walmart

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.1.3

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

library(forecast)

## Warning: package 'forecast' was built under R version 4.1.3

## Registered S3 method overwritten by 'quantmod':  
## method from  
## as.zoo.data.frame zoo

library(partykit)

## Warning: package 'partykit' was built under R version 4.1.3

## Loading required package: grid

## Loading required package: libcoin

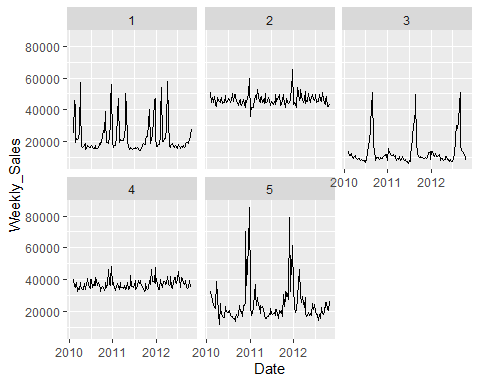
## Warning: package 'libcoin' was built under R version 4.1.3

## Loading required package: mvtnorm

library(rpart)  
train <- read.csv("C:\\Users\\Fitzg\\OneDrive\\Documents\\train.csv", header=TRUE)

train$Date <- as.Date(train$Date)

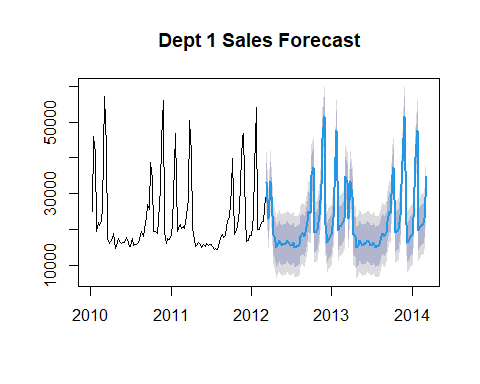
ggplot(train, aes(Date, Weekly\_Sales)) + geom\_line() +facet\_wrap(~Dept)



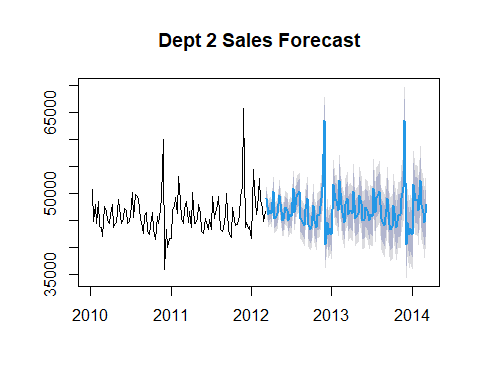
train$IsHoliday <- as.factor(train$IsHoliday)  
model <- glm(Weekly\_Sales ~+IsHoliday+temp+Fuel\_Price+CPI+Unemployment, data = train)  
summary(model)

##   
## Call:  
## glm(formula = Weekly\_Sales ~ +IsHoliday + temp + Fuel\_Price +   
## CPI + Unemployment, data = train)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -26386 -11058 -2397 11874 56606   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -91122.74 75001.63 -1.215 0.2248   
## IsHolidayTRUE 1928.45 2110.97 0.914 0.3613   
## temp -85.00 39.45 -2.155 0.0315 \*  
## Fuel\_Price -2311.75 2025.28 -1.141 0.2541   
## CPI 497.18 290.35 1.712 0.0873 .  
## Unemployment 3343.38 2512.66 1.331 0.1837   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for gaussian family taken to be 196321174)  
##   
## Null deviance: 1.4162e+11 on 714 degrees of freedom  
## Residual deviance: 1.3919e+11 on 709 degrees of freedom  
## AIC: 15690  
##   
## Number of Fisher Scoring iterations: 2

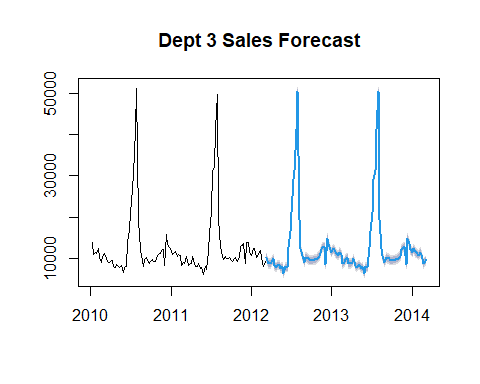
dept\_1 <- train %>%  
 filter(train$Dept == 1)  
  
train.ts <- ts(dept\_1$Weekly\_Sales, frequency = 52, start = c(2010, 2), end = c(2012, 10))  
train.forecast <- forecast(train.ts)  
plot(train.forecast, main = "Dept 1 Sales Forecast")



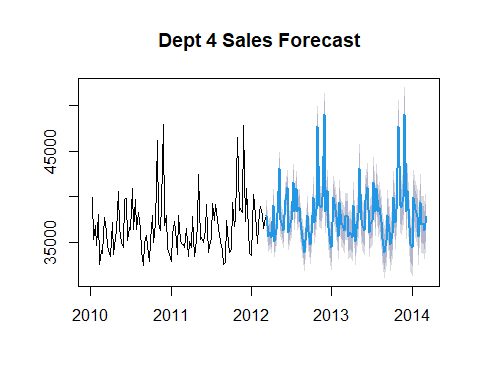
dept\_2 <- train %>%  
 filter(train$Dept == 2)  
  
train.ts2 <- ts(dept\_2$Weekly\_Sales, frequency = 52, start = c(2010, 2), end = c(2012, 10))  
train.forecast2 <- forecast(train.ts2)  
plot(train.forecast2, main = "Dept 2 Sales Forecast")



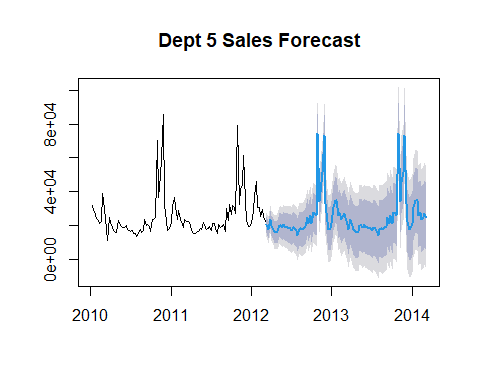
dept\_3 <- train %>%  
 filter(train$Dept == 3)  
  
train.ts3 <- ts(dept\_3$Weekly\_Sales, frequency = 52, start = c(2010, 2), end = c(2012, 10))  
train.forecast3 <- forecast(train.ts3)  
plot(train.forecast3, main = "Dept 3 Sales Forecast")



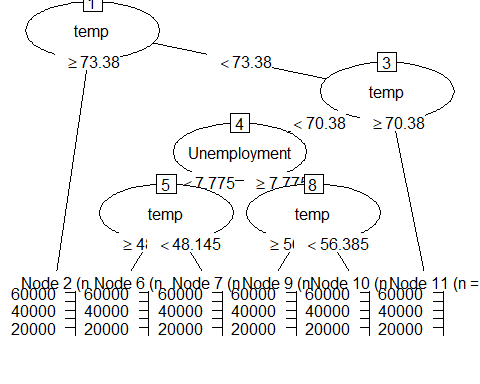
dept\_4 <- train %>%  
 filter(train$Dept == 4)  
  
train.ts4 <- ts(dept\_4$Weekly\_Sales, frequency = 52, start = c(2010, 2), end = c(2012, 10))  
train.forecast4 <- forecast(train.ts4)  
plot(train.forecast4, main = "Dept 4 Sales Forecast")



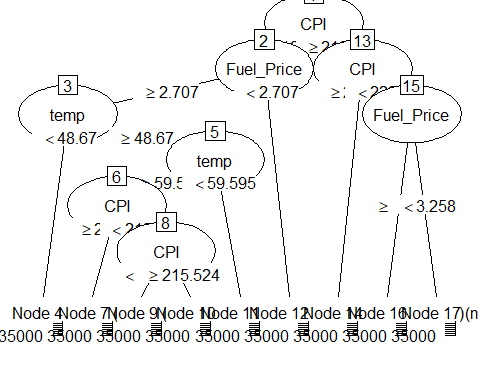
dept\_5 <- train %>%  
 filter(train$Dept == 5)  
  
train.ts5 <- ts(dept\_5$Weekly\_Sales, frequency = 52, start = c(2010, 2), end = c(2012, 10))  
train.forecast5 <- forecast(train.ts5)  
plot(train.forecast5, main = "Dept 5 Sales Forecast")

 Bagging

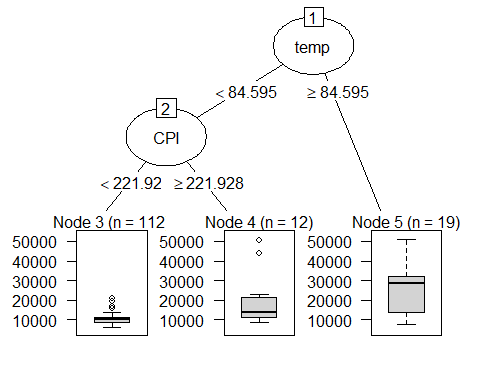
set.seed(8675309)   
train.ct <- rpart (Weekly\_Sales ~ +IsHoliday+temp+Fuel\_Price+CPI+Unemployment, data = dept\_1)  
train.ct.party <- as.party(train.ct)  
plot (train.ct.party)



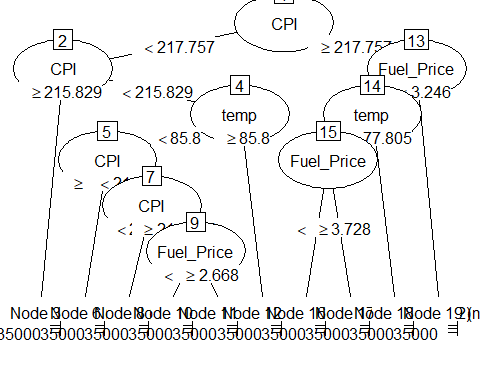
set.seed(8675309)   
train.ct <- rpart (Weekly\_Sales ~ +IsHoliday+temp+Fuel\_Price+CPI+Unemployment, data = dept\_2)  
train.ct.party <- as.party(train.ct)  
plot (train.ct.party)



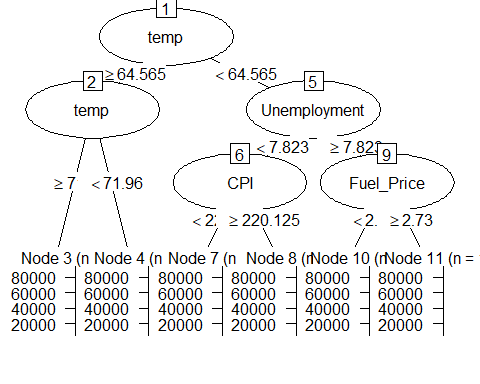
set.seed(8675309)   
train.ct <- rpart (Weekly\_Sales ~ +IsHoliday+temp+Fuel\_Price+CPI+Unemployment, data = dept\_3)  
train.ct.party <- as.party(train.ct)  
plot (train.ct.party)



set.seed(8675309)   
train.ct <- rpart (Weekly\_Sales ~ +IsHoliday+temp+Fuel\_Price+CPI+Unemployment, data = dept\_4)  
train.ct.party <- as.party(train.ct)  
plot (train.ct.party)



set.seed(8675309)   
train.ct <- rpart (Weekly\_Sales ~ +IsHoliday+temp+Fuel\_Price+CPI+Unemployment, data = dept\_5)  
train.ct.party <- as.party(train.ct)  
plot (train.ct.party)



by\_date <- train %>%  
 filter(train$IsHoliday == TRUE)  
   
ggplot(by\_date, aes(Date, Weekly\_Sales)) + geom\_line() +facet\_wrap(~Dept)

