Econ 424 Lab 4, Winter 2020

Introduction

In this lab, you will compute descriptive statistics and find stylized facts for the monthly return data on five Northwest stocks in the **IntroCompFin** package: Amazon (amzn), Boeing (ba), Costco (cost), Nordstrom (jwn), and Starbucks (sbux). I encourage you to go to https://finance.yahoo.com (https://finance.yahoo.com) and research these stocks. This notebook walks you through all of the computations for the lab. You will use the following R packages

- **IntroCompFinR**
- **PerformanceAnalytics package**.
- **zoo**
- **xts**

Make sure to install these packages before you load them into R. As in the previous labs, use this notebook to answer all questions. Insert R chunks where needed. I will provide code hints below.

Reading

- Zivot, chapter 5 (Descriptive statistics)
- Ruppert and Matteson, chapter 4 (Exploratory data analysis)
- Zivot, “Working with Time Series in R”
- **corrplot**, **PerformanceAnalytics**, and **zoo** vignettes

Load packages and set options

```
library(IntroCompFinR)
```

```
## Loading required package: xts
```

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

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

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

```
library(corrplot)
```

```
## corrplot 0.84 loaded
```

```
library(PerformanceAnalytics)
```

```
##
## Attaching package: 'PerformanceAnalytics'
```

```
## The following object is masked from 'package:graphics':
##
##     legend
```

```
library(xts)
options(digits = 3)
Sys.setenv(TZ="UTC")
```

Loading data and computing returns

Load the daily price data from **IntroCompFinR**, and create monthly returns over the period Jan 1998 through Dec 2014:

```
data(amznDailyPrices, baDailyPrices, costDailyPrices, jwnDailyPrices, sbuxDailyPrices)
fiveStocks = merge(amznDailyPrices, baDailyPrices, costDailyPrices, jwnDailyPrices, sbuxDailyPri
ces)
fiveStocks = to.monthly(fiveStocks, OHLC=FALSE)
```

```
## Warning in to.period(x, "months", indexAt = indexAt, name = name, ...): missing
## values removed from data
```

Let's look at the data

```
head(fiveStocks, n=3)
```

```
##          AMZN   BA COST  JWN SBUX
## Jan 1998 4.92 34.0 17.4  9.4 4.24
## Feb 1998 6.42 38.9 19.6 10.6 4.59
## Mar 1998 7.13 37.4 21.4 11.8 5.25
```

```
tail(fiveStocks, n=3)
```

```
##          AMZN  BA COST JWN SBUX
## Oct 2014  305 123  128  72 75.0
## Nov 2014  339 134  137  76 80.9
## Dec 2014  310 129  137  79 81.8
```

Next, let's compute monthly simple returns using the **PerformanceAnalytics** function `Return.Calculate()`

```
fiveStocksRet = na.omit(Return.calculate(fiveStocks, method = "simple"))
head(fiveStocksRet, n=3)
```

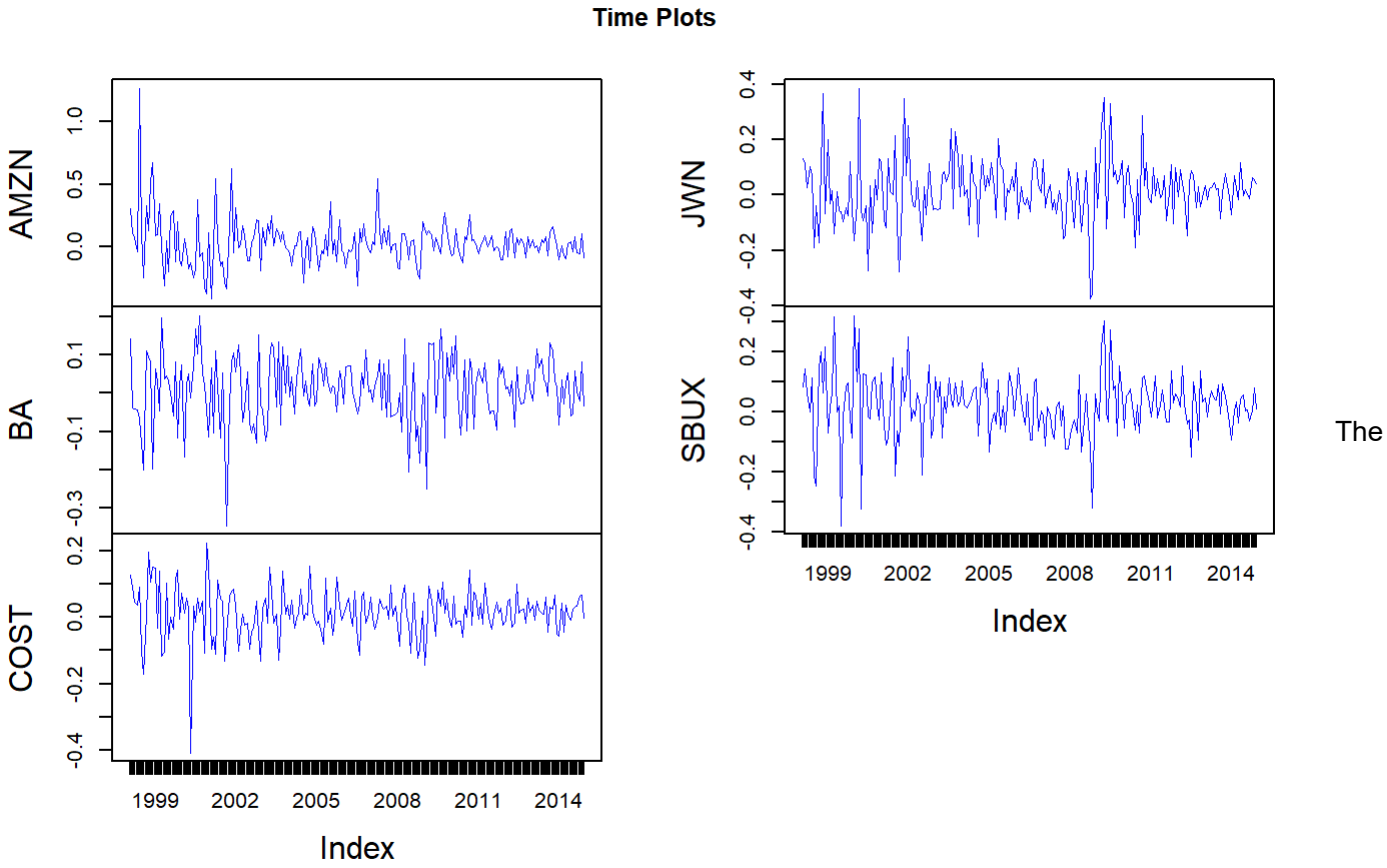
```
##            AMZN      BA    COST    JWN   SBUX
## Feb 1998 0.3049  0.1425 0.1267 0.1298 0.0825
## Mar 1998 0.1106 -0.0393 0.0920 0.1121 0.1438
## Apr 1998 0.0729 -0.0396 0.0468 0.0262 0.0629
```

We removed the missing January return using the function `na.omit()`.

Part I: Univariate Graphical Analysis

1. Make time plots of the return data using the **zoo** function `plot.zoo()`. Comment on any relationships between the returns suggested by the plots.

```
options(digits=3,width=70)
plot.zoo(fiveStocksRet,main="Time Plots",lwd=.75, col="blue")
```

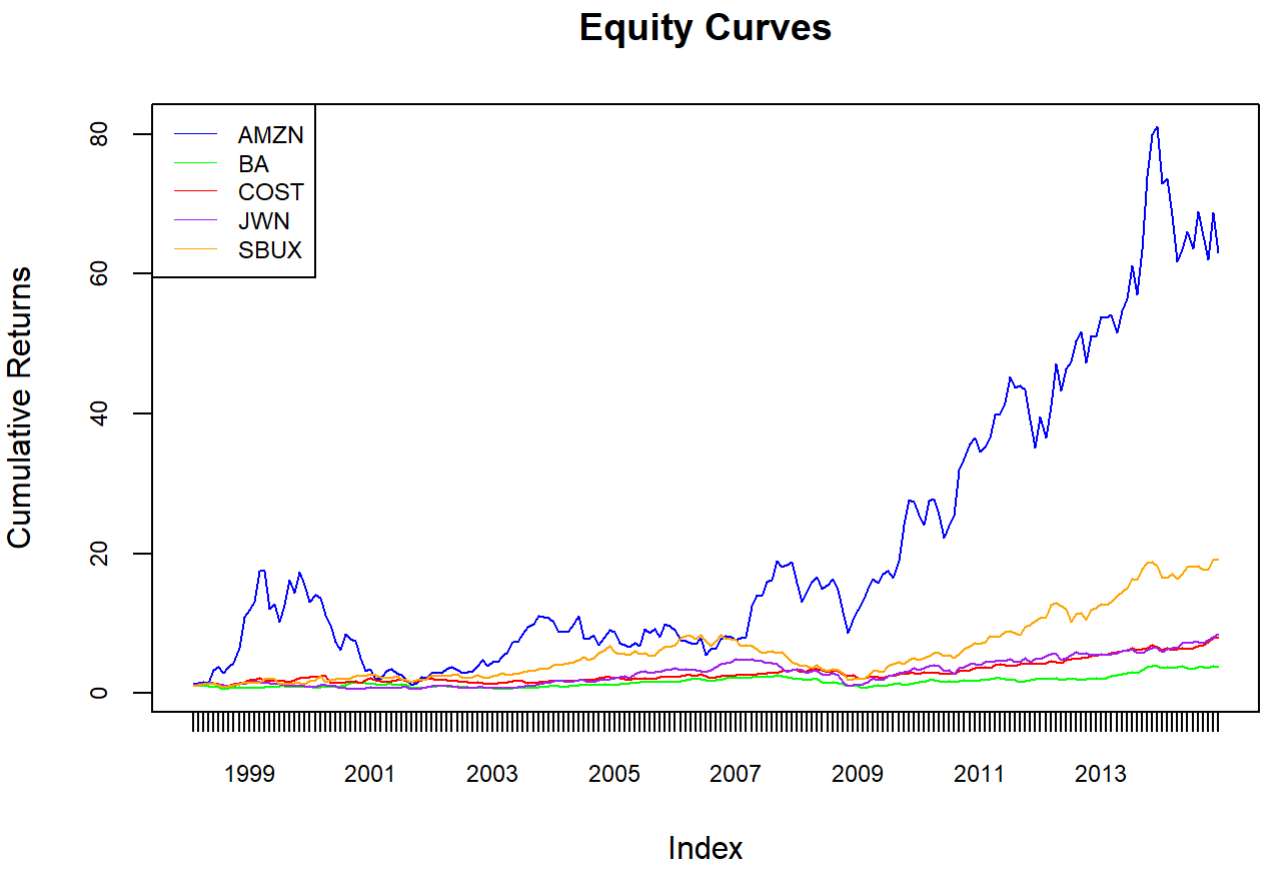


The plots suggest that you have the highest likelyhood of getting positive returns from the Amazon stock since its returns seem to rarely drop below 0. The volotility of returns seems to be the least for Costco and Boeing and seem to be quite high for Starbucks and Nordstrom.

2. Make a cumulative return plot (equity curve) showing the future value of $1 invested in each asset. Which assets gave the best and worst future values over the investment horizon? Hint: you can use the

`cumprod()` function or the **PerformanceAnalytics** function `chart.CumReturns()`.

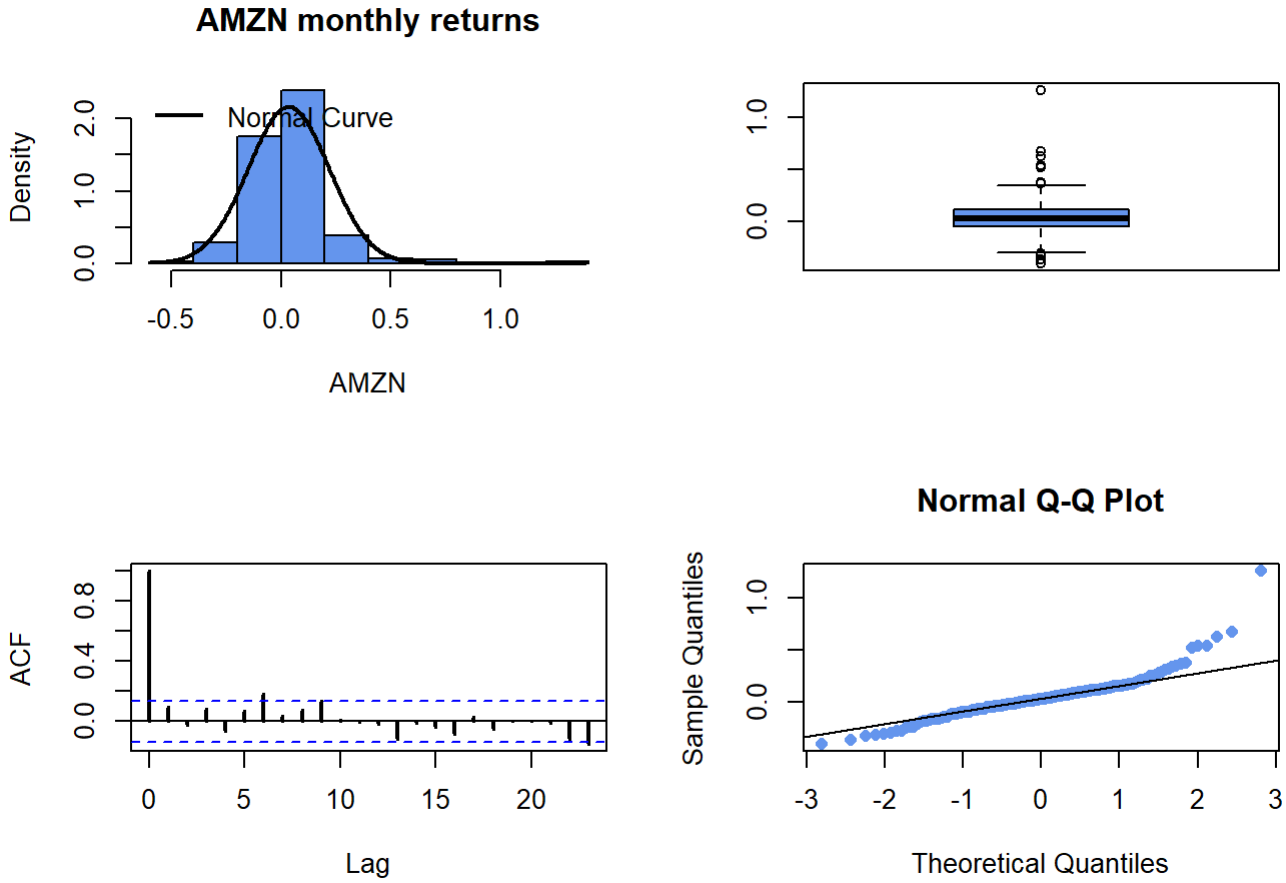
```
equityCurvesFiveStocks=cumprod(1+fiveStocksRet)
plot.zoo(equityCurvesFiveStocks,plot.type = "single",ylab="Cumulative Returns", main="Equity Cur
ves",col=c("blue","green","red","purple","orange"),cex.axis=.75)
legend(x="topleft",legend=c("AMZN","BA","COST","JWN","SBUX"),lwd=.75,col=c("blue","green","red",
"purple","orange"),cex=.75)
```



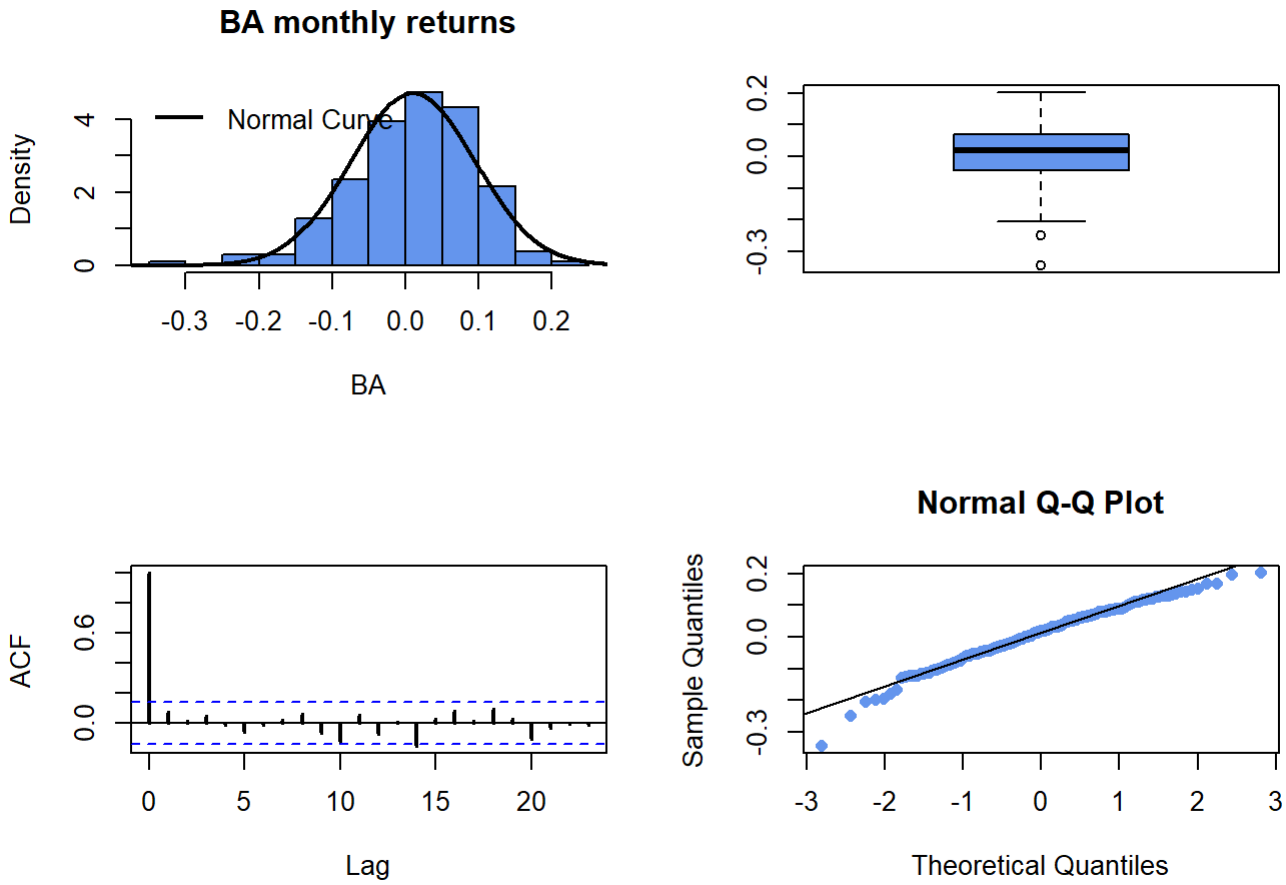
For the future values, Amazon gave the best future returns, taking $1 to over $60. While Boeing gave had the worst future values, turning $1 into just $5.

3. For each return series, make a four panel plot containing a histogram, boxplot, normal QQ-plot and SACF. Do the return series look normally distributed? Is there any evidence of time dependence? Briefly compare the return distributions. Hint: the `fourPanelPlot()` function only works on one asset at a time. So call the function for each asset. For example, `fourPanelPlot(fiveStocksRet[, "AMZN])`

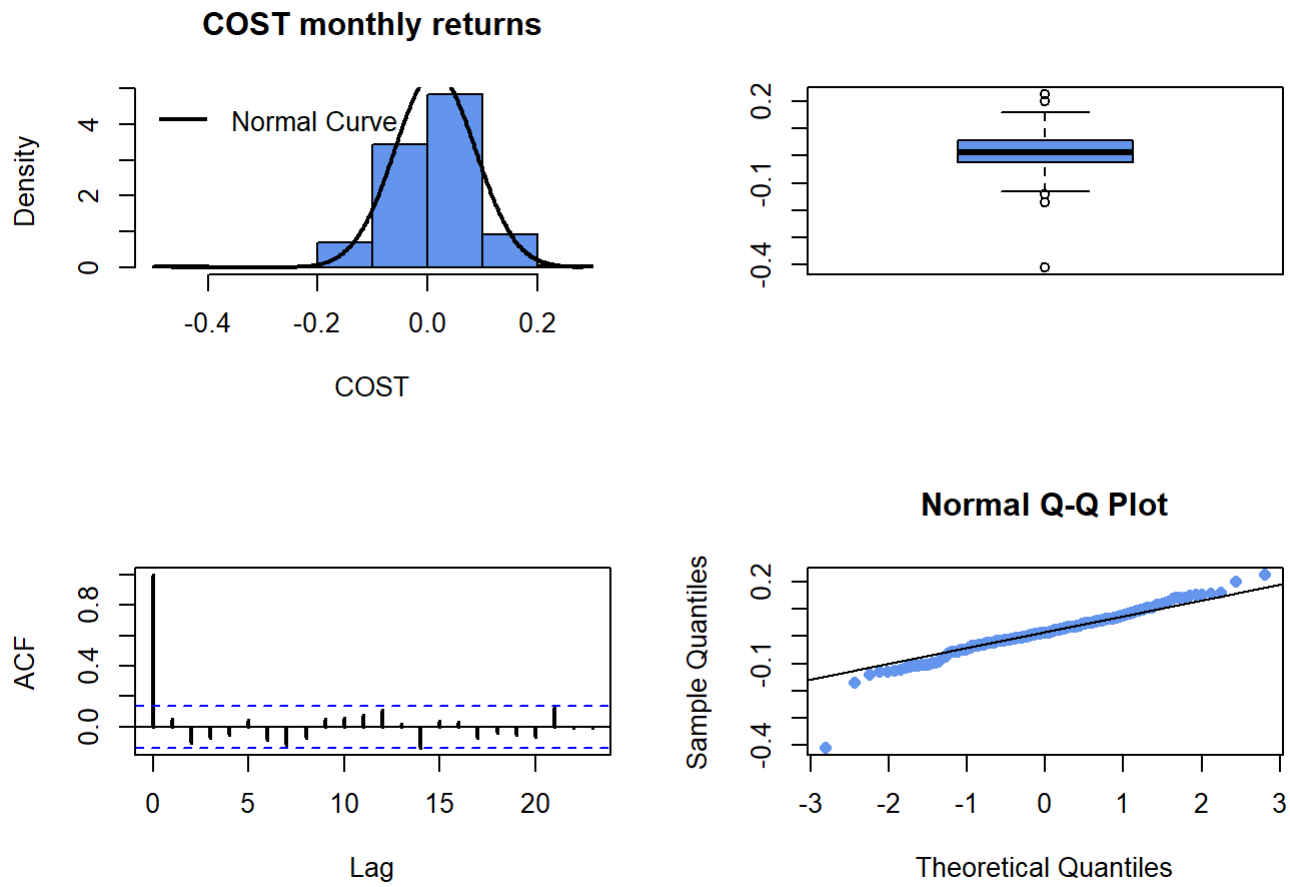
```
fourPanelPlot(fiveStocksRet[,"AMZN"])
```



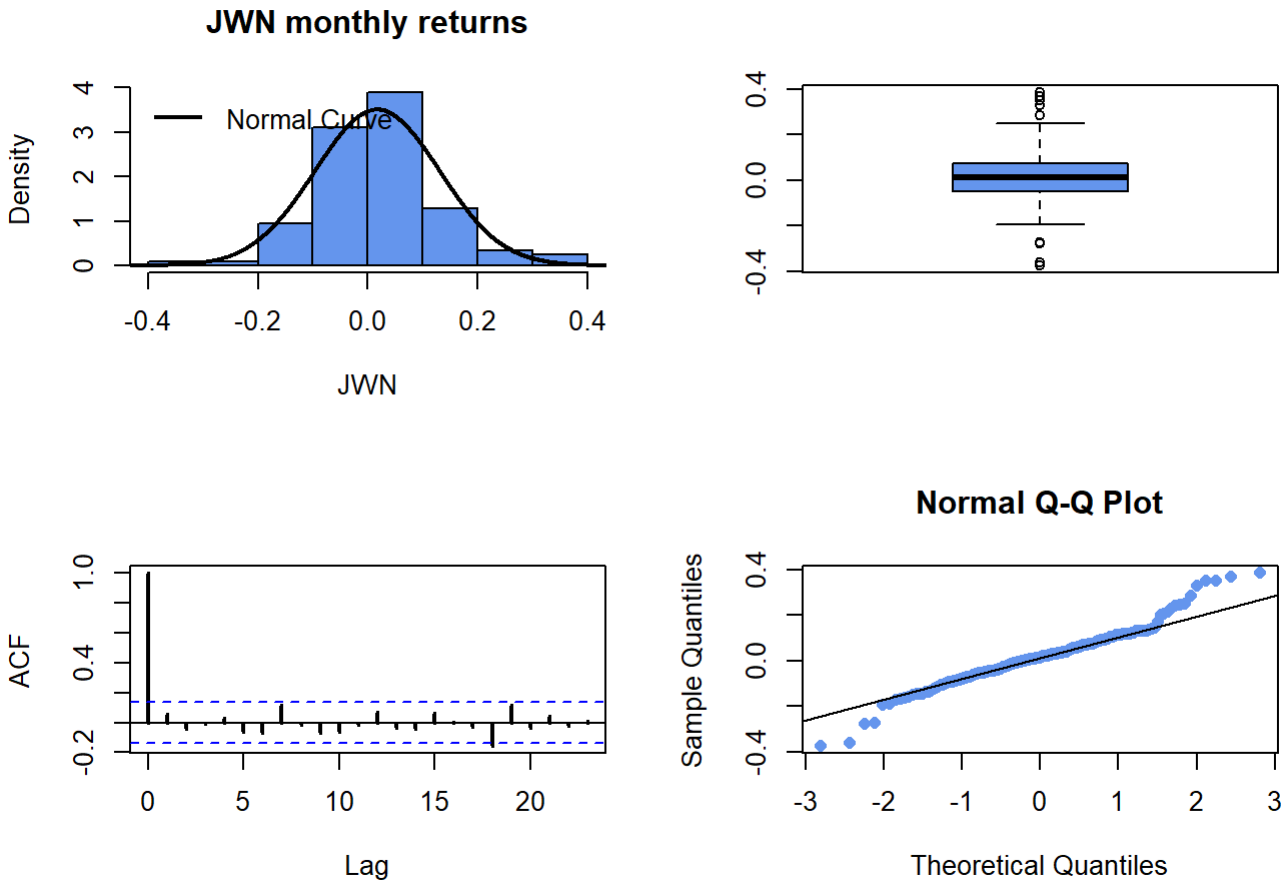

```
fourPanelPlot(fiveStocksRet[,"BA"])
```



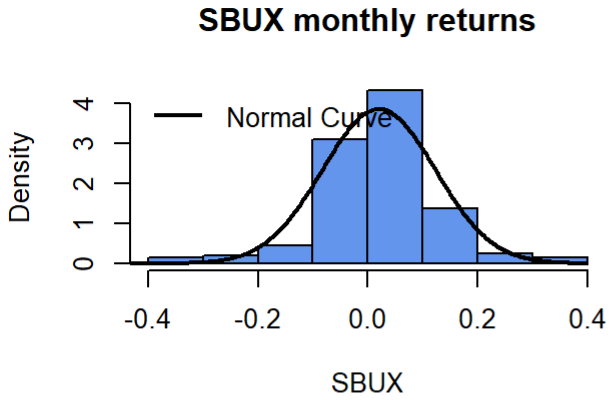

```
fourPanelPlot(fiveStocksRet[,"COST"])
```

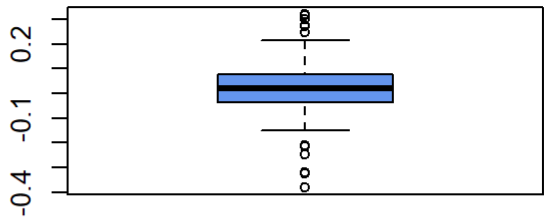



```
fourPanelPlot(fiveStocksRet[,"JWN"])
```

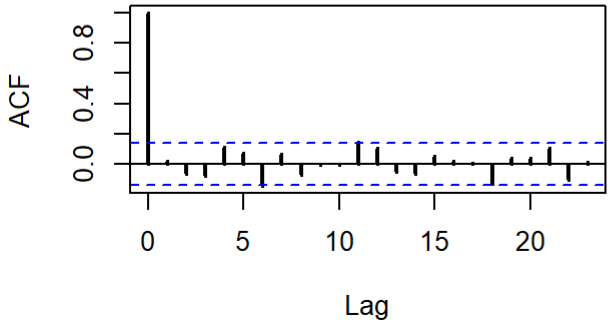


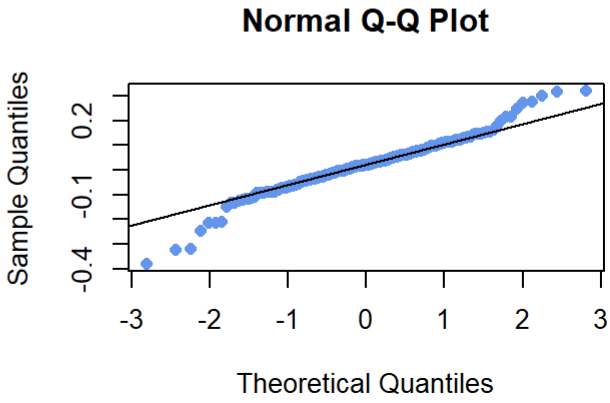

```
fourPanelPlot(fiveStocksRet[,"SBUX"])
```





All of





the stocks seem to be normally distributed with bell shaped curves and all look to be centered around 0. There does not seem to be strong evidence of time dependency, due to the fact that in all of the Lag graphs their lines don't exceed the dotted blue line. Amazon has the largest spread, while Costco has the least amount of spread except for its large negative outlier.

Part II: Univariate Numerical Summaries

1. Compute numerical descriptive statistics for all assets using the R functions `mean()`, `var()`, `stdev()`, and the **PerformanceAnalytics** functions `skewness()` and `kurtosis()`. Compare and contrast the descriptive statistics for the five assets. Which asset appears to be the riskiest asset?

Failed function try at automating data

```
#descript_stat <- function(data) {
  #number = c(1:ncol(data))
  #count = 1
  #mean <- c()
  #variance <- c()
  #sd <- c()
  #skew <- c()
  #kurt <- c()
  #for(val in number){
    #asset[count] <- data[,count]
    #mean[count] <- mean(asset)
    #variance[count] <- var(asset)
    #sd[count]  <- StdDev(asset)
    #skew[count] <- skewness(asset)
    #kurt[count] <- kurtosis(asset)
    #count = count + 1
  #}
  #stats <- data.frame(mean = mean, var = variance, sd = sd, skew = skew, kurt = kurt, stringsAs
Factors = FALSE)
  #print(stats)
#}
#descript_stat(fiveStocksRet)
```

Numerical descriptive statistics

```
statsMonthly=rbind(apply(fiveStocksRet,2,mean), apply(fiveStocksRet,2,var), apply(fiveStocksRet,
2,sd), apply(fiveStocksRet,
2,skewness), apply(fiveStocksRet,2,kurtosis))
rownames(statsMonthly)=c("Mean","Variance","Std Dev","Skewness","Excess Kurtosis")
round(statsMonthly, digits=4)
```

```
##                   AMZN      BA    COST    JWN    SBUX
## Mean             0.0354  0.0103  0.0130 0.0170  0.0201
## Variance         0.0341  0.0072  0.0054 0.0129  0.0106
## Std Dev          0.1845  0.0848  0.0734 0.1137  0.1031
## Skewness         1.8421 -0.6375 -0.8488 0.1674 -0.3198
## Excess Kurtosis 10.0085  1.0912  4.8161 1.9118  2.1383
```

Amazon has the highest mean for its returns meaning that on average it will give the highest returns, although with this it also has the highest standard deviation, which means that it also is the most volatile. Another hint that Amazon might give the highest returns is that its skewness is largely positive compared to the others, which also tells us that it is not normally distributed. The closest to normally distributed is Nordtrom while all others are skewed negitively. The Kurtosis for Amazon and Costco are the only two that have positive kurtosis telling us they are heavy tailed while the other 3 stocks are all negitive, showing us they are light tailed.

2. Using the monthly mean return for each asset, compute an estimate of the mean annual return. Hint: use the square root of time rule for cc returns.

```
12*statsMonthly["Mean",]
```

```
##  AMZN    BA  COST   JWN  SBUX
## 0.424 0.124 0.157 0.204 0.241
```

3. Using the monthly return standard deviation for each asset, compute an estimate of the annual return standard deviation. Hint: use the square root of time rule for cc returns.

```
sqrt(12)*statsMonthly["Std Dev",]
```

```
##  AMZN    BA  COST   JWN  SBUX
## 0.639 0.294 0.254 0.394 0.357
```

Part III: Historical VaR

1. For each asset compute the empirical 1% and 5% quantiles of the returns. Using these quantiles compute the 1% and 5% historical (monthly) VaR values based on an initial $100,000 investment. Which asset has the highest and lowest VaR values? Are you surprised?

```
apply(fiveStocksRet,2,quantile,probs=c(.01,.05))
```

```
##      AMZN     BA   COST    JWN   SBUX
## 1% -0.332 -0.206 -0.142 -0.278 -0.319
## 5% -0.241 -0.123 -0.109 -0.151 -0.125
```

```
w=100000
fiveStocksQuantiles=apply(fiveStocksRet,2,quantile,probs=c(.01,.05))
fiveStocksVaR=w*fiveStocksQuantiles
fiveStocksVaR
```

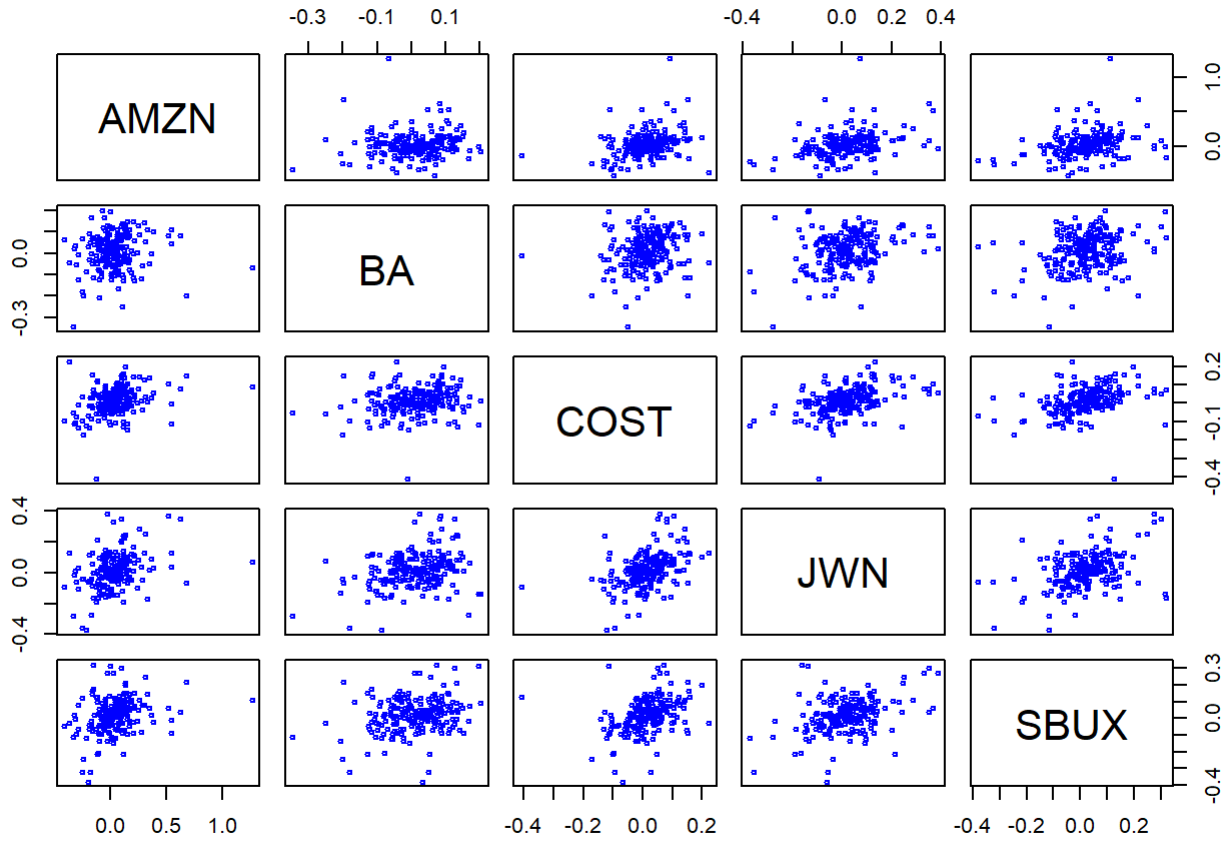
```
##      AMZN     BA   COST    JWN   SBUX
## 1% -33209 -20580 -14224 -27848 -31873
## 5% -24146 -12305 -10890 -15073 -12472
```

Amazon has the least Value at Risk, which is not suprising due to the rest of the data computed. Costco stands with the most Value at Risk, which was also not very suprising due to its negitive skewness and low mean in the descriptive statistics section.

Part IV. Bivariate Graphical Analysis

1. Use the `pairs()` function to create all pair-wise scatterplots of returns. Comment on the direction and strength of the linear relationships in these plots.

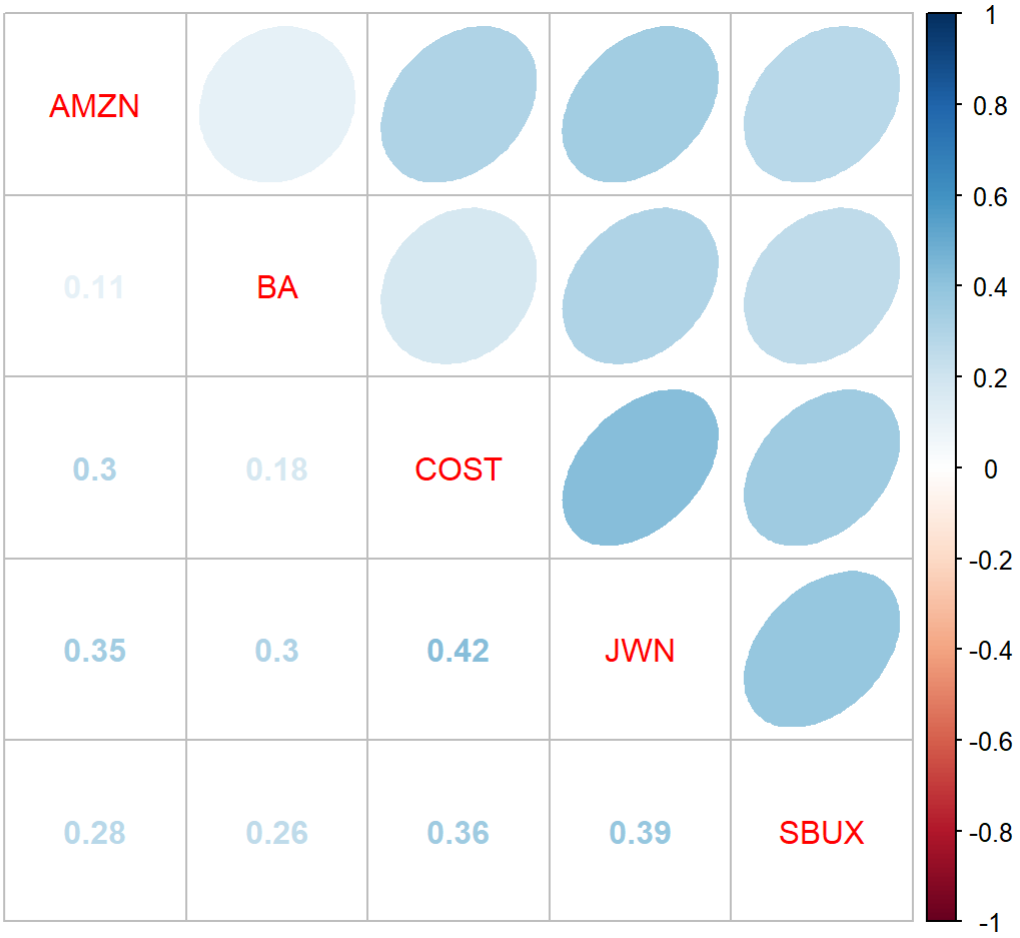
```
pairs(coredata(fiveStocksRet),col="blue",cex.axis=1,cex=.5)
```



Most of the relationships between the stocks seem to be linear with positive correlation. The only relationships that stand out to me to be the least linear are Amazon and Boeing and Boeing and Costco. Although, all stocks seem to have positive correlation with eachother.

2. Use the **corrplot** function `corrplot.mixed()` to plot the correlation matrix of the returns on the five assets. Comment on what you see.

```
corrplot.mixed(cor(fiveStocksRet),lower="number",upper="ellipse")
```



This graph seems to show what I stated above with all graphs being positively correlated and having Amazon and Boeing and Boeing and Costco with the least positive linear correlation.

V. Bivariate Numerical Summary Statistics

Use the R functions `cov()`, and `cor()` to compute the sample covariance matrix and sample correlation matrix of the returns. Comment on the direction and strength of the linear relationships suggested by the values of the covariances and correlations.

```
cov(fiveStocksRet)
```

```
##          AMZN      BA    COST     JWN    SBUX
## AMZN 0.03405 0.00165 0.00411 0.00731 0.00530
## BA   0.00165 0.00719 0.00111 0.00293 0.00223
## COST 0.00411 0.00111 0.00538 0.00352 0.00270
## JWN  0.00731 0.00293 0.00352 0.01294 0.00455
## SBUX 0.00530 0.00223 0.00270 0.00455 0.01063
```

```
cor(fiveStocksRet)
```

```
##       AMZN    BA  COST   JWN  SBUX
## AMZN 1.000 0.106 0.303 0.348 0.278
## BA   0.106 1.000 0.178 0.304 0.255
## COST 0.303 0.178 1.000 0.422 0.357
## JWN  0.348 0.304 0.422 1.000 0.388
## SBUX 0.278 0.255 0.357 0.388 1.000
```

All the stocks have positive relationships shown by the covariance, with Amazon and Nordstrom having the highest. The sample correlation shows also that all stocks have some level of positive linear relationships with Costco and Nordstrom having the highest.

VI. Stylized Facts

1. Based on your analysis of the five Northwest stocks, what stylized facts do you see?

Within the data I feel that Sylized facts: M1, M2, M3, M5 and M6 are all shown within the data. For M1, all the stocks seem to show this. For M2, there is positive and negative skewness and kurtosis. For M3, Amazon shows this to be true. For M5, all of the stocks displayed positive correlation with none having negative correlation. For M6, all of the stocks seemed to show there was little evidence of linear time dependence in asset returns.