Week 1-Zoo

# Time Series using ZOO

#LOAD LIBRARIES #you may need to install packages, if this is the first time you use them. Select Packages > Install Packages in R/RStudio)

library(tseries)

## Registered S3 method overwritten by 'xts':  
## method from  
## as.zoo.xts zoo

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

library(zoo)

##   
## Attaching package: 'zoo'

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

#import dataset into a dataframe)  
cisco <- read.table('C:/Users/Xuan Pham/Dropbox/Fall\_2019/BIA6315/code/Week 1/Data/cisco\_00-10.csv', header=T, sep=',')  
# create time series for cisco prices  
  
ciscots = zoo(cisco$Price, as.Date(as.character(cisco$Date), format = "%m/%d/%y"))  
class(ciscots)

## [1] "zoo"

#To retrieve only dates use  
print("head of TS")

## [1] "head of TS"

head(time(ciscots))

## [1] "2000-01-03" "2000-01-04" "2000-01-05" "2000-01-06" "2000-01-07"  
## [6] "2000-01-10"

print("Retrieve start date")

## [1] "Retrieve start date"

start(ciscots)

## [1] "2000-01-03"

print("Retrieve End date")

## [1] "Retrieve End date"

end(ciscots)

## [1] "2010-12-31"

head(ciscots)

## 2000-01-03 2000-01-04 2000-01-05 2000-01-06 2000-01-07 2000-01-10   
## 54.03 51.00 50.85 50.00 52.94 54.90

#Creating new variables

# create lagged series using function lag(tsobject, k==1);  
pricelag = lag(ciscots, k=-1);  
head(pricelag)

## 2000-01-04 2000-01-05 2000-01-06 2000-01-07 2000-01-10 2000-01-11   
## 54.03 51.00 50.85 50.00 52.94 54.90

# diff = p\_t - p\_(t-1);  
pricedif = diff(ciscots);  
  
  
#compute simple returns ret = (p\_t-p\_(t-1))/p\_(t-1)  
#day-to-day rate of change  
ret=(ciscots-pricelag)/pricelag  
head(ret)

## 2000-01-04 2000-01-05 2000-01-06 2000-01-07 2000-01-10   
## -0.056079956 -0.002941176 -0.016715831 0.058800000 0.037023045   
## 2000-01-11   
## -0.030054645

# sort data in chronological order  
# set variable Date as time/date variable  
cisco$Date=as.Date(as.character(cisco$Date), format = "%m/%d/%y")  
cisco=cisco[order(cisco$Date),]  
head(cisco)

## Date Price  
## 2767 2000-01-03 54.03  
## 2766 2000-01-04 51.00  
## 2765 2000-01-05 50.85  
## 2764 2000-01-06 50.00  
## 2763 2000-01-07 52.94  
## 2762 2000-01-10 54.90

#DEFINE LOG RETURNS  
#Make the data less skewed (more normally distributed).  
#rts is a time series object since it is created from a zoo object  
rts = diff(log(ciscots))  
print("Log Return")

## [1] "Log Return"

head(rts)

## 2000-01-04 2000-01-05 2000-01-06 2000-01-07 2000-01-10 2000-01-11   
## -0.05771382 -0.00294551 -0.01685712 0.05713619 0.03635415 -0.03051554

print("")

## [1] ""

#to retrieve numerical values from time series use coredata()  
# rt is a numerical vector (no date information)  
print("coredata")

## [1] "coredata"

rt=coredata(rts)  
#print first 6 values  
head(rt)

## [1] -0.05771382 -0.00294551 -0.01685712 0.05713619 0.03635415 -0.03051554

# LOAD LIBRARIES

# Load fBasics packages into current session

# To install the package the first time,

# select Tools from top Menu and select Install Packages

library(fBasics)

## Loading required package: timeDate

## Loading required package: timeSeries

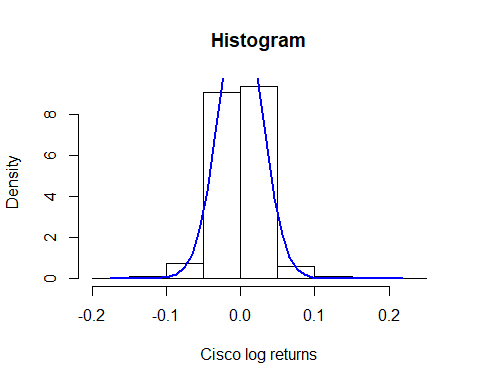
##   
## Attaching package: 'timeSeries'

## The following object is masked from 'package:zoo':  
##   
## time<-

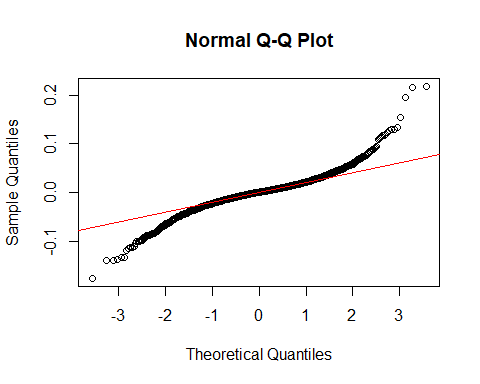
# COMPUTE SUMMARY STATISTICS  
basicStats(rt)

## rt  
## nobs 2766.000000  
## NAs 0.000000  
## Minimum -0.176865  
## Maximum 0.218239  
## 1. Quartile -0.013890  
## 3. Quartile 0.013411  
## Mean -0.000355  
## Median 0.000449  
## Sum -0.982373  
## SE Mean 0.000560  
## LCL Mean -0.001453  
## UCL Mean 0.000742  
## Variance 0.000867  
## Stdev 0.029437  
## Skewness 0.187810  
## Kurtosis 6.053895

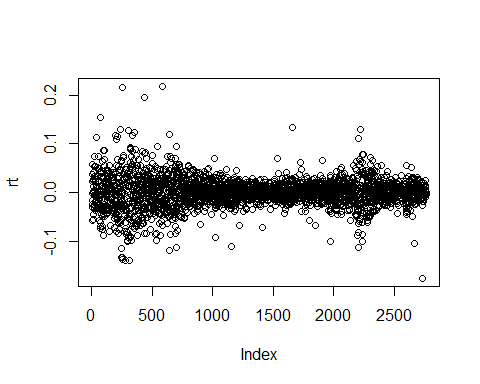
# CREATE HISTOGRAM   
# OPTIONAL creates 2 by 2 display for 4 plots   
# par(mfcol=c(2,2))   
hist(rt, xlab="Cisco log returns", prob=TRUE, main="Histogram")   
# add approximating normal density curve   
xfit<-seq(min(rt),max(rt),length=40)   
yfit<-dnorm(xfit,mean=mean(rt),sd=sd(rt))   
lines(xfit, yfit, col="blue", lwd=2)



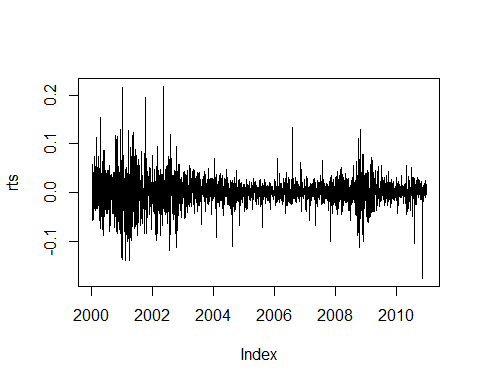
# CREATE NORMAL PROBABILITY PLOT   
qqnorm(rt)   
qqline(rt, col = 2)



# CREATE TIME PLOTS   
# simple plot where x-axis is not labeled with time   
plot(rt)



# use time series object rts to draw time plot indexed with time   
plot(rts)



# creates subsets of data for a certain period of time   
rts\_10 = window(rts, start = as.Date("2010-01-01"), end = as.Date("2010-12-31"))   
# plot the new subset   
plot(rts\_10, type='l', ylab="log returns", main="Plot of 2010 data")

