**Trading Strategy Test & Portfolio Ranking**

Mo Pei

Harvard University, Stats 107 Financial Statistics

Fall 2016 Professor: Michael Parzen

12/08/2016

## Abstract

The stock market has two major characteristics. One is return and the other is risk. Higher return usually associates with higher risk. Many researches and studies describe methods and theory to maximize profit and minimize lose. This project is to conduct three validations to verify different trading and investment ideas. I use training data to get a model of trading or a combination of investment portfolio. By applying test data, I could validate the trading strategy and rank different portfolios. Pairs-Trading applies the Mean-Reverting idea to trading strategy. By selecting an appropriate pair, I could achieve a high return, and the lost could be not canceled off. However, the criteria’s of choosing a pair does not always work. Besides, I took difference, ratio, and log ratio three trading strategies. Relative Strength Index, VWAP, and SMA are popular trading methods. By different time frames, SMA works better in the long run, RSI is good at short to middle range, and VWAP performs well in short period. Ranking different portfolios include benchmarks of Drawdown, DownsideDeviation, SemiDeviation , Omega, Sharpe Ratio, and Trenoy Ratio. The two portfolios I choose make profits and reduce risk in my test.

## Introduction

“The stock price is random walk” Eugen Fama said. The movie inside jobs tells story of 2008 financial crisis. It surprised me that the story begins at Iceland that I have never thought any relationship with finance. People lost jobs and companies went down. Lehman brothers founded in 1850 filed bankruptcy in 2008. Thereby, market has risk. So people fear about market? Maybe. Many hedge fund companies perform exceedingly impressive. John Pualson who is a HBS graduate is a hedge fund industry tycoon donated 400 million dollars to School of Engineering and Applied Science in 2015. He is one of the greatest hedge fund managers ever!

Trading Strategy and investment theory are interesting. My project tests three major ideas in the market including Pairs trading, Portfolio theory, and ranking of different trading strategy. For portfolio and pairs trading part, I focus on small size to middle size of stocks.

For pairs trading method, I want to research pairs selection. For example, by using by using Dickey–Fuller test (ADF) test and correlation, a pair would make profit and reduce risk. Furthermore, I use normalized ratio, difference, and log ratio three trading strategies. Finally, I take test data to validate the pair I chose.

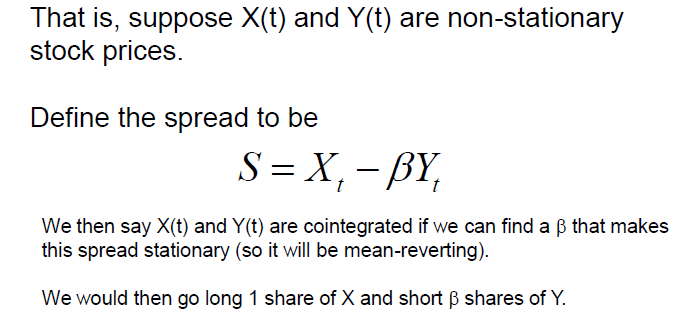
For portfolio investment method,

For comparison of different trading strategy, I checked trading strategies including Relative Strength Index, VWAP, and SMA. I observe their performance by different time frames. Also, I validated their profit and lost in bull and bear markets respectively.

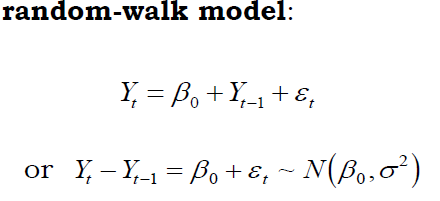
## Method

I first organized data by different conditions. The data is from Nasdaq and includes the information about stocks in NYSE. It has stocks symbols, last traded price, market cap, industry sectors, and other related information. Since I research small to middle cap company’s stocks, I limit the number of data by setting market cap between 6.25th percentile to 50the percentile. Also, because of moving average and pairs trading, I chose stocks’ last trade price greater than two dollars and the IPO year before 2008. Furthermore, I choose AAPL, M, and BA as data to study trading strategy.

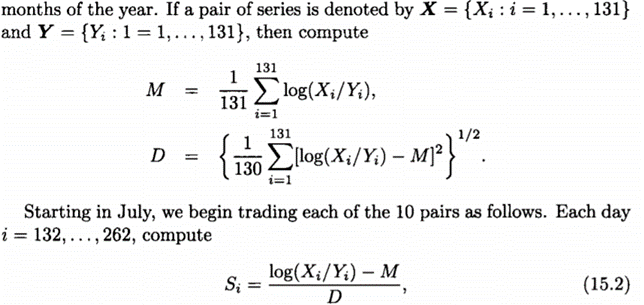
With respect to pairs trading part, the core concept is mean-reverting and the time series of spread is stationary. We need to find a pair of cointegrated stocks in order to implement mean reverting.



Since individual stock’s price is a random walk model we cannot use correlation to describe the linear relationship between stocks price movement. Time series AR(1) model saying when Beta1 smaller than 1 and greater than 0. The time series is mean reverting. By applying The Dickey-Fuller Test(adf), the null hypothesis is that the time series non-stationary. So in order to find a pair of cointegrated stocks, we can need reject the adf test.

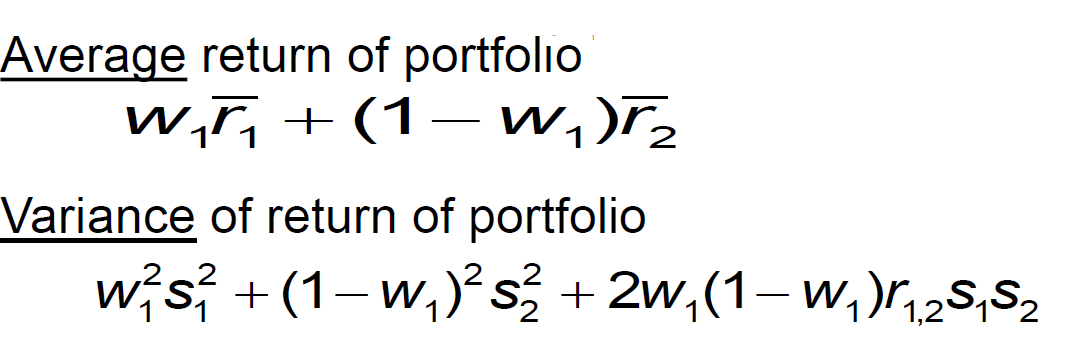


After we got a pair of stocks, we need to decide when to make a transactions. I test three ways including Brazilian difference, Pairlogs ratio, and log ratio method. First of all, all the data are normalized by subtracting a 14 days moving mean and then divide by 14 days moving standard deviation. The log ratio applies 130 days. For pairlogs ratio and Brazilian difference method, when normalized value exceeds 2, it means the ratio or difference becomes larger and would go back because of mean reverting theory. So at this time, it would short-sell. When the ratio is smaller than -2, it would be a buying signal. For log ratio, the trading is rule is similar but with log ratio. The formula below is a description of log ratio.

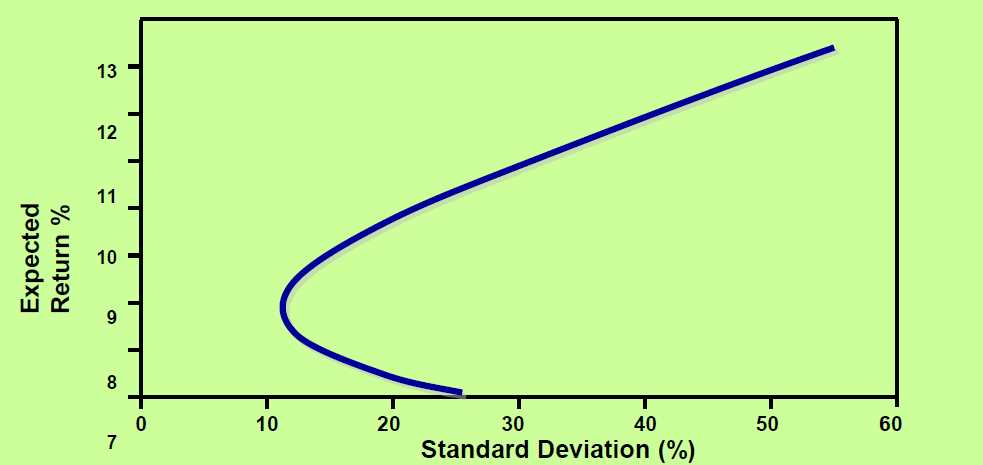


I used past data from last one year to last second year to rank pairs that I selected by running adf and correlation test. I checked their one year correlation and cointegration p-value graph. After all the analysis, I selected two pair of stocks, and use past year data to test the pair trading strategy’s performance.

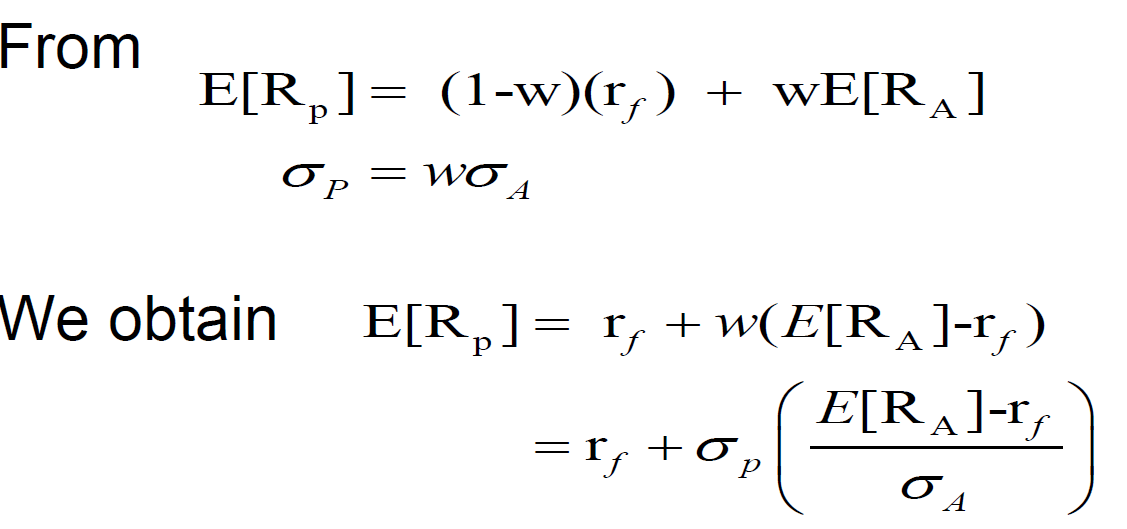
With respect to portfolio investment method, I apply MPT(modern portfolio theory ) to form an efficient portfolio to increase profit and reduce the risk. I computed the return of standard deviation of portfolios.

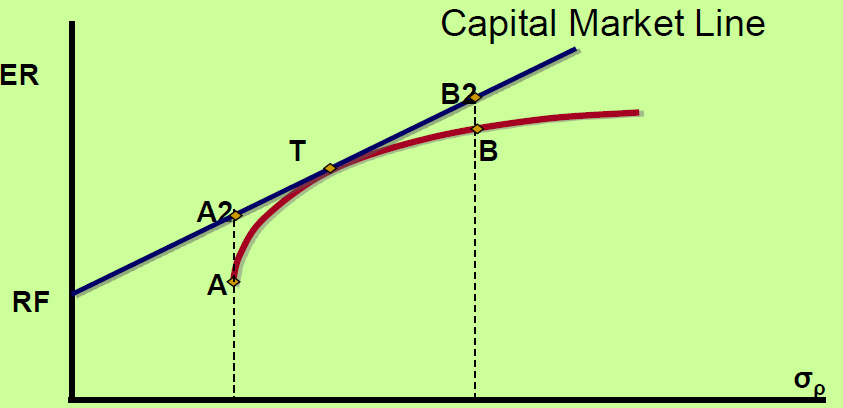


I calculated minimum variance of the portfolio and make a plot of efficient frontier.



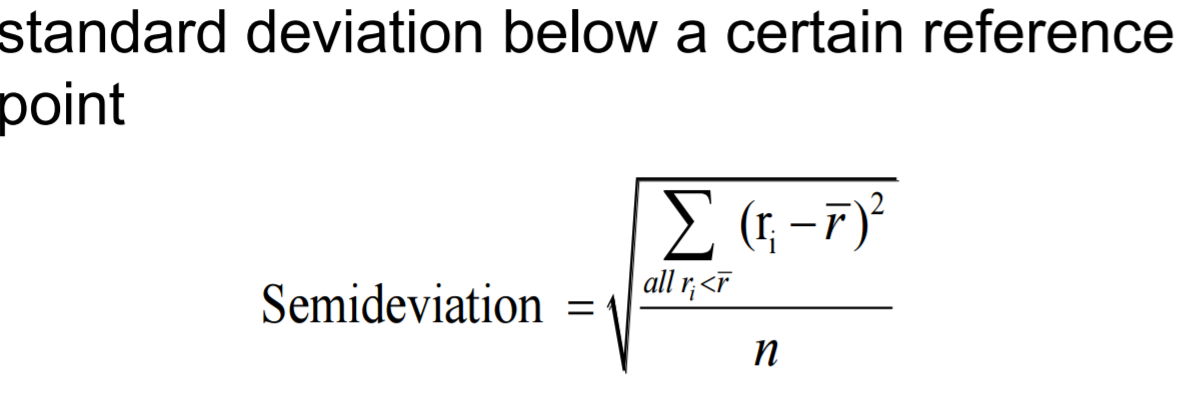
Plus, I added risk free asset to the portfolio to make capital market line and found the tangent portfolio.

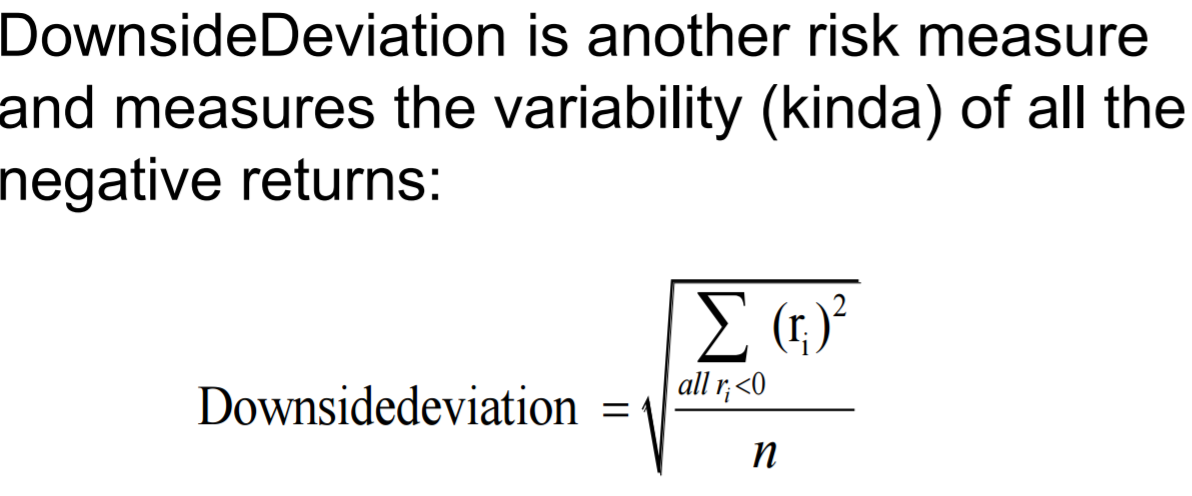


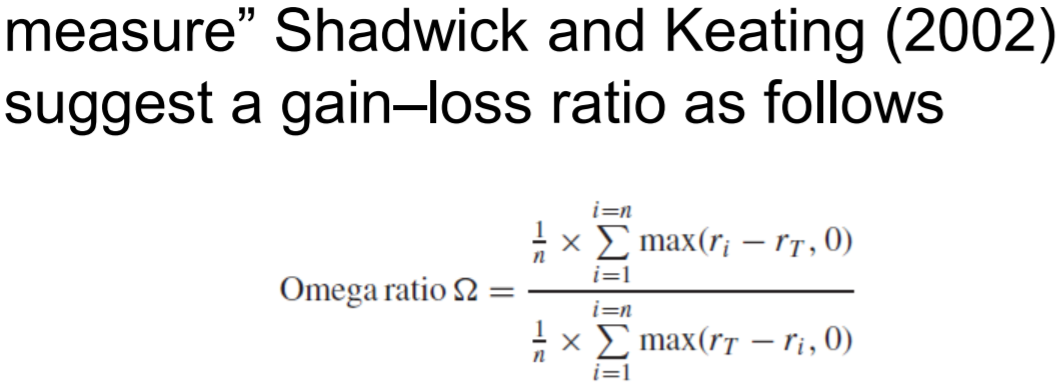


In order to generate a three stocks portfolios I use the small cap size data to randomly select stocks and get 115 combinations of stocks. I calculated the portfolio return, standard deviation, Sharpe ratio, and Trenoy ratio for all of 115 combinations. After compare those conditionss, I selected 4 combinations.

I used the following four ways to rank different portfolios.







A drawdown: is the peak-to-trough decline during a specific recorded period of an investment, fund or commodity. A drawdown is usually quoted as the percentage between the peak and the subsequent trough.

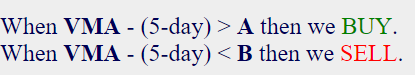
After choose portfolios, I use past year data to the portfolios’ performance.

For comparison of different trading strategy, I checked trading strategies including Bollinger band, Relative Strength Index, VWAP, and SMA. I observe their performance by different time frames. Also, I validated their profit and lost in bull and bear markets respectively.

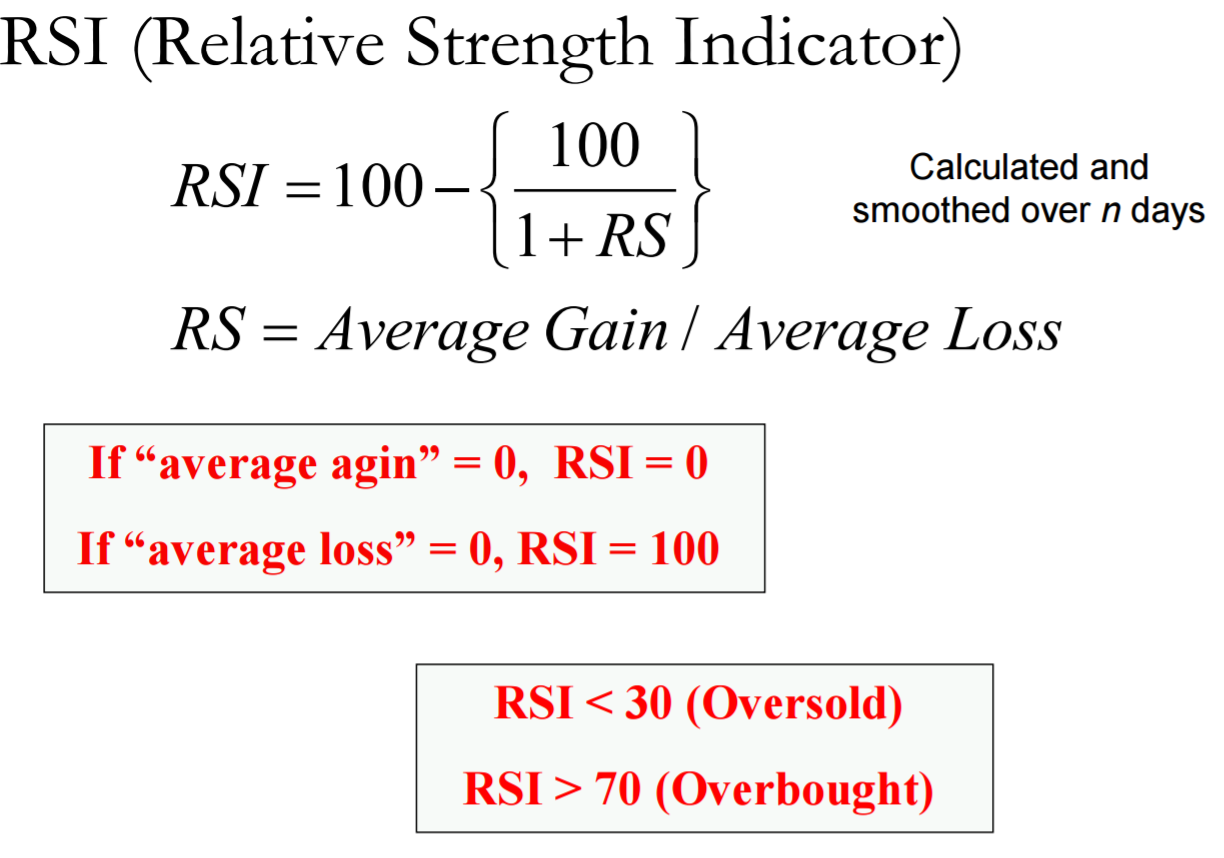
For simple moving average, when stock price above the moving average line, it would be a buy signal. When the price is below the average line, it would be sell signal.

For simple moving average, when stock price above the moving average line, it would be a buy signal. When the price is below the average line, it would be sell signal.

For VWAP method, it would be a buy signal when VWAP line is above SMA line, and would be a sell when is below SMA line.

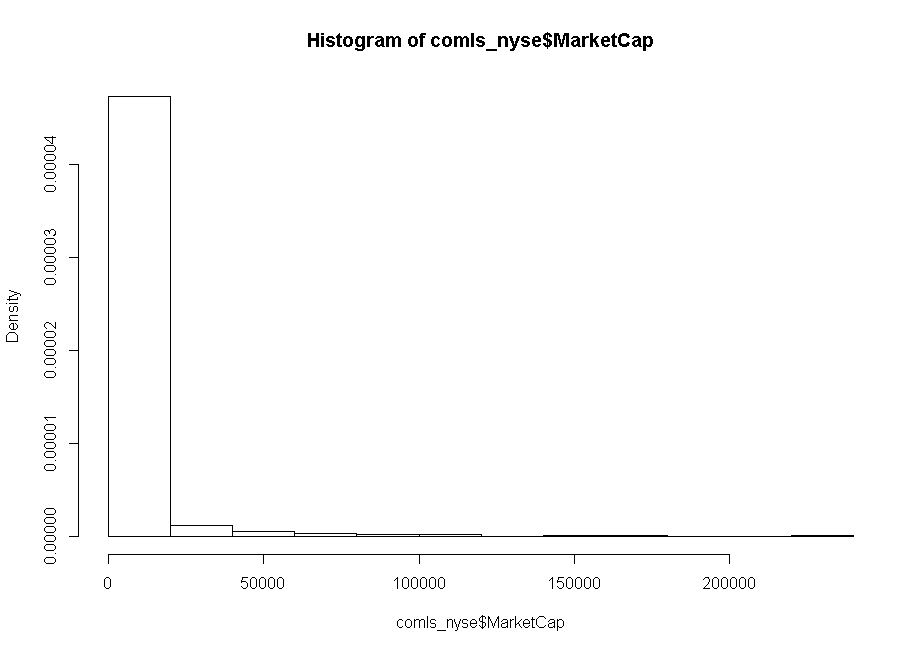


For Relative Strength Index method, it would be a buy when index smaller than 30, and it would be a buy when the index greater than 70.



# Results

Here is the distribution of market cap in NYSE. We can see that most of stock market cap under 50 billion dollars.



Results for pairs trading part, my training data is from

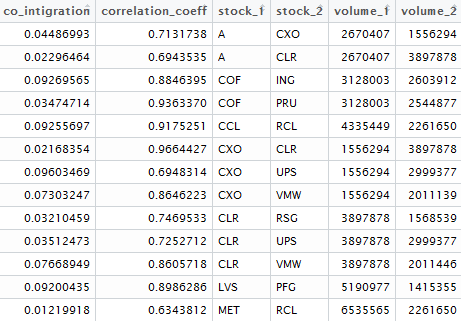


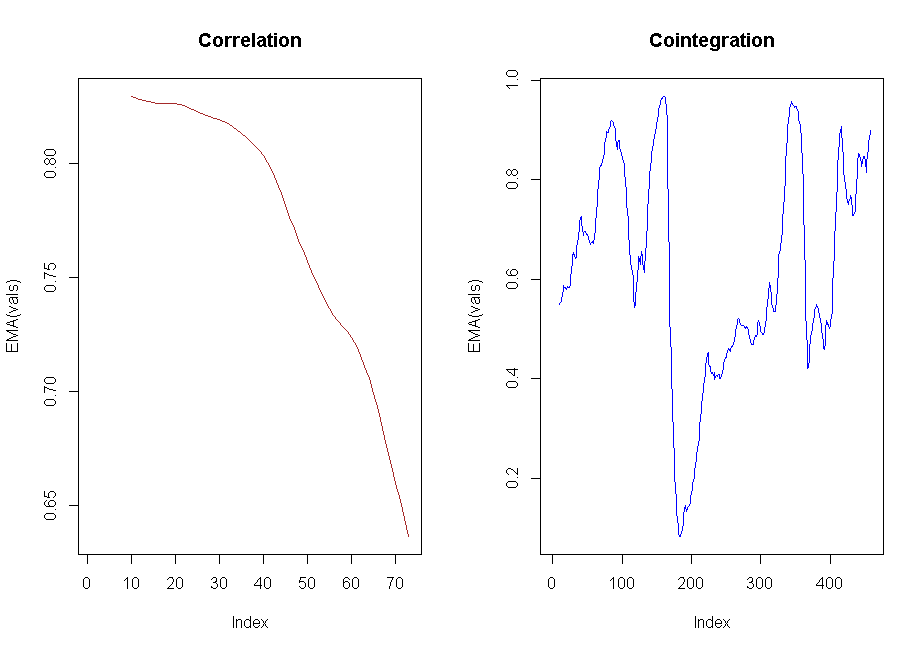
I run a simulation to get 515 combinations of pair stocks. Here is the list by selecting lowest p value(below showing 1-pvalue)

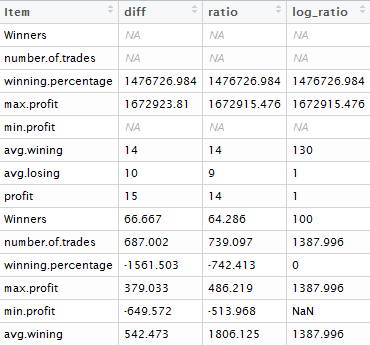
8

[1] "CXO"

[1] "VMW"



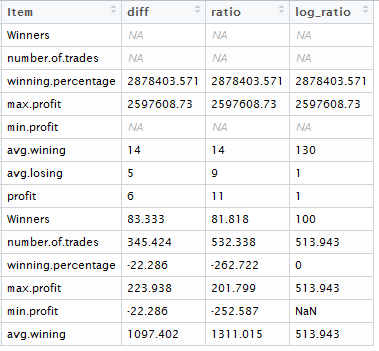


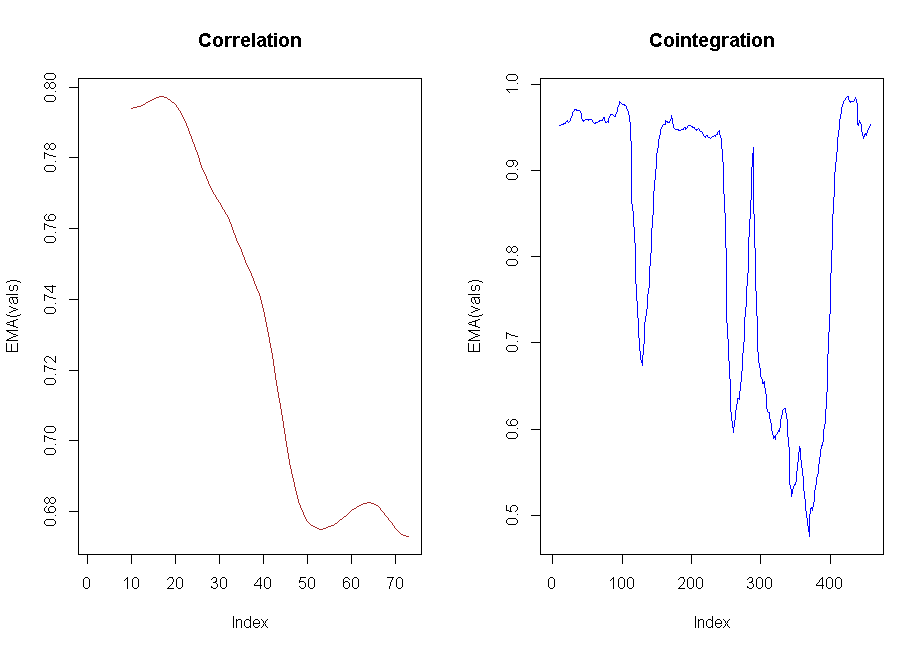


4

[1] "COF"

[1] "PRU"

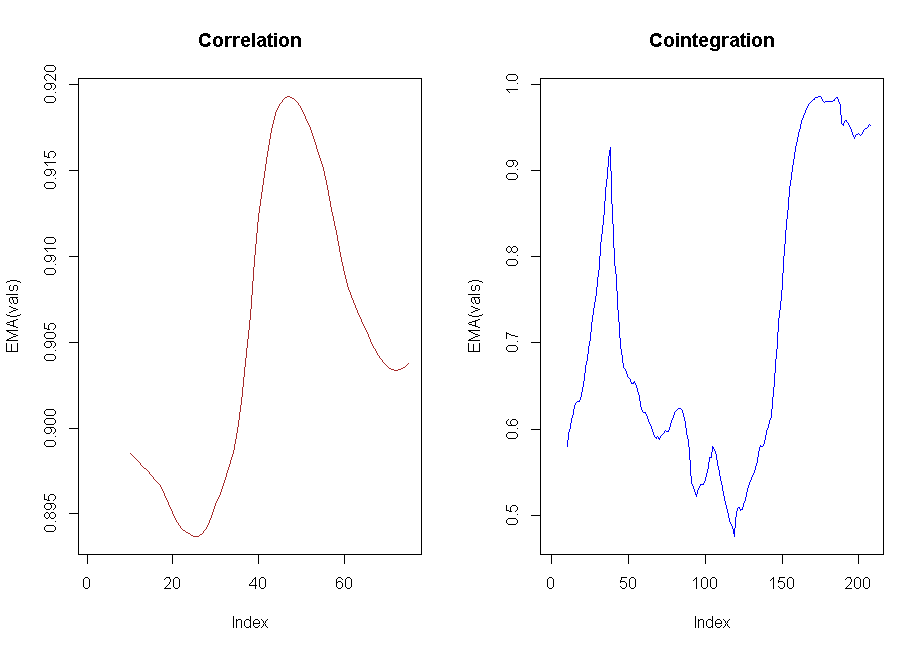


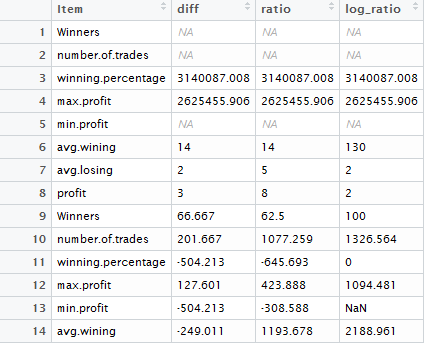


from = "2015-06-01"

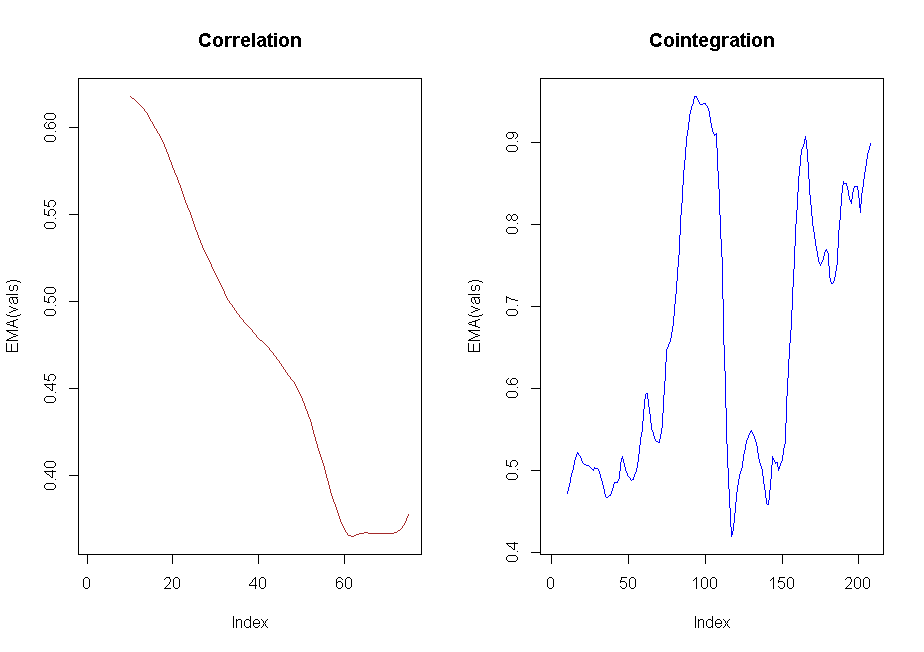
to= "2016-06-01"

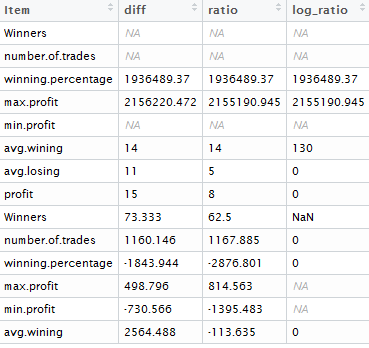
4



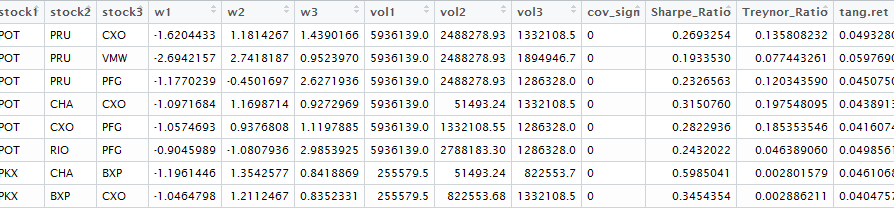


8





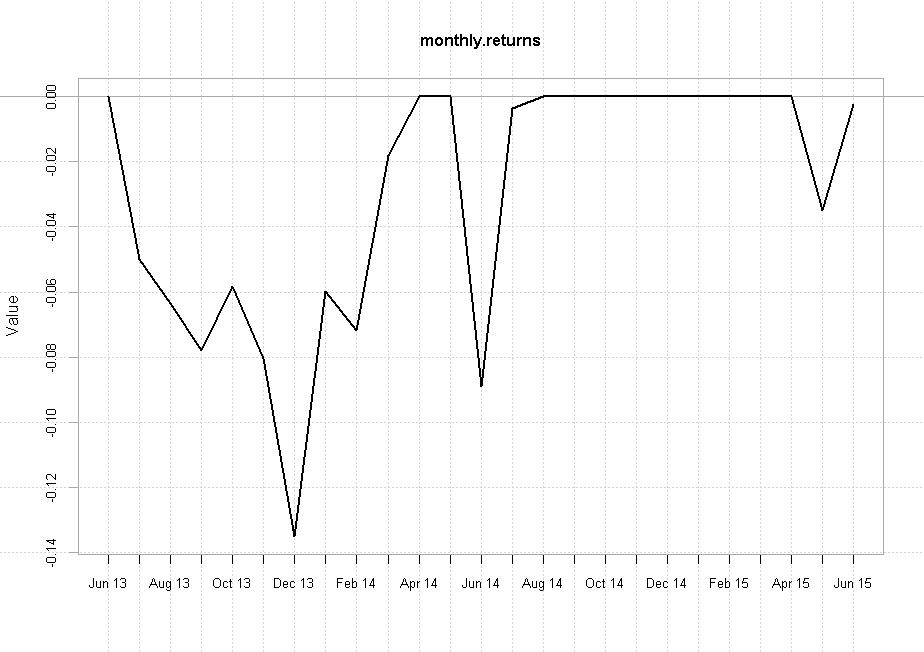
Portfolio



from="2013-06-01"

to="2015-06-01"

PKX CHA BXP -1.196145 1.354258 0.8418869



$omega

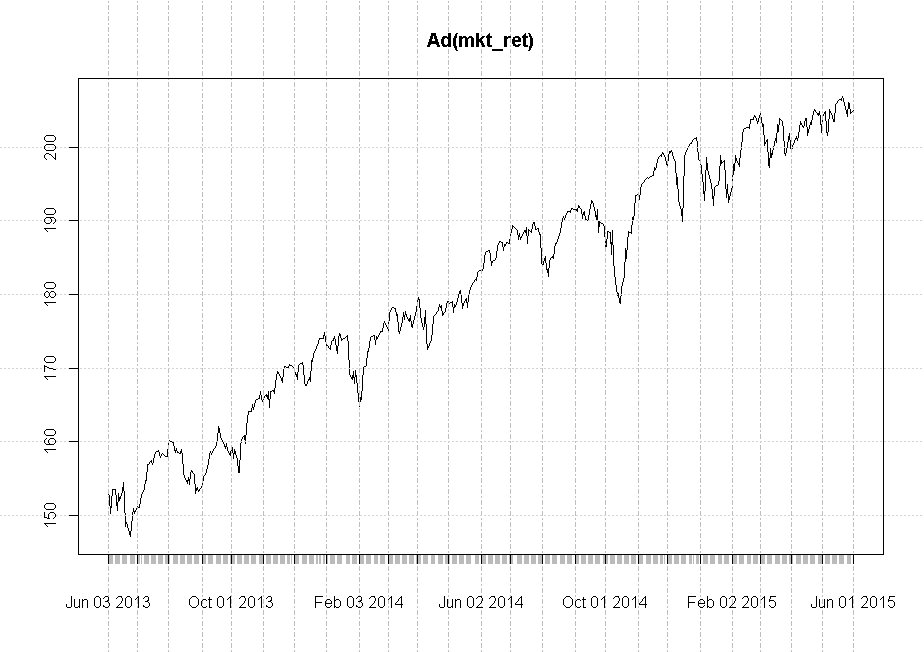
[1] 4.851772

$Downsidedeviation

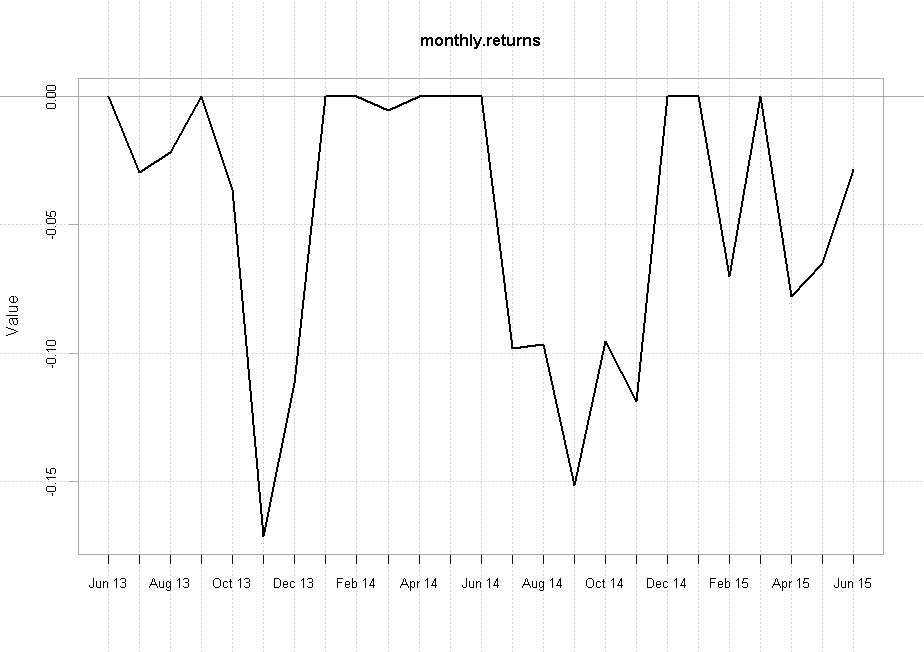
[1] 0.02554041

$SemiDeviation

[1] 0.04976204



PKX BXP CXO -1.04648 1.211247 0.8352331



$omega

[1] 2.86108

$Downsidedeviation

[1] 0.04323092

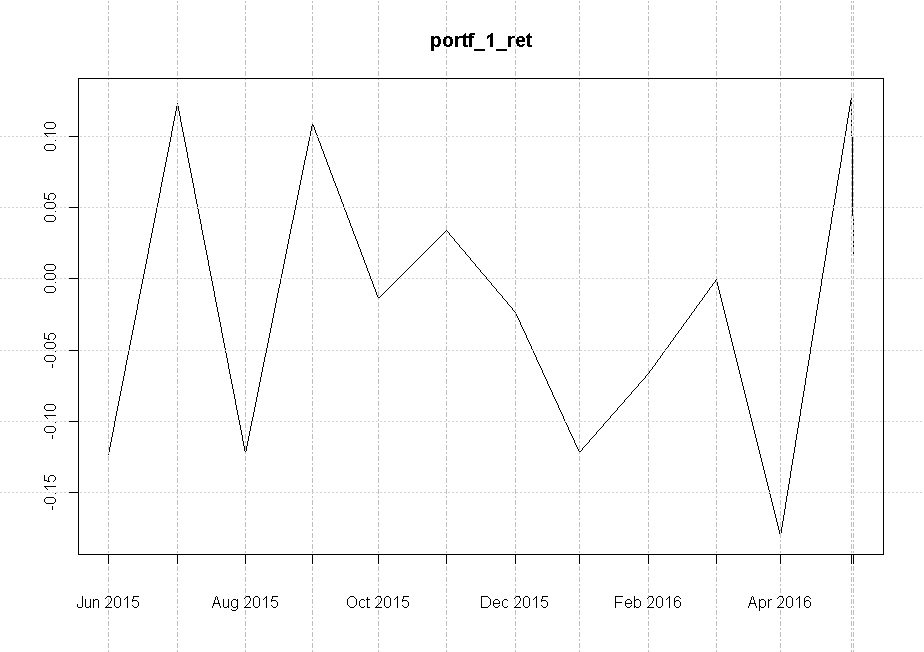
$SemiDeviation

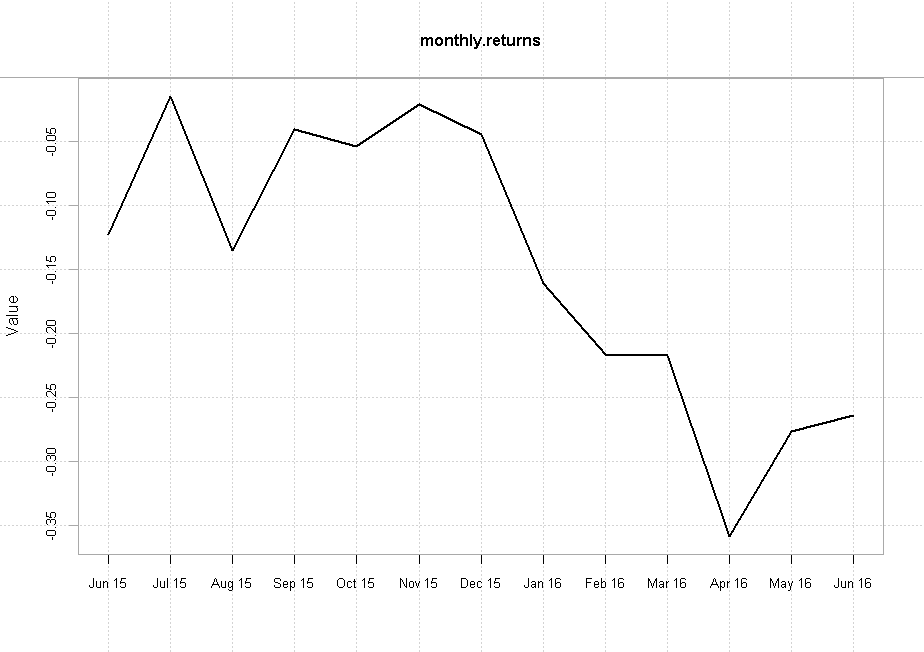
[1] 0.06641435

PKX CHA BXP -1.196145 1.354258 0.8418869

from="2015-06-01"

to="2016-06-01"





$omega

[1] 0.6312936

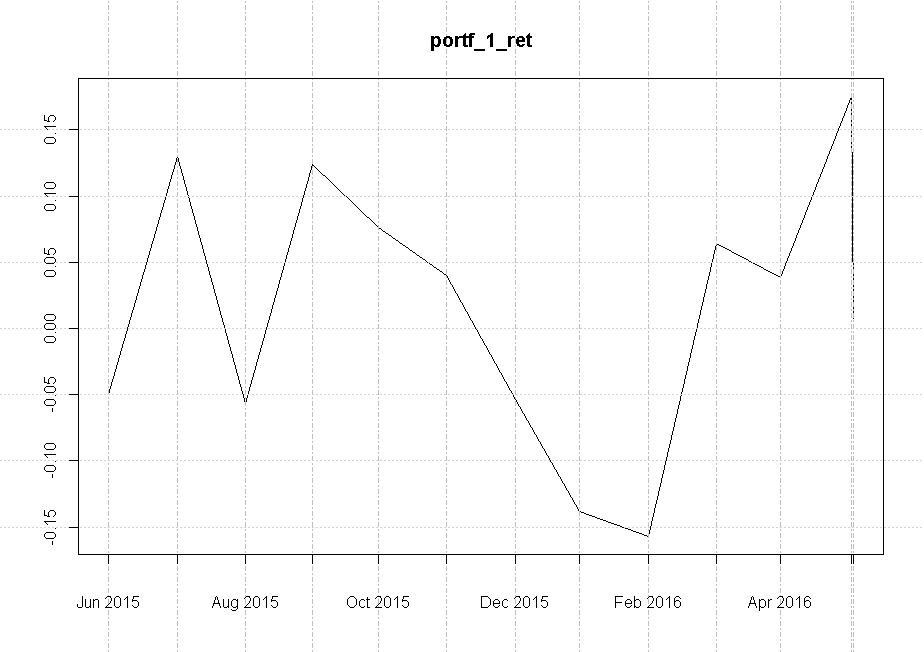
$Downsidedeviation

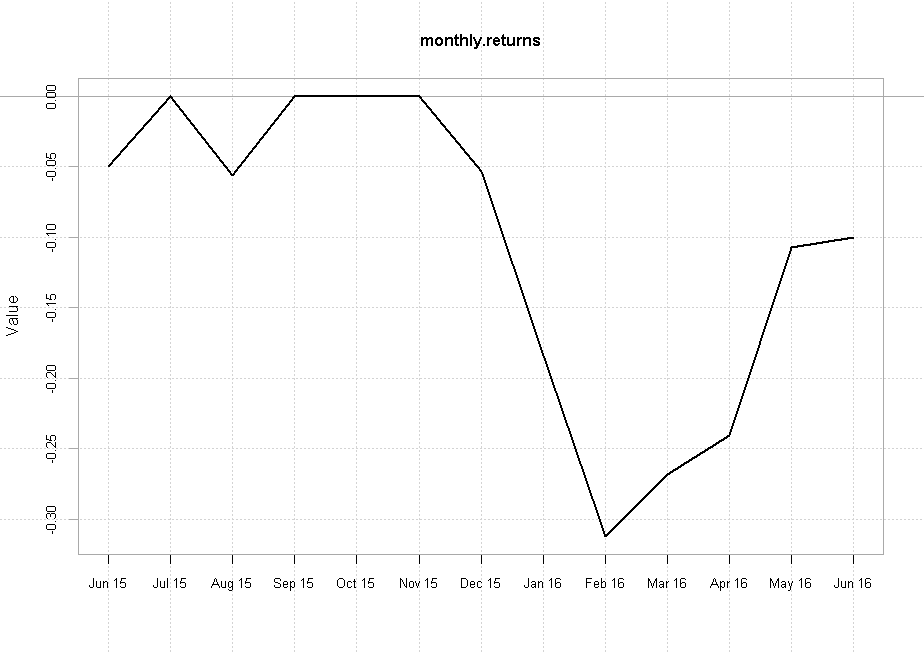
[1] 0.07981627

$SemiDeviation

[1] 0.06854298

PKX BXP CXO -1.04648 1.211247 0.835233





$omega

[1] 1.439156

$Downsidedeviation

[1] 0.0634595

$SemiDeviation

[1] 0.07211157

I made a function of computing configuration adf and correlation for all the 2 stocks combinations. I got total 595 combination of stocks. Finally, I picked up four stocks by limiting p value of adf testing smaller than 10%, volume of two stocks greater than 1,000,000, and correlation great than 0.6.

c) An “**Abstract**” of 100-300 words giving the following: Background, Methods, Results,

Conclusions. The abstract should be written in a style suitable for a general audience.

d) An “**Introduction**” section that states the focus of your analysis and ends with a sentence

stating the primary goal of this analysis, and any possible secondary goals.

e) A “**Methods**” section

f) A “**Results**” section

g) A “**Conclusions and Discussion**” section which should be written in the same “nonstatistical”

style as the abstract (see comments above for the abstract).

h) A “**References**” section that gives 2-5 references (such as software used, data from websites,

research paper you are reproducing, etc.)

i) An “**Appendix**” section. The Appendix may include any R or Excel code (and possible

discussion) that you feel is important in the analysis that you ran.