symbols=c('ADANIPORTS.BO','ASIANPAINT.BO','AXISBANK.BO','BAJAJ-AUTO.BO','BHARTIARTL.BO','COALINDIA.BO','DRREDDY.BO'

,'HDFCBANK.BO','HEROMOTOCO.BO','HINDUNILVR.BO','HDFC.BO','ICICIBANK.BO','INDUSINDBK.BO','INFY.BO','ITC.BO','KOTAKBANK.BO'

,'LT.BO','M&M.BO','MARUTI.BO','NTPC.BO','ONGC.BO','POWERGRID.BO','RELIANCE.BO','SBIN.BO','SUNPHARMA.BO','TCS.BO','TATAMOTORS.BO',

'TATASTEEL.BO','WIPRO.BO','YESBANK.BO','^BSESN')

STOCKS1=lapply(symbols, function(x){ #

dailyReturn(na.omit(getSymbols(x, auto.assign=FALSE)))

})

STOCKS2=do.call(merge.xts,STOCKS1) #do.call calls a function to apply on the stocks1 here

colnames(STOCKS2)=c('ADANIPORTS.BO','ASIANPAINT.BO','AXISBANK.BO','BAJAJ-AUTO.BO','BHARTIARTL.BO','COALINDIA.BO','DRREDDY.BO'

,'HDFCBANK.BO','HEROMOTOCO.BO','HINDUNILVR.BO','HDFC.BO','ICICIBANK.BO','INDUSINDBK.BO','INFY.BO','ITC.BO','KOTAKBANK.BO'

,'LT.BO','M&M.BO','MARUTI.BO','NTPC.BO','ONGC.BO','POWERGRID.BO','RELIANCE.BO','SBIN.BO','SUNPHARMA.BO','TCS.BO','TATAMOTORS.BO',

'TATASTEEL.BO','WIPRO.BO','YESBANK.BO','BSESN') #function called here is merge.xts resulted in a large xts data

stocks3=data.frame(STOCKS2)

sort(colSums(is.na(stocks3)),decreasing = T)

delcols=c('BAJAJ.AUTO.BO','SBIN.BO','COALINDIA.BO')

stocks4=stocks3[!colnames(stocks3) %in% delcols]

sort(colSums(is.na(stocks4)),decreasing = T)

stocks4=sapply(stocks4, function(x) ifelse(is.na(x),mean(x,na.rm = TRUE),x))

str(stocks4)

stocks4=data.frame(stocks4)

str(stocks4)

stocks4$move=ifelse(stocks4$BSESN>0,1,0) #1:up,0:Down

numericcols=Filter(is.numeric,stocks4)

cornum=cor(numericcols)

write.csv(cornum,"cornum.csv")

stocks4=stocks4[-28]

#logistic reg.

stocklogit=glm(move~.,data=stocks4,family = "binomial")

summary(stocklogit)

stocklogit$fitted.values

table(Pred=stocklogit$fitted.values>0.5,Actual=stocks4$move)

(1216+1360)/(1216+84+85+1360)

logitpredict=ifelse(stocklogit$fitted.values>=0.5,1,0)

confusionMatrix(logitpredict,stocks4$move)

#decision tree

stocksrpart=rpart(as.factor(move)~.,data=stocks4,parms = list(split="information"))

summary(stocksrpart)

rpart.plot(stocksrpart,cex=0.8)

rpartpred=predict(stocksrpart,type="class")

table(Pred=rpartpred,Actual=stocks4$move)

confusionMatrix(rpartpred,stocks4$move)

#randomforest

stocksrandforest=randomForest(as.factor(move)~.,data=stocks4,ntree=3000,importance=T,do.trace=200)

print(stocksrandforest)

(100-8.12)