[1:84,]  
str(Train\_MacroEconomics)

## Classes 'tbl\_df', 'tbl' and 'data.frame': 84 obs. of 18 variables:  
## $ Year-Month : chr "2009 - Jan" "2009 - Feb" "2009 - Mar" "2009 - Apr" ...  
## $ Monthly Nominal GDP Index (inMillion$) : num 14422 14389 14341 14327 14346 ...  
## $ Monthly Real GDP Index (inMillion$) : num 14407 14366 14352 14352 14368 ...  
## $ CPI : num 233 235 235 236 236 ...  
## $ PartyInPower : chr "Democrats" "Democrats" "Democrats" "Democrats" ...  
## $ unemployment rate : num 7.8 8.3 8.7 9 9.4 9.5 9.5 9.6 9.8 10 ...  
## $ CommercialBankInterestRateonCreditCardPlans : num 12 13 13 13 13.3 ...  
## $ Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan: num 11.4 11.1 11.1 11.1 11.2 ...  
## $ Earnings or wages in dollars per hour : num 22.1 22.2 22.2 22.1 22 ...  
## $ AdvertisingExpenses (in Thousand Dollars) : chr "137" "200" "?" "214" ...  
## $ Cotton Monthly Price - US cents per Pound(lbs) : num 57.7 55.2 51.5 56.8 62 ...  
## $ Change(in%) : num 4.02 -4.32 -6.72 10.25 9.11 ...  
## $ Average upland planted(million acres) : num 9.3 9.3 9.3 9.3 9.3 ...  
## $ Average upland harvested(million acres) : num 7.56 7.56 7.56 7.56 7.4 ...  
## $ yieldperharvested acre : num 799 799 799 787 803 796 796 809 827 799 ...  
## $ Production (in 480-lb netweright in million bales) : num 12.6 12.6 12.6 12.4 12.4 ...  
## $ Mill use (in 480-lb netweright in million bales) : num 4.17 3.87 3.72 3.62 3.52 3.47 3.47 3.47 3.47 3.37 ...  
## $ Exports : num 11.6 11.1 11.6 12.2 12.3 ...

Test\_MacroEconomics=MacroEconomics[85:96,]  
str(Test\_MacroEconomics)

## Classes 'tbl\_df', 'tbl' and 'data.frame': 12 obs. of 18 variables:  
## $ Year-Month : chr "2016 - Jan" "2016 - Feb" "2016 - Mar" "2016 - Apr" ...  
## $ Monthly Nominal GDP Index (inMillion$) : num 18287 18242 18447 18534 18498 ...  
## $ Monthly Real GDP Index (inMillion$) : num 16546 16504 16665 16694 16627 ...  
## $ CPI : num 260 261 262 263 263 ...  
## $ PartyInPower : chr "Democrats" "Democrats" "Democrats" "Democrats" ...  
## $ unemployment rate : num 4.9 4.9 5 5 4.7 4.9 4.9 4.9 4.9 4.8 ...  
## $ CommercialBankInterestRateonCreditCardPlans : num 12.2 12.3 12.3 12.3 12.2 ...  
## $ Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan: num 9.66 10.03 10.03 10.03 9.65 ...  
## $ Earnings or wages in dollars per hour : num 25.5 25.5 25.5 25.6 25.7 ...  
## $ AdvertisingExpenses (in Thousand Dollars) : chr "?" "?" "?" "?" ...  
## $ Cotton Monthly Price - US cents per Pound(lbs) : num 68.8 66.6 65.5 69.3 70.3 ...  
## $ Change(in%) : num -2.33 -3.17 -1.67 5.84 1.44 5.44 9.39 -0.99 -2.99 0.85 ...  
## $ Average upland planted(million acres) : num 8.42 8.42 8.42 8.42 8.42 ...  
## $ Average upland harvested(million acres) : num 7.92 7.92 7.9 7.9 7.92 ...  
## $ yieldperharvested acre : num 758 758 760 756 755 792 802 787 790 785 ...  
## $ Production (in 480-lb netweright in million bales) : num 12.5 12.5 12.5 12.4 12.5 ...  
## $ Mill use (in 480-lb netweright in million bales) : num 3.58 3.58 3.58 3.58 3.58 ...  
## $ Exports : num 9.53 9.05 9.05 9.03 8.5 ...

Test\_MacroEconomics$`Year-Month`=NULL  
Test\_MacroEconomics$PartyInPower=NULL  
#Test\_MacroEconomics$`AdvertisingExpenses (in Thousand Dollars)`=NULL  
str(Test\_MacroEconomics)

## Classes 'tbl\_df', 'tbl' and 'data.frame': 12 obs. of 16 variables:  
## $ Monthly Nominal GDP Index (inMillion$) : num 18287 18242 18447 18534 18498 ...  
## $ Monthly Real GDP Index (inMillion$) : num 16546 16504 16665 16694 16627 ...  
## $ CPI : num 260 261 262 263 263 ...  
## $ unemployment rate : num 4.9 4.9 5 5 4.7 4.9 4.9 4.9 4.9 4.8 ...  
## $ CommercialBankInterestRateonCreditCardPlans : num 12.2 12.3 12.3 12.3 12.2 ...  
## $ Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan: num 9.66 10.03 10.03 10.03 9.65 ...  
## $ Earnings or wages in dollars per hour : num 25.5 25.5 25.5 25.6 25.7 ...  
## $ AdvertisingExpenses (in Thousand Dollars) : chr "?" "?" "?" "?" ...  
## $ Cotton Monthly Price - US cents per Pound(lbs) : num 68.8 66.6 65.5 69.3 70.3 ...  
## $ Change(in%) : num -2.33 -3.17 -1.67 5.84 1.44 5.44 9.39 -0.99 -2.99 0.85 ...  
## $ Average upland planted(million acres) : num 8.42 8.42 8.42 8.42 8.42 ...  
## $ Average upland harvested(million acres) : num 7.92 7.92 7.9 7.9 7.92 ...  
## $ yieldperharvested acre : num 758 758 760 756 755 792 802 787 790 785 ...  
## $ Production (in 480-lb netweright in million bales) : num 12.5 12.5 12.5 12.4 12.5 ...  
## $ Mill use (in 480-lb netweright in million bales) : num 3.58 3.58 3.58 3.58 3.58 ...  
## $ Exports : num 9.53 9.05 9.05 9.03 8.5 ...

library(readxl)  
Wether1=read\_xlsx('WeatherData.xlsx',sheet = '2009',col\_names=TRUE)  
Wether2=read\_xlsx('WeatherData.xlsx',sheet = '2010',col\_names=TRUE)  
Wether3=read\_xlsx('WeatherData.xlsx',sheet = '2011',col\_names=TRUE)  
Wether4=read\_xlsx('WeatherData.xlsx',sheet = '2012',col\_names=TRUE)  
Wether5=read\_xlsx('WeatherData.xlsx',sheet = '2013',col\_names=TRUE)  
Wether6=read\_xlsx('WeatherData.xlsx',sheet = '2014',col\_names=TRUE)  
Wether7=read\_xlsx('WeatherData.xlsx',sheet = '2015',col\_names=TRUE)  
Wether8=read\_xlsx('WeatherData.xlsx',sheet = '2016',col\_names=TRUE)  
  
  
str(MacroEconomics)

## Classes 'tbl\_df', 'tbl' and 'data.frame': 96 obs. of 18 variables:  
## $ Year-Month : chr "2009 - Jan" "2009 - Feb" "2009 - Mar" "2009 - Apr" ...  
## $ Monthly Nominal GDP Index (inMillion$) : num 14422 14389 14341 14327 14346 ...  
## $ Monthly Real GDP Index (inMillion$) : num 14407 14366 14352 14352 14368 ...  
## $ CPI : num 233 235 235 236 236 ...  
## $ PartyInPower : chr "Democrats" "Democrats" "Democrats" "Democrats" ...  
## $ unemployment rate : num 7.8 8.3 8.7 9 9.4 9.5 9.5 9.6 9.8 10 ...  
## $ CommercialBankInterestRateonCreditCardPlans : num 12 13 13 13 13.3 ...  
## $ Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan: num 11.4 11.1 11.1 11.1 11.2 ...  
## $ Earnings or wages in dollars per hour : num 22.1 22.2 22.2 22.1 22 ...  
## $ AdvertisingExpenses (in Thousand Dollars) : chr "137" "200" "?" "214" ...  
## $ Cotton Monthly Price - US cents per Pound(lbs) : num 57.7 55.2 51.5 56.8 62 ...  
## $ Change(in%) : num 4.02 -4.32 -6.72 10.25 9.11 ...  
## $ Average upland planted(million acres) : num 9.3 9.3 9.3 9.3 9.3 ...  
## $ Average upland harvested(million acres) : num 7.56 7.56 7.56 7.56 7.4 ...  
## $ yieldperharvested acre : num 799 799 799 787 803 796 796 809 827 799 ...  
## $ Production (in 480-lb netweright in million bales) : num 12.6 12.6 12.6 12.4 12.4 ...  
## $ Mill use (in 480-lb netweright in million bales) : num 4.17 3.87 3.72 3.62 3.52 3.47 3.47 3.47 3.47 3.37 ...  
## $ Exports : num 11.6 11.1 11.6 12.2 12.3 ...

EventHolidays=NULL  
  
head(EventHolidays)

## NULL

EventHolidays$MonthDate=as.character(EventHolidays$MonthDate)  
str(EventHolidays)

## List of 1  
## $ MonthDate: chr(0)

Holiday\_Rows=dim(EventHolidays)[1]  
Holiday\_Col=dim(EventHolidays)[2]  
EventHolidays$MonthDate=as.character(EventHolidays$MonthDate)  
  
print(Holiday\_Rows)

## NULL

##Macro Economics  
sum(is.na(Train\_MacroEconomics))

## [1] 0

#droping AdvertisingExpenses (in Thousand Dollars)  
#Train\_MacroEconomics$`AdvertisingExpenses (in Thousand Dollars)=NULL  
  
##BINDING WITH TRAIN DATA  
  
Train\_df=cbind(train\_women,Train\_MacroEconomics)  
  
str(Train\_df)

## 'data.frame': 84 obs. of 21 variables:  
## $ Year : int 2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 ...  
## $ Month : int 1 2 3 4 5 6 7 8 9 10 ...  
## $ Sales.In.ThousandDollars. : int 1755 1729 2256 2662 2732 2220 2164 2371 2421 2579 ...  
## $ Year-Month : chr "2009 - Jan" "2009 - Feb" "2009 - Mar" "2009 - Apr" ...  
## $ Monthly Nominal GDP Index (inMillion$) : num 14422 14389 14341 14327 14346 ...  
## $ Monthly Real GDP Index (inMillion$) : num 14407 14366 14352 14352 14368 ...  
## $ CPI : num 233 235 235 236 236 ...  
## $ PartyInPower : chr "Democrats" "Democrats" "Democrats" "Democrats" ...  
## $ unemployment rate : num 7.8 8.3 8.7 9 9.4 9.5 9.5 9.6 9.8 10 ...  
## $ CommercialBankInterestRateonCreditCardPlans : num 12 13 13 13 13.3 ...  
## $ Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan: num 11.4 11.1 11.1 11.1 11.2 ...  
## $ Earnings or wages in dollars per hour : num 22.1 22.2 22.2 22.1 22 ...  
## $ AdvertisingExpenses (in Thousand Dollars) : chr "137" "200" "?" "214" ...  
## $ Cotton Monthly Price - US cents per Pound(lbs) : num 57.7 55.2 51.5 56.8 62 ...  
## $ Change(in%) : num 4.02 -4.32 -6.72 10.25 9.11 ...  
## $ Average upland planted(million acres) : num 9.3 9.3 9.3 9.3 9.3 ...  
## $ Average upland harvested(million acres) : num 7.56 7.56 7.56 7.56 7.4 ...  
## $ yieldperharvested acre : num 799 799 799 787 803 796 796 809 827 799 ...  
## $ Production (in 480-lb netweright in million bales) : num 12.6 12.6 12.6 12.4 12.4 ...  
## $ Mill use (in 480-lb netweright in million bales) : num 4.17 3.87 3.72 3.62 3.52 3.47 3.47 3.47 3.47 3.37 ...  
## $ Exports : num 11.6 11.1 11.6 12.2 12.3 ...

##checking column wise missing values  
apply(Train\_df,2,function(x) sum(is.na(x)))

## Year   
## 0   
## Month   
## 0   
## Sales.In.ThousandDollars.   
## 4   
## Year-Month   
## 0   
## Monthly Nominal GDP Index (inMillion$)   
## 0   
## Monthly Real GDP Index (inMillion$)   
## 0   
## CPI   
## 0   
## PartyInPower   
## 0   
## unemployment rate   
## 0   
## CommercialBankInterestRateonCreditCardPlans   
## 0   
## Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan   
## 0   
## Earnings or wages in dollars per hour   
## 0   
## AdvertisingExpenses (in Thousand Dollars)   
## 0   
## Cotton Monthly Price - US cents per Pound(lbs)   
## 0   
## Change(in%)   
## 0   
## Average upland planted(million acres)   
## 0   
## Average upland harvested(million acres)   
## 0   
## yieldperharvested acre   
## 0   
## Production (in 480-lb netweright in million bales)   
## 0   
## Mill use (in 480-lb netweright in million bales)   
## 0   
## Exports   
## 0

library(DMwR)  
Train\_df=centralImputation(Train\_df)  
  
##dropping variables  
Train\_df$PartyInPower=NULL  
Train\_df$`Year-Month`=NULL  
Train\_df$Year=NULL  
##converting factors  
Train\_df$Month=as.factor(Train\_df$Month)  
  
  
numerical\_val=c("Monthly Nominal GDP Index (inMillion$)","Monthly Real GDP Index (inMillion$)"  
 ,"CPI","unemployment rate","CommercialBankInterestRateonCreditCardPlans"  
 ,"Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan"  
 ,"Earnings or wages in dollars per hour"  
 ,"Cotton Monthly Price - US cents per Pound(lbs)"  
 ,"Change(in%)","Average upland planted(million acres)"  
 ,"Average upland harvested(million acres)","yieldperharvested acre"  
 ,"Production (in 480-lb netweright in million bales)"  
 ,"Exports")  
  
##Standerdizing Train data  
library(caret)

## Warning: package 'caret' was built under R version 3.4.2

## Loading required package: ggplot2

pre=preProcess(Train\_df[,numerical\_val],method = c("center", "scale"))  
Train\_df=predict(pre,Train\_df)  
names(Train\_df)

## [1] "Month"   
## [2] "Sales.In.ThousandDollars."   
## [3] "Monthly Nominal GDP Index (inMillion$)"   
## [4] "Monthly Real GDP Index (inMillion$)"   
## [5] "CPI"   
## [6] "unemployment rate"   
## [7] "CommercialBankInterestRateonCreditCardPlans"   
## [8] "Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan"  
## [9] "Earnings or wages in dollars per hour"   
## [10] "AdvertisingExpenses (in Thousand Dollars)"   
## [11] "Cotton Monthly Price - US cents per Pound(lbs)"   
## [12] "Change(in%)"   
## [13] "Average upland planted(million acres)"   
## [14] "Average upland harvested(million acres)"   
## [15] "yieldperharvested acre"   
## [16] "Production (in 480-lb netweright in million bales)"   
## [17] "Mill use (in 480-lb netweright in million bales)"   
## [18] "Exports"

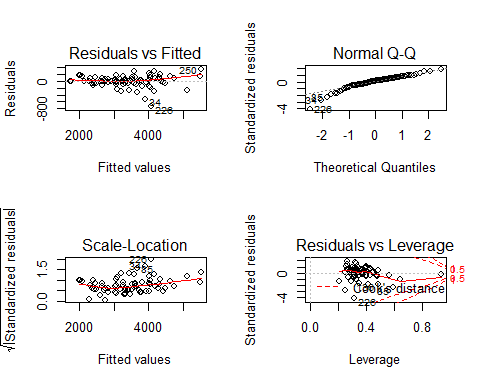
##applying linear model  
model\_basic=lm(Sales.In.ThousandDollars.~.,data=Train\_df)  
  
#model\_basic=lm(Sales.In.ThousandDollars.~.,data=Train\_df)  
  
##checking summary  
summary(model\_basic)

##   
## Call:  
## lm(formula = Sales.In.ThousandDollars. ~ ., data = Train\_df)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -728.91 -53.60 6.85 95.06 343.11   
##   
## Coefficients:  
## Estimate  
## (Intercept) 1076.045  
## Month2 471.788  
## Month3 1071.450  
## Month4 951.294  
## Month5 1277.509  
## Month6 797.111  
## Month7 622.199  
## Month8 836.206  
## Month9 889.834  
## Month10 1074.212  
## Month11 1223.746  
## Month12 2518.583  
## `Monthly Nominal GDP Index (inMillion$)` -1703.788  
## `Monthly Real GDP Index (inMillion$)` 768.606  
## CPI 1343.614  
## `unemployment rate` -673.497  
## CommercialBankInterestRateonCreditCardPlans 494.121  
## `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 94.093  
## `Earnings or wages in dollars per hour` 67.691  
## `AdvertisingExpenses (in Thousand Dollars)`116 430.551  
## `AdvertisingExpenses (in Thousand Dollars)`137 740.156  
## `AdvertisingExpenses (in Thousand Dollars)`150 428.644  
## `AdvertisingExpenses (in Thousand Dollars)`165 89.718  
## `AdvertisingExpenses (in Thousand Dollars)`183 76.635  
## `AdvertisingExpenses (in Thousand Dollars)`200 -135.546  
## `AdvertisingExpenses (in Thousand Dollars)`201 -4.204  
## `AdvertisingExpenses (in Thousand Dollars)`208 -69.897  
## `AdvertisingExpenses (in Thousand Dollars)`214 435.532  
## `AdvertisingExpenses (in Thousand Dollars)`221 9.093  
## `AdvertisingExpenses (in Thousand Dollars)`248 676.518  
## `Cotton Monthly Price - US cents per Pound(lbs)` -69.890  
## `Change(in%)` 9.653  
## `Average upland planted(million acres)` 143.301  
## `Average upland harvested(million acres)` -40.928  
## `yieldperharvested acre` 6.423  
## `Production (in 480-lb netweright in million bales)` 3.803  
## `Mill use (in 480-lb netweright in million bales)` 371.072  
## Exports 95.761  
## Std. Error  
## (Intercept) 1186.296  
## Month2 139.124  
## Month3 144.875  
## Month4 156.393  
## Month5 165.306  
## Month6 196.073  
## Month7 177.628  
## Month8 191.357  
## Month9 168.146  
## Month10 170.643  
## Month11 159.511  
## Month12 165.871  
## `Monthly Nominal GDP Index (inMillion$)` 1957.851  
## `Monthly Real GDP Index (inMillion$)` 1263.594  
## CPI 478.290  
## `unemployment rate` 346.649  
## CommercialBankInterestRateonCreditCardPlans 118.515  
## `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 65.897  
## `Earnings or wages in dollars per hour` 287.270  
## `AdvertisingExpenses (in Thousand Dollars)`116 286.752  
## `AdvertisingExpenses (in Thousand Dollars)`137 402.726  
## `AdvertisingExpenses (in Thousand Dollars)`150 288.808  
## `AdvertisingExpenses (in Thousand Dollars)`165 272.428  
## `AdvertisingExpenses (in Thousand Dollars)`183 253.672  
## `AdvertisingExpenses (in Thousand Dollars)`200 310.282  
## `AdvertisingExpenses (in Thousand Dollars)`201 284.659  
## `AdvertisingExpenses (in Thousand Dollars)`208 259.766  
## `AdvertisingExpenses (in Thousand Dollars)`214 297.670  
## `AdvertisingExpenses (in Thousand Dollars)`221 263.584  
## `AdvertisingExpenses (in Thousand Dollars)`248 264.526  
## `Cotton Monthly Price - US cents per Pound(lbs)` 66.169  
## `Change(in%)` 40.165  
## `Average upland planted(million acres)` 74.332  
## `Average upland harvested(million acres)` 320.924  
## `yieldperharvested acre` 81.985  
## `Production (in 480-lb netweright in million bales)` 327.067  
## `Mill use (in 480-lb netweright in million bales)` 332.075  
## Exports 76.603  
## t value  
## (Intercept) 0.907  
## Month2 3.391  
## Month3 7.396  
## Month4 6.083  
## Month5 7.728  
## Month6 4.065  
## Month7 3.503  
## Month8 4.370  
## Month9 5.292  
## Month10 6.295  
## Month11 7.672  
## Month12 15.184  
## `Monthly Nominal GDP Index (inMillion$)` -0.870  
## `Monthly Real GDP Index (inMillion$)` 0.608  
## CPI 2.809  
## `unemployment rate` -1.943  
## CommercialBankInterestRateonCreditCardPlans 4.169  
## `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 1.428  
## `Earnings or wages in dollars per hour` 0.236  
## `AdvertisingExpenses (in Thousand Dollars)`116 1.501  
## `AdvertisingExpenses (in Thousand Dollars)`137 1.838  
## `AdvertisingExpenses (in Thousand Dollars)`150 1.484  
## `AdvertisingExpenses (in Thousand Dollars)`165 0.329  
## `AdvertisingExpenses (in Thousand Dollars)`183 0.302  
## `AdvertisingExpenses (in Thousand Dollars)`200 -0.437  
## `AdvertisingExpenses (in Thousand Dollars)`201 -0.015  
## `AdvertisingExpenses (in Thousand Dollars)`208 -0.269  
## `AdvertisingExpenses (in Thousand Dollars)`214 1.463  
## `AdvertisingExpenses (in Thousand Dollars)`221 0.034  
## `AdvertisingExpenses (in Thousand Dollars)`248 2.557  
## `Cotton Monthly Price - US cents per Pound(lbs)` -1.056  
## `Change(in%)` 0.240  
## `Average upland planted(million acres)` 1.928  
## `Average upland harvested(million acres)` -0.128  
## `yieldperharvested acre` 0.078  
## `Production (in 480-lb netweright in million bales)` 0.012  
## `Mill use (in 480-lb netweright in million bales)` 1.117  
## Exports 1.250  
## Pr(>|t|)  
## (Intercept) 0.369101  
## Month2 0.001439  
## Month3 2.33e-09  
## Month4 2.18e-07  
## Month5 7.46e-10  
## Month6 0.000186  
## Month7 0.001037  
## Month8 7.02e-05  
## Month9 3.28e-06  
## Month10 1.04e-07  
## Month11 9.05e-10  
## Month12 < 2e-16  
## `Monthly Nominal GDP Index (inMillion$)` 0.388691  
## `Monthly Real GDP Index (inMillion$)` 0.546001  
## CPI 0.007267  
## `unemployment rate` 0.058166  
## CommercialBankInterestRateonCreditCardPlans 0.000134  
## `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 0.160081  
## `Earnings or wages in dollars per hour` 0.814762  
## `AdvertisingExpenses (in Thousand Dollars)`116 0.140066  
## `AdvertisingExpenses (in Thousand Dollars)`137 0.072545  
## `AdvertisingExpenses (in Thousand Dollars)`150 0.144580  
## `AdvertisingExpenses (in Thousand Dollars)`165 0.743402  
## `AdvertisingExpenses (in Thousand Dollars)`183 0.763936  
## `AdvertisingExpenses (in Thousand Dollars)`200 0.664266  
## `AdvertisingExpenses (in Thousand Dollars)`201 0.988279  
## `AdvertisingExpenses (in Thousand Dollars)`208 0.789073  
## `AdvertisingExpenses (in Thousand Dollars)`214 0.150229  
## `AdvertisingExpenses (in Thousand Dollars)`221 0.972630  
## `AdvertisingExpenses (in Thousand Dollars)`248 0.013905  
## `Cotton Monthly Price - US cents per Pound(lbs)` 0.296380  
## `Change(in%)` 0.811140  
## `Average upland planted(million acres)` 0.060061  
## `Average upland harvested(million acres)` 0.899075  
## `yieldperharvested acre` 0.937896  
## `Production (in 480-lb netweright in million bales)` 0.990772  
## `Mill use (in 480-lb netweright in million bales)` 0.269611  
## Exports 0.217585  
##   
## (Intercept)   
## Month2 \*\*   
## Month3 \*\*\*  
## Month4 \*\*\*  
## Month5 \*\*\*  
## Month6 \*\*\*  
## Month7 \*\*   
## Month8 \*\*\*  
## Month9 \*\*\*  
## Month10 \*\*\*  
## Month11 \*\*\*  
## Month12 \*\*\*  
## `Monthly Nominal GDP Index (inMillion$)`   
## `Monthly Real GDP Index (inMillion$)`   
## CPI \*\*   
## `unemployment rate` .   
## CommercialBankInterestRateonCreditCardPlans \*\*\*  
## `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan`   
## `Earnings or wages in dollars per hour`   
## `AdvertisingExpenses (in Thousand Dollars)`116   
## `AdvertisingExpenses (in Thousand Dollars)`137 .   
## `AdvertisingExpenses (in Thousand Dollars)`150   
## `AdvertisingExpenses (in Thousand Dollars)`165   
## `AdvertisingExpenses (in Thousand Dollars)`183   
## `AdvertisingExpenses (in Thousand Dollars)`200   
## `AdvertisingExpenses (in Thousand Dollars)`201   
## `AdvertisingExpenses (in Thousand Dollars)`208   
## `AdvertisingExpenses (in Thousand Dollars)`214   
## `AdvertisingExpenses (in Thousand Dollars)`221   
## `AdvertisingExpenses (in Thousand Dollars)`248 \*   
## `Cotton Monthly Price - US cents per Pound(lbs)`   
## `Change(in%)`   
## `Average upland planted(million acres)` .   
## `Average upland harvested(million acres)`   
## `yieldperharvested acre`   
## `Production (in 480-lb netweright in million bales)`   
## `Mill use (in 480-lb netweright in million bales)`   
## Exports   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 218.9 on 46 degrees of freedom  
## Multiple R-squared: 0.9607, Adjusted R-squared: 0.9291   
## F-statistic: 30.41 on 37 and 46 DF, p-value: < 2.2e-16

##plotting model  
par(mfrow = c(2,2))  
plot(model\_basic)

## Warning: not plotting observations with leverage one:  
## 1, 2, 4, 13, 16, 21, 23, 30, 42, 44, 82

## Warning: not plotting observations with leverage one:  
## 1, 2, 4, 13, 16, 21, 23, 30, 42, 44, 82



##feature selection  
library(MASS)

##   
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':  
##   
## select

model\_aic=stepAIC(model\_basic)

## Start: AIC=930.67  
## Sales.In.ThousandDollars. ~ Month + `Monthly Nominal GDP Index (inMillion$)` +   
## `Monthly Real GDP Index (inMillion$)` + CPI + `unemployment rate` +   
## CommercialBankInterestRateonCreditCardPlans + `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` +   
## `Earnings or wages in dollars per hour` + `AdvertisingExpenses (in Thousand Dollars)` +   
## `Cotton Monthly Price - US cents per Pound(lbs)` + `Change(in%)` +   
## `Average upland planted(million acres)` + `Average upland harvested(million acres)` +   
## `yieldperharvested acre` + `Production (in 480-lb netweright in million bales)` +   
## `Mill use (in 480-lb netweright in million bales)` + Exports  
##   
## Df  
## - `Production (in 480-lb netweright in million bales)` 1  
## - `yieldperharvested acre` 1  
## - `Average upland harvested(million acres)` 1  
## - `Earnings or wages in dollars per hour` 1  
## - `Change(in%)` 1  
## - `Monthly Real GDP Index (inMillion$)` 1  
## - `Monthly Nominal GDP Index (inMillion$)` 1  
## <none>   
## - `Cotton Monthly Price - US cents per Pound(lbs)` 1  
## - `Mill use (in 480-lb netweright in million bales)` 1  
## - Exports 1  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 1  
## - `AdvertisingExpenses (in Thousand Dollars)` 11  
## - `Average upland planted(million acres)` 1  
## - `unemployment rate` 1  
## - CPI 1  
## - CommercialBankInterestRateonCreditCardPlans 1  
## - Month 11  
## Sum of Sq  
## - `Production (in 480-lb netweright in million bales)` 6  
## - `yieldperharvested acre` 294  
## - `Average upland harvested(million acres)` 779  
## - `Earnings or wages in dollars per hour` 2660  
## - `Change(in%)` 2767  
## - `Monthly Real GDP Index (inMillion$)` 17723  
## - `Monthly Nominal GDP Index (inMillion$)` 36275  
## <none>   
## - `Cotton Monthly Price - US cents per Pound(lbs)` 53438  
## - `Mill use (in 480-lb netweright in million bales)` 59810  
## - Exports 74855  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 97660  
## - `AdvertisingExpenses (in Thousand Dollars)` 745847  
## - `Average upland planted(million acres)` 178024  
## - `unemployment rate` 180811  
## - CPI 378007  
## - CommercialBankInterestRateonCreditCardPlans 832635  
## - Month 16645656  
## RSS  
## - `Production (in 480-lb netweright in million bales)` 2203393  
## - `yieldperharvested acre` 2203681  
## - `Average upland harvested(million acres)` 2204166  
## - `Earnings or wages in dollars per hour` 2206046  
## - `Change(in%)` 2206154  
## - `Monthly Real GDP Index (inMillion$)` 2221109  
## - `Monthly Nominal GDP Index (inMillion$)` 2239662  
## <none> 2203387  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 2256825  
## - `Mill use (in 480-lb netweright in million bales)` 2263197  
## - Exports 2278242  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 2301047  
## - `AdvertisingExpenses (in Thousand Dollars)` 2949234  
## - `Average upland planted(million acres)` 2381410  
## - `unemployment rate` 2384198  
## - CPI 2581393  
## - CommercialBankInterestRateonCreditCardPlans 3036022  
## - Month 18849043  
## AIC  
## - `Production (in 480-lb netweright in million bales)` 928.67  
## - `yieldperharvested acre` 928.69  
## - `Average upland harvested(million acres)` 928.70  
## - `Earnings or wages in dollars per hour` 928.78  
## - `Change(in%)` 928.78  
## - `Monthly Real GDP Index (inMillion$)` 929.35  
## - `Monthly Nominal GDP Index (inMillion$)` 930.05  
## <none> 930.67  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 930.69  
## - `Mill use (in 480-lb netweright in million bales)` 930.92  
## - Exports 931.48  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 932.32  
## - `AdvertisingExpenses (in Thousand Dollars)` 933.16  
## - `Average upland planted(million acres)` 935.20  
## - `unemployment rate` 935.30  
## - CPI 941.97  
## - CommercialBankInterestRateonCreditCardPlans 955.60  
## - Month 1088.98  
##   
## Step: AIC=928.67  
## Sales.In.ThousandDollars. ~ Month + `Monthly Nominal GDP Index (inMillion$)` +   
## `Monthly Real GDP Index (inMillion$)` + CPI + `unemployment rate` +   
## CommercialBankInterestRateonCreditCardPlans + `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` +   
## `Earnings or wages in dollars per hour` + `AdvertisingExpenses (in Thousand Dollars)` +   
## `Cotton Monthly Price - US cents per Pound(lbs)` + `Change(in%)` +   
## `Average upland planted(million acres)` + `Average upland harvested(million acres)` +   
## `yieldperharvested acre` + `Mill use (in 480-lb netweright in million bales)` +   
## Exports  
##   
## Df  
## - `yieldperharvested acre` 1  
## - `Earnings or wages in dollars per hour` 1  
## - `Change(in%)` 1  
## - `Average upland harvested(million acres)` 1  
## - `Monthly Real GDP Index (inMillion$)` 1  
## - `Monthly Nominal GDP Index (inMillion$)` 1  
## <none>   
## - `Cotton Monthly Price - US cents per Pound(lbs)` 1  
## - `Mill use (in 480-lb netweright in million bales)` 1  
## - Exports 1  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 1  
## - `AdvertisingExpenses (in Thousand Dollars)` 11  
## - `Average upland planted(million acres)` 1  
## - `unemployment rate` 1  
## - CPI 1  
## - CommercialBankInterestRateonCreditCardPlans 1  
## - Month 11  
## Sum of Sq  
## - `yieldperharvested acre` 1661  
## - `Earnings or wages in dollars per hour` 2712  
## - `Change(in%)` 2762  
## - `Average upland harvested(million acres)` 14323  
## - `Monthly Real GDP Index (inMillion$)` 17724  
## - `Monthly Nominal GDP Index (inMillion$)` 36327  
## <none>   
## - `Cotton Monthly Price - US cents per Pound(lbs)` 54411  
## - `Mill use (in 480-lb netweright in million bales)` 63089  
## - Exports 74914  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 108935  
## - `AdvertisingExpenses (in Thousand Dollars)` 745881  
## - `Average upland planted(million acres)` 182204  
## - `unemployment rate` 182826  
## - CPI 379229  
## - CommercialBankInterestRateonCreditCardPlans 836538  
## - Month 17018541  
## RSS  
## - `yieldperharvested acre` 2205054  
## - `Earnings or wages in dollars per hour` 2206106  
## - `Change(in%)` 2206155  
## - `Average upland harvested(million acres)` 2217716  
## - `Monthly Real GDP Index (inMillion$)` 2221117  
## - `Monthly Nominal GDP Index (inMillion$)` 2239720  
## <none> 2203393  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 2257804  
## - `Mill use (in 480-lb netweright in million bales)` 2266482  
## - Exports 2278308  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 2312329  
## - `AdvertisingExpenses (in Thousand Dollars)` 2949274  
## - `Average upland planted(million acres)` 2385597  
## - `unemployment rate` 2386219  
## - CPI 2582622  
## - CommercialBankInterestRateonCreditCardPlans 3039931  
## - Month 19221934  
## AIC  
## - `yieldperharvested acre` 926.74  
## - `Earnings or wages in dollars per hour` 926.78  
## - `Change(in%)` 926.78  
## - `Average upland harvested(million acres)` 927.22  
## - `Monthly Real GDP Index (inMillion$)` 927.35  
## - `Monthly Nominal GDP Index (inMillion$)` 928.05  
## <none> 928.67  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 928.72  
## - `Mill use (in 480-lb netweright in million bales)` 929.05  
## - Exports 929.48  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 930.73  
## - `AdvertisingExpenses (in Thousand Dollars)` 931.17  
## - `Average upland planted(million acres)` 933.35  
## - `unemployment rate` 933.37  
## - CPI 940.01  
## - CommercialBankInterestRateonCreditCardPlans 953.71  
## - Month 1088.62  
##   
## Step: AIC=926.74  
## Sales.In.ThousandDollars. ~ Month + `Monthly Nominal GDP Index (inMillion$)` +   
## `Monthly Real GDP Index (inMillion$)` + CPI + `unemployment rate` +   
## CommercialBankInterestRateonCreditCardPlans + `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` +   
## `Earnings or wages in dollars per hour` + `AdvertisingExpenses (in Thousand Dollars)` +   
## `Cotton Monthly Price - US cents per Pound(lbs)` + `Change(in%)` +   
## `Average upland planted(million acres)` + `Average upland harvested(million acres)` +   
## `Mill use (in 480-lb netweright in million bales)` + Exports  
##   
## Df  
## - `Earnings or wages in dollars per hour` 1  
## - `Change(in%)` 1  
## - `Average upland harvested(million acres)` 1  
## - `Monthly Real GDP Index (inMillion$)` 1  
## - `Monthly Nominal GDP Index (inMillion$)` 1  
## <none>   
## - `Cotton Monthly Price - US cents per Pound(lbs)` 1  
## - `Mill use (in 480-lb netweright in million bales)` 1  
## - Exports 1  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 1  
## - `AdvertisingExpenses (in Thousand Dollars)` 11  
## - `unemployment rate` 1  
## - `Average upland planted(million acres)` 1  
## - CPI 1  
## - CommercialBankInterestRateonCreditCardPlans 1  
## - Month 11  
## Sum of Sq  
## - `Earnings or wages in dollars per hour` 2172  
## - `Change(in%)` 3624  
## - `Average upland harvested(million acres)` 12826  
## - `Monthly Real GDP Index (inMillion$)` 17066  
## - `Monthly Nominal GDP Index (inMillion$)` 35736  
## <none>   
## - `Cotton Monthly Price - US cents per Pound(lbs)` 66186  
## - `Mill use (in 480-lb netweright in million bales)` 70978  
## - Exports 92065  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 111794  
## - `AdvertisingExpenses (in Thousand Dollars)` 767704  
## - `unemployment rate` 181165  
## - `Average upland planted(million acres)` 188544  
## - CPI 389674  
## - CommercialBankInterestRateonCreditCardPlans 836910  
## - Month 17563543  
## RSS  
## - `Earnings or wages in dollars per hour` 2207226  
## - `Change(in%)` 2208678  
## - `Average upland harvested(million acres)` 2217881  
## - `Monthly Real GDP Index (inMillion$)` 2222121  
## - `Monthly Nominal GDP Index (inMillion$)` 2240791  
## <none> 2205054  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 2271240  
## - `Mill use (in 480-lb netweright in million bales)` 2276032  
## - Exports 2297120  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 2316849  
## - `AdvertisingExpenses (in Thousand Dollars)` 2972759  
## - `unemployment rate` 2386219  
## - `Average upland planted(million acres)` 2393598  
## - CPI 2594728  
## - CommercialBankInterestRateonCreditCardPlans 3041964  
## - Month 19768597  
## AIC  
## - `Earnings or wages in dollars per hour` 924.82  
## - `Change(in%)` 924.88  
## - `Average upland harvested(million acres)` 925.22  
## - `Monthly Real GDP Index (inMillion$)` 925.39  
## - `Monthly Nominal GDP Index (inMillion$)` 926.09  
## <none> 926.74  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 927.22  
## - `Mill use (in 480-lb netweright in million bales)` 927.40  
## - Exports 928.17  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 928.89  
## - `AdvertisingExpenses (in Thousand Dollars)` 929.83  
## - `unemployment rate` 931.37  
## - `Average upland planted(million acres)` 931.63  
## - CPI 938.41  
## - CommercialBankInterestRateonCreditCardPlans 951.76  
## - Month 1088.98  
##   
## Step: AIC=924.82  
## Sales.In.ThousandDollars. ~ Month + `Monthly Nominal GDP Index (inMillion$)` +   
## `Monthly Real GDP Index (inMillion$)` + CPI + `unemployment rate` +   
## CommercialBankInterestRateonCreditCardPlans + `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` +   
## `AdvertisingExpenses (in Thousand Dollars)` + `Cotton Monthly Price - US cents per Pound(lbs)` +   
## `Change(in%)` + `Average upland planted(million acres)` +   
## `Average upland harvested(million acres)` + `Mill use (in 480-lb netweright in million bales)` +   
## Exports  
##   
## Df  
## - `Change(in%)` 1  
## - `Average upland harvested(million acres)` 1  
## - `Monthly Real GDP Index (inMillion$)` 1  
## - `Monthly Nominal GDP Index (inMillion$)` 1  
## <none>   
## - `Cotton Monthly Price - US cents per Pound(lbs)` 1  
## - `Mill use (in 480-lb netweright in million bales)` 1  
## - Exports 1  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 1  
## - `AdvertisingExpenses (in Thousand Dollars)` 11  
## - `unemployment rate` 1  
## - `Average upland planted(million acres)` 1  
## - CPI 1  
## - CommercialBankInterestRateonCreditCardPlans 1  
## - Month 11  
## Sum of Sq  
## - `Change(in%)` 3842  
## - `Average upland harvested(million acres)` 12796  
## - `Monthly Real GDP Index (inMillion$)` 15682  
## - `Monthly Nominal GDP Index (inMillion$)` 33565  
## <none>   
## - `Cotton Monthly Price - US cents per Pound(lbs)` 64148  
## - `Mill use (in 480-lb netweright in million bales)` 69015  
## - Exports 90086  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 109706  
## - `AdvertisingExpenses (in Thousand Dollars)` 791625  
## - `unemployment rate` 179548  
## - `Average upland planted(million acres)` 186389  
## - CPI 387607  
## - CommercialBankInterestRateonCreditCardPlans 835511  
## - Month 18078277  
## RSS  
## - `Change(in%)` 2211068  
## - `Average upland harvested(million acres)` 2220022  
## - `Monthly Real GDP Index (inMillion$)` 2222908  
## - `Monthly Nominal GDP Index (inMillion$)` 2240791  
## <none> 2207226  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 2271375  
## - `Mill use (in 480-lb netweright in million bales)` 2276241  
## - Exports 2297312  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 2316932  
## - `AdvertisingExpenses (in Thousand Dollars)` 2998852  
## - `unemployment rate` 2386775  
## - `Average upland planted(million acres)` 2393615  
## - CPI 2594834  
## - CommercialBankInterestRateonCreditCardPlans 3042737  
## - Month 20285504  
## AIC  
## - `Change(in%)` 922.97  
## - `Average upland harvested(million acres)` 923.31  
## - `Monthly Real GDP Index (inMillion$)` 923.41  
## - `Monthly Nominal GDP Index (inMillion$)` 924.09  
## <none> 924.82  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 925.23  
## - `Mill use (in 480-lb netweright in million bales)` 925.41  
## - Exports 926.18  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 926.89  
## - `AdvertisingExpenses (in Thousand Dollars)` 928.57  
## - `unemployment rate` 929.39  
## - `Average upland planted(million acres)` 929.63  
## - CPI 936.41  
## - CommercialBankInterestRateonCreditCardPlans 949.79  
## - Month 1089.15  
##   
## Step: AIC=922.97  
## Sales.In.ThousandDollars. ~ Month + `Monthly Nominal GDP Index (inMillion$)` +   
## `Monthly Real GDP Index (inMillion$)` + CPI + `unemployment rate` +   
## CommercialBankInterestRateonCreditCardPlans + `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` +   
## `AdvertisingExpenses (in Thousand Dollars)` + `Cotton Monthly Price - US cents per Pound(lbs)` +   
## `Average upland planted(million acres)` + `Average upland harvested(million acres)` +   
## `Mill use (in 480-lb netweright in million bales)` + Exports  
##   
## Df  
## - `Monthly Real GDP Index (inMillion$)` 1  
## - `Average upland harvested(million acres)` 1  
## - `Monthly Nominal GDP Index (inMillion$)` 1  
## <none>   
## - `Cotton Monthly Price - US cents per Pound(lbs)` 1  
## - `Mill use (in 480-lb netweright in million bales)` 1  
## - Exports 1  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 1  
## - `unemployment rate` 1  
## - `AdvertisingExpenses (in Thousand Dollars)` 11  
## - `Average upland planted(million acres)` 1  
## - CPI 1  
## - CommercialBankInterestRateonCreditCardPlans 1  
## - Month 11  
## Sum of Sq  
## - `Monthly Real GDP Index (inMillion$)` 12552  
## - `Average upland harvested(million acres)` 14760  
## - `Monthly Nominal GDP Index (inMillion$)` 29800  
## <none>   
## - `Cotton Monthly Price - US cents per Pound(lbs)` 60383  
## - `Mill use (in 480-lb netweright in million bales)` 69235  
## - Exports 92031  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 106063  
## - `unemployment rate` 175780  
## - `AdvertisingExpenses (in Thousand Dollars)` 824130  
## - `Average upland planted(million acres)` 193137  
## - CPI 405585  
## - CommercialBankInterestRateonCreditCardPlans 841523  
## - Month 19077350  
## RSS  
## - `Monthly Real GDP Index (inMillion$)` 2223620  
## - `Average upland harvested(million acres)` 2225828  
## - `Monthly Nominal GDP Index (inMillion$)` 2240868  
## <none> 2211068  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 2271451  
## - `Mill use (in 480-lb netweright in million bales)` 2280303  
## - Exports 2303099  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 2317132  
## - `unemployment rate` 2386848  
## - `AdvertisingExpenses (in Thousand Dollars)` 3035198  
## - `Average upland planted(million acres)` 2404205  
## - CPI 2616653  
## - CommercialBankInterestRateonCreditCardPlans 3052591  
## - Month 21288418  
## AIC  
## - `Monthly Real GDP Index (inMillion$)` 921.44  
## - `Average upland harvested(million acres)` 921.53  
## - `Monthly Nominal GDP Index (inMillion$)` 922.09  
## <none> 922.97  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 923.23  
## - `Mill use (in 480-lb netweright in million bales)` 923.56  
## - Exports 924.39  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 924.90  
## - `unemployment rate` 927.39  
## - `AdvertisingExpenses (in Thousand Dollars)` 927.58  
## - `Average upland planted(million acres)` 928.00  
## - CPI 935.11  
## - CommercialBankInterestRateonCreditCardPlans 948.06  
## - Month 1091.20  
##   
## Step: AIC=921.44  
## Sales.In.ThousandDollars. ~ Month + `Monthly Nominal GDP Index (inMillion$)` +   
## CPI + `unemployment rate` + CommercialBankInterestRateonCreditCardPlans +   
## `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` +   
## `AdvertisingExpenses (in Thousand Dollars)` + `Cotton Monthly Price - US cents per Pound(lbs)` +   
## `Average upland planted(million acres)` + `Average upland harvested(million acres)` +   
## `Mill use (in 480-lb netweright in million bales)` + Exports  
##   
## Df  
## - `Average upland harvested(million acres)` 1  
## <none>   
## - `Cotton Monthly Price - US cents per Pound(lbs)` 1  
## - `Mill use (in 480-lb netweright in million bales)` 1  
## - Exports 1  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 1  
## - `Monthly Nominal GDP Index (inMillion$)` 1  
## - `AdvertisingExpenses (in Thousand Dollars)` 11  
## - `unemployment rate` 1  
## - `Average upland planted(million acres)` 1  
## - CommercialBankInterestRateonCreditCardPlans 1  
## - CPI 1  
## - Month 11  
## Sum of Sq  
## - `Average upland harvested(million acres)` 18647  
## <none>   
## - `Cotton Monthly Price - US cents per Pound(lbs)` 54856  
## - `Mill use (in 480-lb netweright in million bales)` 60829  
## - Exports 87733  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 106947  
## - `Monthly Nominal GDP Index (inMillion$)` 124956  
## - `AdvertisingExpenses (in Thousand Dollars)` 811742  
## - `unemployment rate` 230154  
## - `Average upland planted(million acres)` 243986  
## - CommercialBankInterestRateonCreditCardPlans 848061  
## - CPI 929981  
## - Month 21232513  
## RSS  
## - `Average upland harvested(million acres)` 2242267  
## <none> 2223620  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 2278476  
## - `Mill use (in 480-lb netweright in million bales)` 2284449  
## - Exports 2311353  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 2330567  
## - `Monthly Nominal GDP Index (inMillion$)` 2348576  
## - `AdvertisingExpenses (in Thousand Dollars)` 3035362  
## - `unemployment rate` 2453774  
## - `Average upland planted(million acres)` 2467606  
## - CommercialBankInterestRateonCreditCardPlans 3071681  
## - CPI 3153601  
## - Month 23456133  
## AIC  
## - `Average upland harvested(million acres)` 920.14  
## <none> 921.44  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 921.49  
## - `Mill use (in 480-lb netweright in million bales)` 921.71  
## - Exports 922.69  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 923.39  
## - `Monthly Nominal GDP Index (inMillion$)` 924.03  
## - `AdvertisingExpenses (in Thousand Dollars)` 925.58  
## - `unemployment rate` 927.71  
## - `Average upland planted(million acres)` 928.19  
## - CommercialBankInterestRateonCreditCardPlans 946.58  
## - CPI 948.79  
## - Month 1097.35  
##   
## Step: AIC=920.14  
## Sales.In.ThousandDollars. ~ Month + `Monthly Nominal GDP Index (inMillion$)` +   
## CPI + `unemployment rate` + CommercialBankInterestRateonCreditCardPlans +   
## `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` +   
## `AdvertisingExpenses (in Thousand Dollars)` + `Cotton Monthly Price - US cents per Pound(lbs)` +   
## `Average upland planted(million acres)` + `Mill use (in 480-lb netweright in million bales)` +   
## Exports  
##   
## Df  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 1  
## <none>   
## - Exports 1  
## - `Mill use (in 480-lb netweright in million bales)` 1  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 1  
## - `Monthly Nominal GDP Index (inMillion$)` 1  
## - `AdvertisingExpenses (in Thousand Dollars)` 11  
## - `unemployment rate` 1  
## - `Average upland planted(million acres)` 1  
## - CommercialBankInterestRateonCreditCardPlans 1  
## - CPI 1  
## - Month 11  
## Sum of Sq  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 52960  
## <none>   
## - Exports 74808  
## - `Mill use (in 480-lb netweright in million bales)` 77240  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 88602  
## - `Monthly Nominal GDP Index (inMillion$)` 131448  
## - `AdvertisingExpenses (in Thousand Dollars)` 833432  
## - `unemployment rate` 211507  
## - `Average upland planted(million acres)` 338814  
## - CommercialBankInterestRateonCreditCardPlans 833492  
## - CPI 956514  
## - Month 21233080  
## RSS  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 2295228  
## <none> 2242267  
## - Exports 2317076  
## - `Mill use (in 480-lb netweright in million bales)` 2319507  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 2330869  
## - `Monthly Nominal GDP Index (inMillion$)` 2373716  
## - `AdvertisingExpenses (in Thousand Dollars)` 3075700  
## - `unemployment rate` 2453774  
## - `Average upland planted(million acres)` 2581082  
## - CommercialBankInterestRateonCreditCardPlans 3075759  
## - CPI 3198782  
## - Month 23475348  
## AIC  
## - `Cotton Monthly Price - US cents per Pound(lbs)` 920.10  
## <none> 920.14  
## - Exports 920.90  
## - `Mill use (in 480-lb netweright in million bales)` 920.99  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 921.40  
## - `Monthly Nominal GDP Index (inMillion$)` 922.93  
## - `AdvertisingExpenses (in Thousand Dollars)` 924.69  
## - `unemployment rate` 925.71  
## - `Average upland planted(million acres)` 929.96  
## - CommercialBankInterestRateonCreditCardPlans 944.69  
## - CPI 947.99  
## - Month 1095.41  
##   
## Step: AIC=920.1  
## Sales.In.ThousandDollars. ~ Month + `Monthly Nominal GDP Index (inMillion$)` +   
## CPI + `unemployment rate` + CommercialBankInterestRateonCreditCardPlans +   
## `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` +   
## `AdvertisingExpenses (in Thousand Dollars)` + `Average upland planted(million acres)` +   
## `Mill use (in 480-lb netweright in million bales)` + Exports  
##   
## Df  
## - Exports 1  
## - `Mill use (in 480-lb netweright in million bales)` 1  
## <none>   
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 1  
## - `Monthly Nominal GDP Index (inMillion$)` 1  
## - `AdvertisingExpenses (in Thousand Dollars)` 11  
## - `unemployment rate` 1  
## - `Average upland planted(million acres)` 1  
## - CommercialBankInterestRateonCreditCardPlans 1  
## - CPI 1  
## - Month 11  
## Sum of Sq  
## - Exports 23073  
## - `Mill use (in 480-lb netweright in million bales)` 32509  
## <none>   
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 71721  
## - `Monthly Nominal GDP Index (inMillion$)` 101689  
## - `AdvertisingExpenses (in Thousand Dollars)` 863487  
## - `unemployment rate` 232259  
## - `Average upland planted(million acres)` 333959  
## - CommercialBankInterestRateonCreditCardPlans 790378  
## - CPI 1014930  
## - Month 21328675  
## RSS  
## - Exports 2318301  
## - `Mill use (in 480-lb netweright in million bales)` 2327737  
## <none> 2295228  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 2366948  
## - `Monthly Nominal GDP Index (inMillion$)` 2396916  
## - `AdvertisingExpenses (in Thousand Dollars)` 3158714  
## - `unemployment rate` 2527486  
## - `Average upland planted(million acres)` 2629187  
## - CommercialBankInterestRateonCreditCardPlans 3085605  
## - CPI 3310158  
## - Month 23623902  
## AIC  
## - Exports 918.94  
## - `Mill use (in 480-lb netweright in million bales)` 919.29  
## <none> 920.10  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 920.69  
## - `Monthly Nominal GDP Index (inMillion$)` 921.75  
## - `AdvertisingExpenses (in Thousand Dollars)` 924.93  
## - `unemployment rate` 926.20  
## - `Average upland planted(million acres)` 929.51  
## - CommercialBankInterestRateonCreditCardPlans 942.96  
## - CPI 948.86  
## - Month 1093.94  
##   
## Step: AIC=918.94  
## Sales.In.ThousandDollars. ~ Month + `Monthly Nominal GDP Index (inMillion$)` +   
## CPI + `unemployment rate` + CommercialBankInterestRateonCreditCardPlans +   
## `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` +   
## `AdvertisingExpenses (in Thousand Dollars)` + `Average upland planted(million acres)` +   
## `Mill use (in 480-lb netweright in million bales)`  
##   
## Df  
## - `Mill use (in 480-lb netweright in million bales)` 1  
## <none>   
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 1  
## - `Monthly Nominal GDP Index (inMillion$)` 1  
## - `AdvertisingExpenses (in Thousand Dollars)` 11  
## - `unemployment rate` 1  
## - `Average upland planted(million acres)` 1  
## - CommercialBankInterestRateonCreditCardPlans 1  
## - CPI 1  
## - Month 11  
## Sum of Sq  
## - `Mill use (in 480-lb netweright in million bales)` 39337  
## <none>   
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 73631  
## - `Monthly Nominal GDP Index (inMillion$)` 88110  
## - `AdvertisingExpenses (in Thousand Dollars)` 870271  
## - `unemployment rate` 209880  
## - `Average upland planted(million acres)` 334132  
## - CommercialBankInterestRateonCreditCardPlans 829516  
## - CPI 1011599  
## - Month 21309111  
## RSS  
## - `Mill use (in 480-lb netweright in million bales)` 2357637  
## <none> 2318301  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 2391931  
## - `Monthly Nominal GDP Index (inMillion$)` 2406410  
## - `AdvertisingExpenses (in Thousand Dollars)` 3188572  
## - `unemployment rate` 2528180  
## - `Average upland planted(million acres)` 2652432  
## - CommercialBankInterestRateonCreditCardPlans 3147816  
## - CPI 3329899  
## - Month 23627411  
## AIC  
## - `Mill use (in 480-lb netweright in million bales)` 918.36  
## <none> 918.94  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 919.57  
## - `Monthly Nominal GDP Index (inMillion$)` 920.08  
## - `AdvertisingExpenses (in Thousand Dollars)` 923.72  
## - `unemployment rate` 924.22  
## - `Average upland planted(million acres)` 928.25  
## - CommercialBankInterestRateonCreditCardPlans 942.64  
## - CPI 947.36  
## - Month 1091.96  
##   
## Step: AIC=918.36  
## Sales.In.ThousandDollars. ~ Month + `Monthly Nominal GDP Index (inMillion$)` +   
## CPI + `unemployment rate` + CommercialBankInterestRateonCreditCardPlans +   
## `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` +   
## `AdvertisingExpenses (in Thousand Dollars)` + `Average upland planted(million acres)`  
##   
## Df  
## <none>   
## - `Monthly Nominal GDP Index (inMillion$)` 1  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 1  
## - `unemployment rate` 1  
## - `AdvertisingExpenses (in Thousand Dollars)` 11  
## - `Average upland planted(million acres)` 1  
## - CommercialBankInterestRateonCreditCardPlans 1  
## - CPI 1  
## - Month 11  
## Sum of Sq  
## <none>   
## - `Monthly Nominal GDP Index (inMillion$)` 99568  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 102023  
## - `unemployment rate` 291616  
## - `AdvertisingExpenses (in Thousand Dollars)` 1009857  
## - `Average upland planted(million acres)` 447945  
## - CommercialBankInterestRateonCreditCardPlans 940485  
## - CPI 1021298  
## - Month 21275163  
## RSS  
## <none> 2357637  
## - `Monthly Nominal GDP Index (inMillion$)` 2457206  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 2459660  
## - `unemployment rate` 2649253  
## - `AdvertisingExpenses (in Thousand Dollars)` 3367495  
## - `Average upland planted(million acres)` 2805583  
## - CommercialBankInterestRateonCreditCardPlans 3298122  
## - CPI 3378936  
## - Month 23632800  
## AIC  
## <none> 918.36  
## - `Monthly Nominal GDP Index (inMillion$)` 919.83  
## - `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan` 919.92  
## - `unemployment rate` 926.15  
## - `AdvertisingExpenses (in Thousand Dollars)` 926.30  
## - `Average upland planted(million acres)` 930.97  
## - CommercialBankInterestRateonCreditCardPlans 944.56  
## - CPI 946.59  
## - Month 1089.98

model\_aic$coefficients

## (Intercept)   
## 2411.517944   
## Month2   
## 444.333007   
## Month3   
## 1069.757690   
## Month4   
## 946.802672   
## Month5   
## 1291.504654   
## Month6   
## 762.806034   
## Month7   
## 605.000230   
## Month8   
## 863.613192   
## Month9   
## 874.840144   
## Month10   
## 1045.525840   
## Month11   
## 1171.136262   
## Month12   
## 2434.984787   
## `Monthly Nominal GDP Index (inMillion$)`   
## -438.464944   
## CPI   
## 990.705858   
## `unemployment rate`   
## -523.991070   
## CommercialBankInterestRateonCreditCardPlans   
## 463.497672   
## `Finance Rate on Personal Loans at Commercial Banks, 24 Month Loan`   
## 72.048085   
## `AdvertisingExpenses (in Thousand Dollars)`116   
## 433.971248   
## `AdvertisingExpenses (in Thousand Dollars)`137   
## 972.963675   
## `AdvertisingExpenses (in Thousand Dollars)`150   
## 404.218946   
## `AdvertisingExpenses (in Thousand Dollars)`165   
## 214.190974   
## `AdvertisingExpenses (in Thousand Dollars)`183   
## 68.240311   
## `AdvertisingExpenses (in Thousand Dollars)`200   
## 4.223114   
## `AdvertisingExpenses (in Thousand Dollars)`201   
## -158.259002   
## `AdvertisingExpenses (in Thousand Dollars)`208   
## -63.705960   
## `AdvertisingExpenses (in Thousand Dollars)`214   
## 544.077578   
## `AdvertisingExpenses (in Thousand Dollars)`221   
## 45.573792   
## `AdvertisingExpenses (in Thousand Dollars)`248   
## 579.866430   
## `Average upland planted(million acres)`   
## 130.815994

Test=read.csv('template.csv',header = TRUE,sep=",")  
Women\_Test=Test[Test$ProductCategory=='WomenClothing',]  
Women\_Test$ProductCategory=NULL  
Women\_Test$Year=NULL  
  
##CONVERTING TO FACTOR  
Women\_Test$Month=as.factor(Women\_Test$Month)  
  
str(Women\_Test)

## 'data.frame': 12 obs. of 2 variables:  
## $ Month : Factor w/ 12 levels "1","2","3","4",..: 1 2 3 4 5 6 7 8 9 10 ...  
## $ target: int 119 151 72 78 130 190 200 166 62 132 ...

##binding macroeconmic data with testdata  
Test\_df=cbind(Women\_Test,Test\_MacroEconomics)  
  
##taking mean and sd from train and applying on test for standardization  
Test\_df=predict(pre,Test\_df)  
  
##prediction on test data  
pred\_model=predict(model\_basic,Test\_df)  
##Error Metrics  
library(DMwR)  
#regr.eval(Test\_df, pred\_model)  
  
  
  
##prediction on test data for step aic  
pred\_aic=predict(model\_aic,Test\_df)  
##Error Metrics  
library(DMwR)  
regr.eval(Test\_df$Sales.In.ThousandDollars., pred\_aic)

## mae mse rmse mape   
## NaN NaN NaN NaN

#library(DAAG)  
#cvResults <- suppressWarnings(CVlm(df=Train\_df, form.lm=Sales.In.ThousandDollars.~, m=5, dots=FALSE, seed=29, legend.pos="topleft", printit=FALSE, main="Small symbols are predicted values while bigger ones are actuals."))  
#attr(cvResults, 'ms')  
  
  
  
  
  
library(rpart)  
rpart(Sales.In.ThousandDollars.~.,data =Train\_df)->rf  
summary(rf)

## Call:  
## rpart(formula = Sales.In.ThousandDollars. ~ ., data = Train\_df)  
## n= 84   
##   
## CP nsplit rel error xerror xstd  
## 1 0.39714529 0 1.0000000 1.0248973 0.16785292  
## 2 0.12226542 1 0.6028547 0.8500280 0.14583938  
## 3 0.12131551 2 0.4805893 0.8098735 0.13464322  
## 4 0.05118376 3 0.3592738 0.6174753 0.09798942  
## 5 0.04710283 4 0.3080900 0.5398058 0.08449735  
## 6 0.02808224 5 0.2609872 0.5138491 0.07879263  
## 7 0.01945655 6 0.2329049 0.5133193 0.08129346  
## 8 0.01000000 7 0.2134484 0.5086439 0.08144212  
##   
## Variable importance  
## Monthly Nominal GDP Index (inMillion$)   
## 16   
## Monthly Real GDP Index (inMillion$)   
## 16   
## CPI   
## 15   
## Earnings or wages in dollars per hour   
## 15   
## CommercialBankInterestRateonCreditCardPlans   
## 12   
## unemployment rate   
## 11   
## Month   
## 7   
## Average upland planted(million acres)   
## 4   
## Production (in 480-lb netweright in million bales)   
## 1   
## yieldperharvested acre   
## 1   
## Change(in%)   
## 1   
##   
## Node number 1: 84 observations, complexity param=0.3971453  
## mean=3406.964, MSE=667899.6   
## left son=2 (48 obs) right son=3 (36 obs)  
## Primary splits:  
## Monthly Real GDP Index (inMillion$) < 0.0900057 to the left, improve=0.3971453, (0 missing)  
## Monthly Nominal GDP Index (inMillion$) < -0.8068302 to the left, improve=0.3970211, (0 missing)  
## unemployment rate < -0.1040174 to the right, improve=0.3913350, (0 missing)  
## Earnings or wages in dollars per hour < 0.04327362 to the left, improve=0.3854259, (0 missing)  
## CPI < 0.615269 to the left, improve=0.3776358, (0 missing)  
## Surrogate splits:  
## Monthly Nominal GDP Index (inMillion$) < 0.1292032 to the left, agree=0.988, adj=0.972, (0 split)  
## Earnings or wages in dollars per hour < 0.2141489 to the left, agree=0.988, adj=0.972, (0 split)  
## CPI < 0.43785 to the left, agree=0.976, adj=0.944, (0 split)  
## unemployment rate < -0.1040174 to the right, agree=0.976, adj=0.944, (0 split)  
## CommercialBankInterestRateonCreditCardPlans < -0.3861181 to the right, agree=0.893, adj=0.750, (0 split)  
##   
## Node number 2: 48 observations, complexity param=0.1213155  
## mean=2960.938, MSE=357481.8   
## left son=4 (23 obs) right son=5 (25 obs)  
## Primary splits:  
## Monthly Nominal GDP Index (inMillion$) < -0.8068302 to the left, improve=0.3966539, (0 missing)  
## Monthly Real GDP Index (inMillion$) < -0.7319077 to the left, improve=0.3966539, (0 missing)  
## CPI < -1.271988 to the left, improve=0.3853968, (0 missing)  
## Cotton Monthly Price - US cents per Pound(lbs) < -0.3683744 to the left, improve=0.3853968, (0 missing)  
## Earnings or wages in dollars per hour < -0.9768003 to the left, improve=0.3362920, (0 missing)  
## Surrogate splits:  
## Monthly Real GDP Index (inMillion$) < -0.7319077 to the left, agree=1.000, adj=1.000, (0 split)  
## CPI < -1.025193 to the left, agree=0.958, adj=0.913, (0 split)  
## Earnings or wages in dollars per hour < -0.9146638 to the left, agree=0.958, adj=0.913, (0 split)  
## Average upland planted(million acres) < -0.0656004 to the left, agree=0.938, adj=0.870, (0 split)  
## CommercialBankInterestRateonCreditCardPlans < 0.4999834 to the right, agree=0.875, adj=0.739, (0 split)  
##   
## Node number 3: 36 observations, complexity param=0.1222654  
## mean=4001.667, MSE=462866   
## left son=6 (29 obs) right son=7 (7 obs)  
## Primary splits:  
## Month splits as LLLLRLLLLLLR, improve=0.4116578, (0 missing)  
## Monthly Nominal GDP Index (inMillion$) < 0.6957071 to the left, improve=0.2132552, (0 missing)  
## Monthly Real GDP Index (inMillion$) < 0.5940022 to the left, improve=0.2079452, (0 missing)  
## Exports < -0.8434762 to the right, improve=0.1606711, (0 missing)  
## yieldperharvested acre < -0.06630414 to the right, improve=0.1602246, (0 missing)  
## Surrogate splits:  
## Monthly Nominal GDP Index (inMillion$) < 0.2313332 to the right, agree=0.861, adj=0.286, (0 split)  
## Monthly Real GDP Index (inMillion$) < 0.1954439 to the right, agree=0.833, adj=0.143, (0 split)  
## yieldperharvested acre < -1.103392 to the right, agree=0.833, adj=0.143, (0 split)  
## Production (in 480-lb netweright in million bales) < 0.8863121 to the left, agree=0.833, adj=0.143, (0 split)  
##   
## Node number 4: 23 observations, complexity param=0.02808224  
## mean=2568.348, MSE=193089.6   
## left son=8 (14 obs) right son=9 (9 obs)  
## Primary splits:  
## Month splits as LLRRRLLLRLLR, improve=0.3547607, (0 missing)  
## Monthly Nominal GDP Index (inMillion$) < -1.369315 to the left, improve=0.3234273, (0 missing)  
## Monthly Real GDP Index (inMillion$) < -1.374455 to the left, improve=0.3234273, (0 missing)  
## Earnings or wages in dollars per hour < -1.328907 to the left, improve=0.3234273, (0 missing)  
## Cotton Monthly Price - US cents per Pound(lbs) < -0.7839217 to the left, improve=0.3234273, (0 missing)  
## Surrogate splits:  
## AdvertisingExpenses (in Thousand Dollars) splits as LRL-R-L--R-L, agree=0.739, adj=0.333, (0 split)  
## Average upland harvested(million acres) < -1.12348 to the right, agree=0.739, adj=0.333, (0 split)  
## Production (in 480-lb netweright in million bales) < -1.080408 to the right, agree=0.739, adj=0.333, (0 split)  
## Monthly Nominal GDP Index (inMillion$) < -1.465963 to the right, agree=0.696, adj=0.222, (0 split)  
## Change(in%) < 0.8431257 to the left, agree=0.696, adj=0.222, (0 split)  
##   
## Node number 5: 25 observations, complexity param=0.04710283  
## mean=3322.12, MSE=236473.2   
## left son=10 (14 obs) right son=11 (11 obs)  
## Primary splits:  
## Month splits as LLRRRLLLLLRR, improve=0.44700820, (0 missing)  
## CPI < -0.2705274 to the right, improve=0.16947880, (0 missing)  
## Earnings or wages in dollars per hour < -0.1379578 to the right, improve=0.11714190, (0 missing)  
## Change(in%) < -0.1856634 to the right, improve=0.08220731, (0 missing)  
## Monthly Nominal GDP Index (inMillion$) < -0.2653809 to the right, improve=0.08005015, (0 missing)  
## Surrogate splits:  
## CPI < -0.2705274 to the right, agree=0.76, adj=0.455, (0 split)  
## CommercialBankInterestRateonCreditCardPlans < -0.3079327 to the left, agree=0.72, adj=0.364, (0 split)  
## Cotton Monthly Price - US cents per Pound(lbs) < 1.671907 to the left, agree=0.72, adj=0.364, (0 split)  
## Average upland planted(million acres) < 0.3478548 to the right, agree=0.72, adj=0.364, (0 split)  
## Production (in 480-lb netweright in million bales) < 0.9778613 to the left, agree=0.72, adj=0.364, (0 split)  
##   
## Node number 6: 29 observations, complexity param=0.05118376  
## mean=3787.207, MSE=250532.6   
## left son=12 (8 obs) right son=13 (21 obs)  
## Primary splits:  
## Month splits as LLRR-RLRRRR-, improve=0.3952395, (0 missing)  
## Monthly Nominal GDP Index (inMillion$) < 0.682752 to the left, improve=0.3550517, (0 missing)  
## Monthly Real GDP Index (inMillion$) < 0.5940022 to the left, improve=0.3550517, (0 missing)  
## CPI < 0.9956782 to the left, improve=0.3375080, (0 missing)  
## Change(in%) < -0.01213272 to the right, improve=0.2626150, (0 missing)  
## Surrogate splits:  
## Change(in%) < 0.3920836 to the right, agree=0.793, adj=0.250, (0 split)  
## yieldperharvested acre < 0.2601866 to the right, agree=0.759, adj=0.125, (0 split)  
##   
## Node number 7: 7 observations  
## mean=4890.143, MSE=362600.4   
##   
## Node number 8: 14 observations  
## mean=2358.5, MSE=110710.1   
##   
## Node number 9: 9 observations  
## mean=2894.778, MSE=146178.4   
##   
## Node number 10: 14 observations  
## mean=3033.929, MSE=72192.64   
##   
## Node number 11: 11 observations  
## mean=3688.909, MSE=205317.9   
##   
## Node number 12: 8 observations  
## mean=3277.375, MSE=214970.2   
##   
## Node number 13: 21 observations, complexity param=0.01945655  
## mean=3981.429, MSE=127337.8   
## left son=26 (7 obs) right son=27 (14 obs)  
## Primary splits:  
## Monthly Nominal GDP Index (inMillion$) < 0.7151312 to the left, improve=0.4082064, (0 missing)  
## CPI < 0.919077 to the left, improve=0.4082064, (0 missing)  
## Earnings or wages in dollars per hour < 0.737131 to the left, improve=0.4082064, (0 missing)  
## unemployment rate < -0.6461228 to the right, improve=0.4082064, (0 missing)  
## Mill use (in 480-lb netweright in million bales) < 3.65 to the left, improve=0.2993566, (0 missing)  
## Surrogate splits:  
## CPI < 0.919077 to the left, agree=1.000, adj=1.000, (0 split)  
## unemployment rate < -0.6461228 to the right, agree=1.000, adj=1.000, (0 split)  
## Earnings or wages in dollars per hour < 0.737131 to the left, agree=1.000, adj=1.000, (0 split)  
## Monthly Real GDP Index (inMillion$) < 0.5698506 to the left, agree=0.952, adj=0.857, (0 split)  
## Mill use (in 480-lb netweright in million bales) < 3.53 to the left, agree=0.952, adj=0.857, (0 split)  
##   
## Node number 26: 7 observations  
## mean=3659, MSE=27695.43   
##   
## Node number 27: 14 observations  
## mean=4142.643, MSE=99188.8

prediction1 = predict(rf, Train\_df)  
  
library(DMwR)  
  
regr.eval(Train\_df$Sales.In.ThousandDollars., prediction1)

## mae mse rmse mape   
## 2.954718e+02 1.425621e+05 3.775740e+02 9.231719e-02

prediction.test = predict(rf, Test\_df)  
  
prediction.test

## 1 2 3 4 5 6 7 8   
## 3277.375 3277.375 4142.643 4142.643 4890.143 4142.643 3277.375 4142.643   
## 9 10 11 12   
## 4142.643 4142.643 4142.643 4890.143

regr.eval(Test\_df$Sales.In.ThousandDollars., prediction.test)

## mae mse rmse mape   
## NaN NaN NaN NaN

rf$cptable[which.min(rf$cptable[,"xerror"]),"CP"]->min.xerror  
  
rt.pruned <- prune(rf,cp = min.xerror)   
  
train.pred.rtree.p <- predict(rt.pruned,Train\_df)  
  
#test.pred.rtree.p <- predict(train.pred.rtree.p,Test\_df)  
  
test.pred.rtree.p <- predict(rt.pruned,Test\_df)