Final Project

Decision Support System for Portfolio Optimization

SYST 542 – Decision Support Systems Engineering

By

GROUP 3

Hemanth Kiran Dharmagadda Swamy

Sai Harsha Yagnamurthy

Ravi Teja Konda

Anirudh Myakala

Guided By

Professor Paulo C.G. Costa

CONTENTS

1. GOAL OF THE PROJECT

2. INTRODUCTION TO THE DECISION SUPPORT SYSTEM

3. FLOW CHART OF DECISION SUPPORT SYSTEM

4. INTERNAL PROCESS OF DECISION SUPPORT SYSTEM

5. INTRODUCTION TO MODERN PORTFOLIO THEORY

6. R MODEL : DATA SOURCE, STOCKS SELECTED AND PACKAGE USED IN R

7. PORTFOLIO OPTIMIZATION & BRIEF DESCRIPTION OF R CODE

8. RESULTS

9. FUTURE WORK

10. REFERENCES

GOAL OF THE PROJECT

The Goal of the Project is to make decisions based on the Decision Support System based on the user inputs. Once the inputs are submitted then the DSS access the database and optimize the results based on the user requirements. Comparing the Returns and Volatility of the Portfolio which contains few assets one selects based on the risk and return one is able to live with. In these project we will be learning how to maximize expected return subject to an upper bound on the risk, or to minimize the risk subject to a lower bound on the expected return.

Given a set of equities where I selected few assets randomly based on their beta values which would be derived from SML (4 Stocks), our code in R will calculate different efficient portfolios associated with risk and returns. It also involves calculating Mean Variance Portfolio (MVP) and Tangency Portfolio of the combination of all the assets.

We will also be calculating the beta of the portfolio to see how aggressive one is in selecting his portfolio.

The user can either split his investments based on equally distributed method or unequally distributed method.

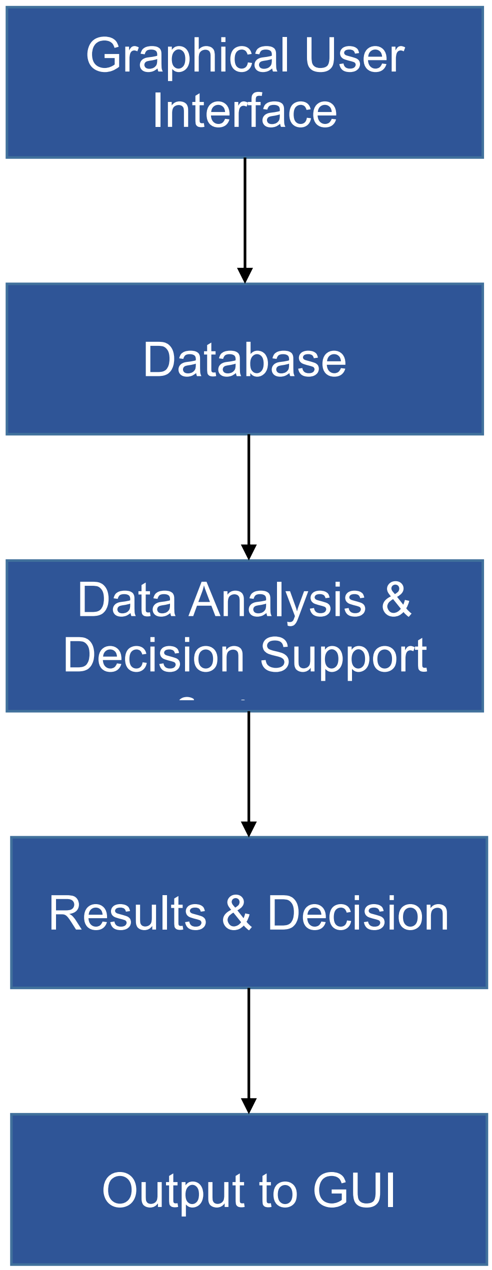
INTRODUCTION TO THE DECISION SUPPORT SYSTEM

Every user would have a secured login account. The user who wants to invest would definitely have a good knowledge on investment because they are investing their hard earned money. We can consider this as one of the assumption as well. So the user will enter his inputs in the dashboard. The input which the user enters would be the stock names as well as the percentage of investment. We have prepared a sample dashboard using some online web designing tools. Once after entering the asset names the user would need to give the percentage of investments on the entire portfolio. Portfolio in the sense all the assets which the user enters would be considered as a single or multiple portfolios. After entering the percentages the user should select the mode of investments. That is whether the user want to take some risk and invest on the risky assets or invest on the non-risky assets i.e., non- aggressive assets. We know that investing on the aggressive assets would be risky but at the same time if the asset price increases the returns would also be higher.

Once the user clicks on the "optimize" the portfolio then the data would be loaded in to the historical database. The present day adjusted values would be pulled from the different data centers such as Yahoo, Google in order to normalize the values. The data of the past history would be converted to the normalized values through log normal functions. The Entire process would be done in the backend process using the R statistical tools where we can connect the database to R through some secured packages. Once the values are normalized and based on the distribution entered by the user the R selects the algorithm. In the R package we have multiple algorithms. So we will be considering the algorithm based on the risk percentage entered by the user. We will be using several multiple libraries in order to use multiple algorithms.

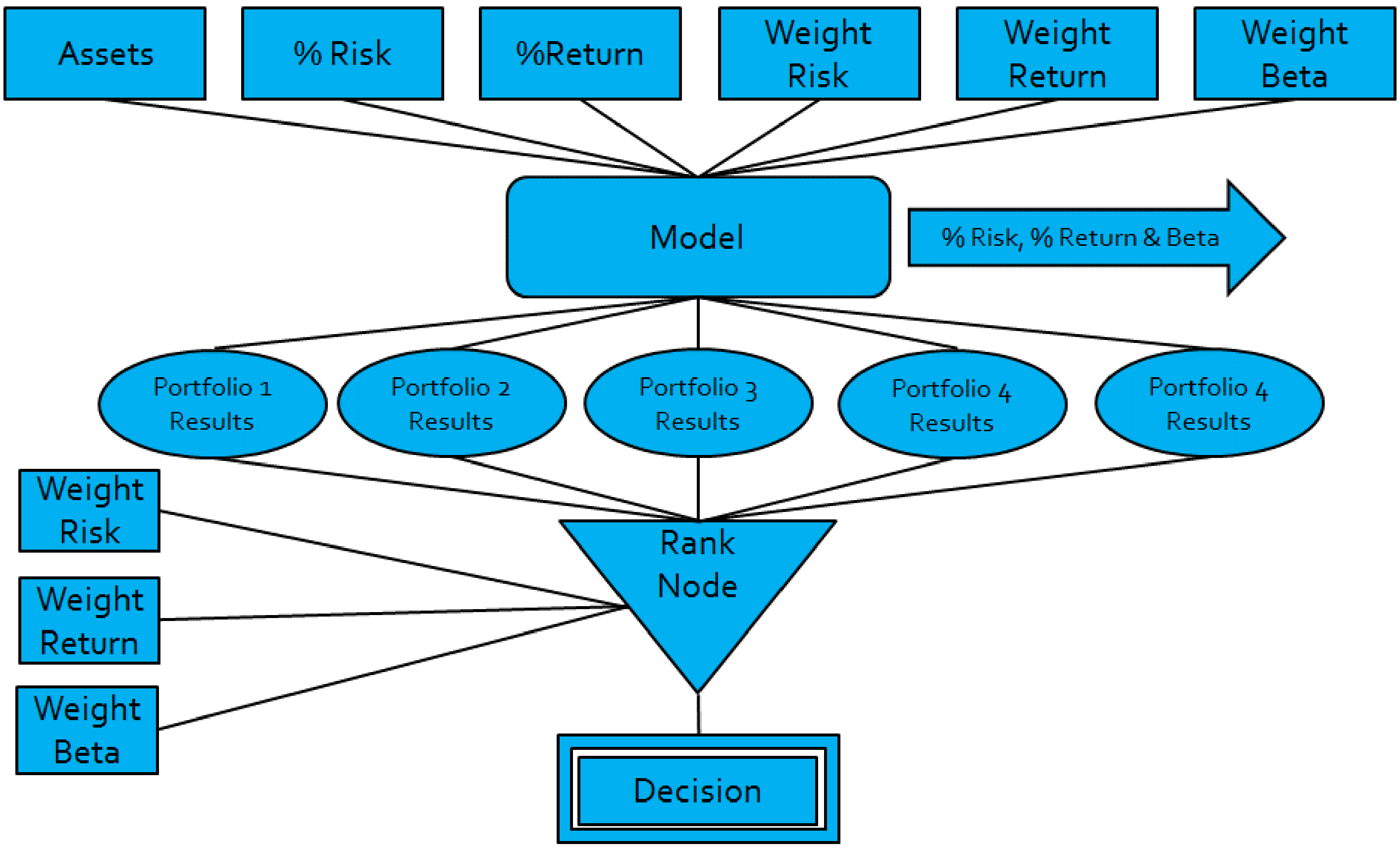
Data Analysis would be checked based on the normalized data. The points would be taken based on the one to one mapping for the efficient portfolio as well as the frontiers. Based on this the decisions would be pulled and populated to the dashboard where the user can see the results and based on the results the user can take decisions whether he can invest on this factor or he needs to change the risk factors or the assets. The Value at Risk as well as expected shortfall would be given so that user can have a picture of the distribution in his mind based on the confidence intervals as well.

Flow Chart of the DSS Process:



Internal Process of Decision Support System:

The Top line is the inputs entered by the user through the interface when the user submits after giving their inputs. The model block does the optimization of the portfolios by using different algorithms and with different probabilities after optimization it provides various results from different perspectives and provides the rankings accordingly. Based on the financial factors, the individual's investment returns the decisions can be made.



The individual is provided with a user interface where he can input different assets that he wish to invest along with the percentage return and risk or either of them and also the weights that he would like to allocate for different decision parameters in the decision support system. Here the Assets along with percentage return and risk are purely used in developing different kind of portfolios using the model in R. Once various numerous portfolios are developed with different percentages of Risk, Return and Beta values, the results are fed in to the Rank Node.

Rank Node is where our decision support system works. The Rank Node takes the results of various numerous portfolios generated using the model and ranks them according to the weights provided by the user at the beginning and returns the one which tops the list. Though this decision support system looks simple but it works almost all the times but when the user would like to have weights assigned very close then there may be a bias over the top three portfolios which may affect a bit on the DSS. To solve this we would like to add conditional probability theory which eliminates the bias which may be created during certain cases as our future work.

INTRODUCTION TO MODERN PORTFOLIO THEORY (Concepts of OR 538 – Analytics for Financial Engineering and Econometrics)

Almost everyone in the world is interested in investing in the Stock Market, but the question's which stops them in investing are

How should one invest his/her hard earned money?

What's going to happen to their investments In future?

What's the risk I am incurring & what's the return I am going to get?

Is it worth investing in stock market?

Portfolio Theory provides an answer to this questions based on two principles.

a. We want to maximize the expected return, and
b. We want to minimize the risk.

MODERN PORTFOLIO THEORY

Modern Portfolio Theory generally called MPT is a theory of finance that attempts to maximize portfolio expected return for a given amount of portfolio risk, or equally minimize risk for a given level of expected return, by carefully choosing the proportions of various assets.

It can also be explained as, according to the MPT, it's possible to construct an efficient frontier of optimal portfolios offering the maximum possible expected return for a given level of risk.

There are four basic steps involved in portfolio construction.

a. Security Valuation
b. Asset Allocation
c. Portfolio Optimization
d. Performance Measurement

The general objective in portfolio optimization is to make investments in assets (e.g. equities) so that overall portfolio provides good return, that is to say, it has a high growth rate per dollar invested while providing good protection against risk.

IMPORTANT CONCEPTS & DEFINITIONS INVOLVED IN PORTFOLIO CONSTRUCTION

Stochastic Process: Is a sequence of random variables and can be viewed as the "theoretical" or ""population" analog of a time series.

Time Series: A time series is a sequence of observations in chronological order. A time series can be considered a sample from the stochastic process.

Time Invariant Behavior: Behavior / Distribution of time series Y_1, Y_2……, Y_N is the same as that of Y_{M+1}, Y_{M+2},…., Y_{M+N}

Stationary Stochastic Processes: Are probability models for time series with time-invariant behavior.

There are two types of Stationary Processes, Strictly Stationary and Weakly Stationary. Strictly stationary are probability distributions that are time-invariant whereas weakly stationary are probability distributions with only the first two moments namely Mean, Variance & Covariance remain time-invariant.

When a time series is observed, a natural question is whether it appears to be stationary. This is not an easy question to address, and we can never be absolutely certain of the answer. However visual inspection of time series and changes in time series ca be helpful.

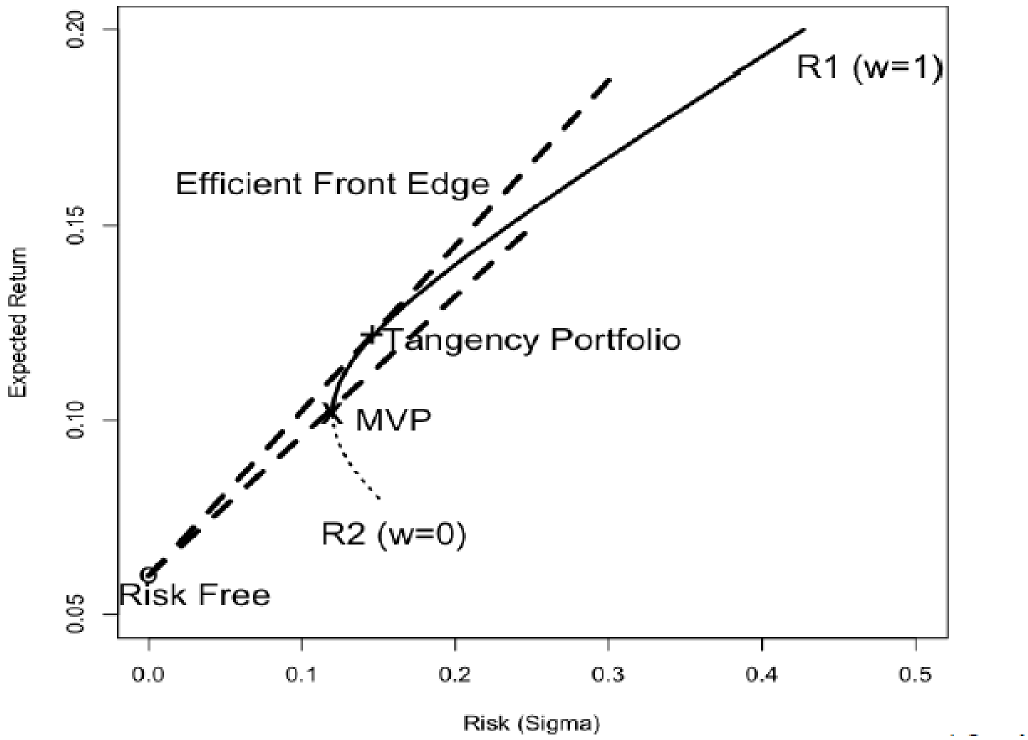
A time series plot of stationary series should show oscillation around some fixed level, a phenomenon called "mean-variance". If the series wanderers without returning repeatedly to some fixed level, then the series should not be modeled as a stationary series.

There are other ways to test like autocorrelation function (ACF) with test bounds and tests like augmented Dickey Fuller test etc. to test whether the time series are stationary or not.

Most statistical software will plot a sample ACF with test bounds. These bounds are used to test the null hypothesis that an autocorrelation coefficient is 0. The null hypothesis is rejected if the sample autocorrelation is outside the bounds. The usual level of the test is 0.05, so one can expect to see about 1 out of 20 sample autocorrelations outside the test bounds simply by chance.

The augmented Dickey Fuller test suggests that the time series is stationary if the null hypothesis of a uni-root is rejected

Some of the important concepts (Involved in Portfolio Optimization) of portfolios can be explained by the graph below which is a risk-return graph of a portfolio involving two risky assets. Concepts explained below are a combination of two risky assets and a risk free asset and same can be generalized for "N" no. of risky assets and one risk free asset.



The somewhat parabolic curve is the locus of values Sigma and Returns of the Portfolio. The left most point on this locus achieves the minimum value of the risk and is called "Minimum Variance Portfolio (MVP)". The points on the locus that have an expected return at least as large as the minimum variance portfolio are called "Efficient Frontier". Portfolios on the Efficient Frontier are called "Efficient Portfolios".

The Line that is tangent to the efficient frontier provides the highest return (slope) than any other line and the portfolio at this tangent point is called "Tangency Portfolio", where the slope of the line is called "Sharpe's Ratio" and is given by

$$S = \frac{E(R) - r_f}{\sigma_R}$$

The optimal or efficient portfolios mix the tangency portfolio with risk free asset. Each efficient portfolio has two properties:

1. It has higher expected return than any other portfolio with the same or smaller risk and
2. It has a smaller risk than any other portfolio with the same or higher expected return.

Thus we can only improve (reduce) the risk of an efficient portfolio by accepting a worse (smaller) expected return, and we can only improve (increase) the expected return of an efficient portfolio by accepting worse (higher) risk.

All efficient portfolios use the same mix of the two risky assets, namely, the tangency portfolio Only the proportion allocated to the tangency portfolio and the proportion allocated to the risk-free asset vary.

Given the importance of Tangency Portfolio, one should know how to find it.

$$w_T = \frac{V_1\sigma_2^2 - V_2\rho_{12}\,\sigma_1\sigma_2}{V_1\sigma_2^2 + V_2\sigma_1^2 - (V_1 + V_2)\rho_{12}\,\sigma_1\sigma_2}$$

Given expected means and standard deviations of the risky assets one can find the tangency portfolio weights for two risky assets using the above formulae for weight of 1^{st} risky asset and 1-1^{st} risky asset will give the weight of the 2^{nd} risky asset.

As I have mentioned earlier combining a tangency portfolio with risk free asset will gives us the efficient portfolio, the weights at that point are calculated using the below formulas.

$$E(R_P) = \omega\, E(R_T) + (1 - \omega)\, \mu_f$$
$$\sigma_P = \omega\, \sigma_T$$

Mostly expected returns and risk of an asset are calculated using the historical data assuming stationary time series (not the equity prices, but the returns on the equities) and assuming short term stationary is more reasonable. 1 to 2 years is representative for short term future prediction, long term prediction is not reliable.

An important concept called “Short Selling” is most important during portfolio optimization. Some weights in an efficient frontier may be negative, selling stocks without owing it. Selling short is a way to profit if a stock price goes down. To sell a stock short, one sells the stock without owning it. The stock must be borrowed from a broker or another customer of the broker. At a later point in time, one buys the stock and gives it back to the lender. This closes the short position.

CAPITAL ASSET PRCING MODEL (CAPM) & SECURITY MARKET LINE

A model for pricing an individual security or a portfolio. For individual securities, we make use of the Security Market Line (SML) and its relation to expected return to show how the market must price individual securities in relation to their security risk class. The SML enables us to calculate the reward-to-risk ratio for any security in relation to that of the overall market.

If the market portfolio M is efficient, the expected return of any asset “I” satisfies

$$\mu_i = \mu_f + \beta_i(\mu_M - \mu_f)$$

$$\beta_i = \frac{\mu_i - \mu_f}{\mu_M - \mu_f} = \frac{\sigma_{iM}}{\sigma_M^2} = \rho_{iM}\frac{\sigma_i}{\sigma_M}$$

Where beta (i) measures Riskness of the asset in other words how aggressive the asset is. Below is how beta value measures the aggressiveness of the asset.

Beta > 1 → Aggressive
Beta = 1 → Average risk
Beta < 1 → Not Aggressive

Total Risk is given by the below formula where the 1st component is called the market risk which is also called the symmetric risk (which cannot be diversified) and the 2nd component is the unique risk associated with the asset called unsymmetric risk (which can be diversified).

$$\sigma_i^2 = \beta_i^2\sigma_M^2 + \sigma_{\varepsilon,i}^2$$

$\beta_i^2\sigma_M^2$: Market risk (systematic risk)

$\sigma_{\varepsilon,i}^2$: Unique risk (unsystematic risk)

R- Code : DATA SOURCE, STOCKS SELECTED P AND PACKAGE USED IN R

I have used "GOOGLE FINANCE" as my data source to get the historical data from 1st Oct 2012 to 20th May 2015

The equities that I have considered are "Google", "Wal-Mart", "Bank of America" and "Microsoft". The package that we used for my portfolio optimization in R is "fPortfolio". There are other packages like "Portfolio", "performanceAnalytics" which can be used for portfolio optimization, but "fPortfolio" is exclusively designed for Financial Engineering and Computational Finance. It has very powerful collection of functions to optimize portfolios and to analyze them from different points of view.

Some of the functions which we have used in construction of my project work are like "portfolioFrontier () – Plots the Frontier Curve of all the assets included in the portfolio", "plot(portfolioFrontier ()) – This is an interactive plot having 8 options to plot", "tangencyPortfolio () – which returns the weights at tangency portfolio", "getPortfolio (), "getWeights ()", "getStatistics ()", "periodReturn () – Used for calculating returns on equities", etc.

PORTFOLIO OPTIMIZATION AND BRIEF DESCRIPTION OF R CODE

Portfolio Optimization can be done using the R Code provided below where the efficient frontier of the portfolio is seen.

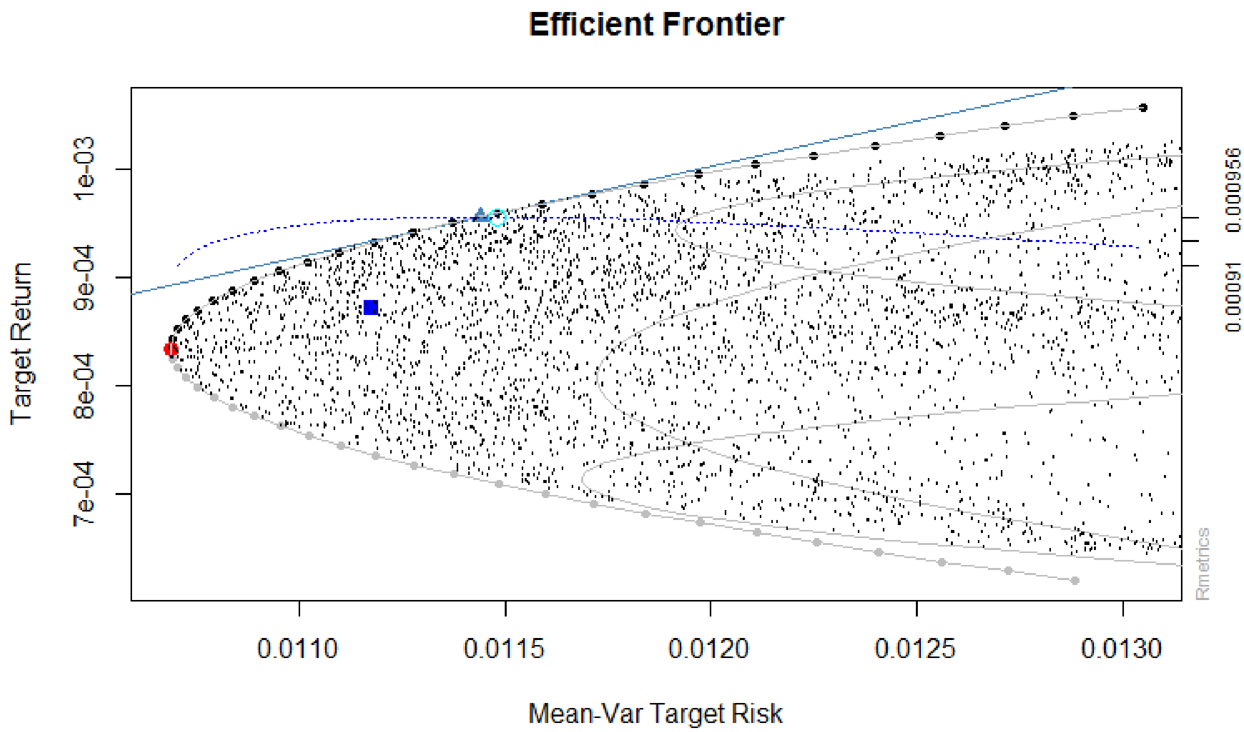
R Code:

```
getSymbols("GOOG",from = "2012-10-01", to = "2015-05-20", src = "google")
GOOG.Clo <- GOOG[,4]
GOOG.Ret <- periodReturn(GOOG.Clo, period = "daily")
GOOG.Mean=mean(GOOG.Ret)
GOOG.StdDev=sd(GOOG.Ret)
getSymbols("BAC",from = "2012-10-01", to = "2015-05-20", src = "google")
BAC.Clo <- BAC[,4]
BAC.Ret <- periodReturn(BAC.Clo, period <- "daily")
getSymbols("MSFT",from = "2012-10-01", to = "2015-05-20", src = "google")
MSFT.Clo <- MSFT[,4]
MSFT.Ret <- periodReturn(MSFT.Clo, period = "daily")
getSymbols("AMZN",from = "2012-10-01", to = "2015-05-20", src = "google")
AMZN.Clo <- AMZN[,4]
```

```
AMZN.Ret <- periodReturn(AMZN.Clo, period = "daily")
getSymbols("^IRX",from = "2012-10-01",to = "2015-05-20")
IRX1 <- IRX
head(IRX1)
IRX1 <- IRX1[,6]
head(IRX1)
IRX1 <- IRX1/(100*253)
head(IRX1)
mean(IRX1)
riskFreeRate <- mean(IRX1)*253
riskFreeRate # Yearly Rate

Portfolio.Stocks <- as.timeSeries(cbind(GOOG.Ret,BAC.Ret,MSFT.Ret,AMZN.Ret))
colnames(Portfolio.Stocks) <- c("GOOG.Ret","BAC.Ret","MSFT.Ret","AMZN.Ret")
head(Portfolio.Stocks)
StocksList=c("GOOG.Ret","BAC.Ret","MSFT.Ret","AMZN.Ret")
Spec <- portfolioSpec(portfolio=list(riskFreeRate=mean(IRX1),nFrontierPoints=50))
Const1 <- c("minW[1:length(StocksList)]=0.1")
Const2 <- c("minW[1:length(StocksList)]=0.2")
Const3 <- c("minW[1:length(StocksList)]=-1")
Const4 <- c("maxW[1:length(StocksList)]=0.05")
Frontier <- portfolioFrontier(Portfolio.Stocks,spec = Spec, constraints = Const3)
#plot(Frontier,xlimit=range(0,0.2))
allocations <- getWeights(Frontier@portfolio)
# Tangency Portfolio
tangency.portfolio <- tangencyPortfolio(Portfolio.Stocks, spec = Spec, constraints = Const3)
tangency.portfolio
tangency.portfolio.weights=getPortfolio(tangency.portfolio)$weights
tangency.portfolio.weights
# Mean / Return at Tangency Portfolio
mean.tangency.portfolio=getPortfolio(tangency.portfolio)$targetReturn
```

```
return.TP=mean.tangency.portfolio[1]*253 # Converting it to yearly return
return.TP
# variance or Risk at Tangency Portfolio
sigma.tangency.portfolio=getPortfolio(tangency.portfolio)$targetRisk
volatality.TP=sigma.tangency.portfolio[2]*sqrt(253) # Converting it to yearly risk
volatality.TP
# Calculating Weights of Tangency Portfolio & Risk Free Asset given risk=10%/year
sigma.portfolio=0.10
riskFreeRate
weight.at.tangency.portfolio=sigma.portfolio/volatality.TP
names(weight.at.tangency.portfolio)=c("Weight at Tangency Portfolio")
weight.at.tangency.portfolio
tangency.portfolio.return1=getPortfolio(tangency.portfolio)$targetReturn
tangency.portfolio.return1
tangency.portfolio.return=tangency.portfolio.return1[1]*253
tangency.portfolio.return
return.portfolio=(weight.at.tangency.portfolio*tangency.portfolio.return)+
  ((1-weight.at.tangency.portfolio)*riskFreeRate)
names(return.portfolio)=c("portfolio return")
return.portfolio
Weight.at.riskfreeassest=1-weight.at.tangency.portfolio
names(Weight.at.riskfreeassest)=c("Weight.at.riskfreeassest")
Weight.at.riskfreeassest
Weights.of.stocks.TP=weight.at.tangency.portfolio*tangency.portfolio.weights
Weights.of.stocks.TP
# Calculating Weights of Tangency Portfolio & Risk Free Asset given return=21.15276%/Year
mean.portfolio=0.2115276
riskFreeRate
tangency.portfolio.return1=getPortfolio(tangency.portfolio)$targetReturn
tangency.portfolio.return1
tangency.portfolio.return=tangency.portfolio.return1[1]*253 # Yearly Return
```

```
tangency.portfolio.return

weight.at.tangency.portfolio=(mean.portfolio-riskFreeRate)/(tangency.portfolio.return-
riskFreeRate)

names(weight.at.tangency.portfolio)=c("weight of Tangency Portfolio")

weight.at.tangency.portfolio

Weight.at.riskfreeassest=1-weight.at.tangency.portfolio

names(Weight.at.riskfreeassest)=c("Weight.at.riskfreeassest")

Weight.at.riskfreeassest

Weights.of.stocks.TP=weight.at.tangency.portfolio*tangency.portfolio.weights

Weights.of.stocks.TP

# Calculating Beta of portfolio which tells you how aggressive your portfolio


getSymbols("^GSPC",from="2012-10-01",to="2015-05-20",src="yahoo",colnames=0)

SP500.Clo=GSPC[,6]

SP500.Ret=periodReturn(SP500.Clo,period="daily")

head(SP500.Ret)

adf.test(SP500.Ret)

mean(SP500.Ret)

market.returns=mean(SP500.Ret)*253 # Converting daily rteurns to yearly returns of S & P 500

market.returns

market.risk=sd(SP500.Ret)*sqrt(253) # Converting daily risk to yearly returns of S & P 500

market.risk

Beta.Google=(GOOG.Mean-riskFreeRate)/(mean(SP500.Ret)-riskFreeRate)

Beta.Google

Beta.Portfolio=((mean.portfolio/253)-riskFreeRate)/(mean(SP500.Ret)-riskFreeRate)

Beta.Portfolio
```



RESULTS

We tried to develop a prototype of the Graphical User Interface and can be seen at http://harshayagnamurthy.wix.com/start-from-scratch. Though we have not developed a connection to the database with the GUI we developed but we tried to analyze three different portfolios when a user has input the risky assets as (Google, Microsoft, Walmart, Amazon & Risk Free Rate) and the risk or volatility that he can live with as 10%.

The results of the three different portfolios can be seen below in the table.

Portfolio Constraint	Risk (%)	Return (%)	Beta
Portfolio 1 (Minimum 10%)	**10%**	**13.20**	**1.44**
Portfolio 2 (Minimum 20%)	**10%**	**12.90**	**1.44**
Portfolio 3 (Minimum -100%)	**10%**	**13.30**	**1.44**

As said there will be numerous portfolios that are developed when the user has his inputs in place in GUI and let's think that the above are from those numerous portfolios generated and see how our decision support system (**Rank Node uses Ranking Algorithm**) works here in ranking the portfolios. If the user has more weight in the return percentage the rank node orders the portfolios in descending order with the highest return portfolio on top and returns that portfolio weights as the output along with return, risk percentages and beta value. The same if the beta has go the highest weight.

Now let's look at the three portfolios where we have given huge weight for the beta value and one can see there are same beta values here and now what the rank node does is it looks for the second highest weight and arranges them accordingly and returns the results. So in our portfolio the second weight is for returns so will return the portfolio 3 as the best portfolio.

FUTURE WORK

We would like to develop more on the Decision Support System as the present decision support system is purely based on the ranking (Rank Node) of various numerous portfolios developed during the process, based on the weights provided by the user. Though it works most of the cases but at some point there may be bias if the weights provided by the user are very nearer to each other. We would like to develop a decision support system based on the decision trees and using conditional probability which may be more sound system that the present one.

REFERENCES

OR 538 - Statistics and Data Analysis for Financial Engineering & Econometrics by KC Chang (Class Notes)

Statistics & Data Analysis for Financial Engineering by David Ruppet

OR 538 – Project on Portfolio Optimization

www.rmetrics.org

"fPortfolio" package document downloaded from www.rmetrics.org

www.stackoverflow.com