PORTFOLIO CONSTRUCTION FOR DIFFERENT COMMUNITIES

GR5261 Final Project
Instructor: Zhiliang Ying

Leanne Li (lyl2118) report writing: intro, dataset, stock selection, conclusion
Sizhu Chen (sc4248) coding: asset allocation & portfolio assessment
Mengjia Huang (mh3781) report writing: asset allocation & portfolio assessment
Yuchen Ma (ym2644) coding : asset allocation & optimal portfolio construction
Wenjie Jing (wj2271) report writing: optimal portfolio construction
Weibo Zhang (wz2353) coding: risk management
Shenzhi Li (sl4250) report writing: risk management

Introduction & Motivation

After the 2008 financial crisis, many investors who used to invest most of their wealth in stocks have become aware of riskiness of the stock market. Many start to allocate their assets into less risky assets such as Treasury bonds in order to eliminate their investment risk. How they allocate their assets depend on their investment choices and their financial backgrounds. Aggressive investors choose to take the high volatility while risk-averse investors tend to put less money in the stock market. But how exactly should they make their choices to achieve the optimal portfolios they want?

Our group is passionate about helping these investors make their choices. In this project, we will focus on examining different investment strategies and constructing portfolios for different types of investors. Our goal is to make all investors happy by providing them investment plans that fit their characteristics the best. To do so, we will first construct the tangency portfolio of stocks by utilizing mean-variance analysis and risk parity. We will compare the portfolios created by the two techniques and evaluate the resulting portfolio. After evaluating and coming up with the best tangency portfolio, we are going to mix it with risk-free assets to find the best portfolios that can fit different types of investors well. At the end, we will evaluate our mixed portfolios using risk management strategies to see how we have done.

Dataset

The datasets we will be using in this project are S&P 500 Index and T-bill. We obtain our dataset from Yahoo Finance, taken time period from April 23rd, 2013 to April 20th, 2018 (daily adjusted close price). We use S&P500 because it is one of the most common followed equity indices and many people consider it as one of the best representations of the U.S. stock market. For T-bill, since it has short maturities and no credit risk, we will need it to balance our portfolio and negate stock market volatility. Software we use for this project is R.

Stock Selection

In order to construct our portfolio, we will begin with cleaning the dataset of S&P 500 Index. We first calculate the expected return and standard deviation of each stock using the adjusted close price. Then we find Sharpe ratio of each stock using:

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

We will then rank the 500 stocks from the highest to the lowest by their Sharpe ratio values. Also, we delete the stocks which have PE (Price-Earnings) Ratio larger than 100. After the deletion, we keep the top 50 companies with the largest Sharpe Ratios for our portfolio.

Asset Allocation

Finishing with cleaning the dataset, now we are ready to begin the process of asset allocation. Choices relating to asset allocation have major effect on the risk/return characteristics of investment portfolios. We utilize the traditional mean-variance optimization methodology and risk parity to construct tangency portfolios with 50 stocks, 30 stocks, and 10 stocks. Then we choose one with the highest Sharpe Ratio.

Mean-Variance Analysis – Focus on allocation of capital

Traditional strategic asset allocation theory is deeply rooted in the mean–variance portfolio optimization framework developed by Markowitz [1952] for constructing equity portfolios. The idea behind the influential economic theory is that "risk-averse investors can construct portfolios to maximize expected return based on a given level of market risk" (Markowitz). However, Since the Modern Portfolio Theory assumes that investors are rational and risk-averse, the mean–variance optimization is difficult to implement due to "the challenges associated with estimating the expected returns and covariance for asset classes with accuracy" (Chaves).

Risk Parity – Focus on allocation of risk

"Starting in the late 2000s financial crisis, investment professionals increasingly began to favor risk-based solutions in determining their strategic asset allocation" (Molenaar). Risk parity became popular as a method of setting up portfolios in such a way that the different investment categories held in a portfolio made up an equal risk budget. In other words, the approach of building a risk parity portfolio is similar to creating a minimum-variance portfolio subject to the constraint that each asset contributes equally to the portfolio overall volatility. In this way, it can not only achieve a higher Sharpe Ratio, but it also can be more resistant to market downturns than the traditional portfolio.

How to obtain the Optimal Portfolio and What do we find?

Under the method of Mean-variance Optimization, we use *solve.QP* function to solve the quadratic programming problem:

$$\text{minimize } w^T \Sigma w \text{ subject to } \sum_{j=1}^{J} w_j = 1 \text{ and } \sum_{j=1}^{J} w_j \times \mu_j = \mu$$

where w_j is the weight, μ_j is the net return of asset j.

By solving the optimization problem, we have obtained the optimal weights of the portfolios with 50 stocks, 30 stocks, and 10 stocks (table 1-3) and then we plotted three efficient frontiers (figure 1-3)

On the other hand, under the method of Risk Parity, we find optimal weights (table 4-6) of the three portfolios using *nloptr* function to address a nonlinear optimization problem through gradient descent:

$$\arg\min_{w} \sum_{i=1}^{N} \left[w_i - \frac{\sigma(w)^2}{(\Sigma w)_i N} \right]^2$$

where $\sigma(w) = \sqrt{w'\Sigma w}$ is the volatility of the portfolio, $w_i = \frac{\sigma(w)^2}{(\Sigma w)_i N}$ is the weight of asset i, N is the total number of assets.

Usually, as the number of stocks in a portfolio increase, we'd consider that the expected return and volatility of the portfolio should decrease due to the effect of diversition. However, after obtaining these optimal portfolios, from table 7 we surprisingly find that under the mean-variance method, the expected return and volatility of 10 stocks are the smallest. This result violates our convention. After inspecting the component stocks and their weights, we find that this result could be interpreted as follow:

By applying mean-variance method to 10-stock-portfolio, six stocks whose weights are non-zero are NOC, NVDA, LMT, DXC, STZ and HII, which are the biggest corporations in aerospace, military technology and shipbuilding, government-related technology and wine production industries. Big corporations in these industries are extremely stable no matter in terms of their gross revenues or of their stock prices. Even business cycles or other special events cannot be able to cast a huge effect on their operations. As we expand the number of component stocks in our portfolio, we include more leading corporations or corporations that have good performance these years in other industries. There are two possible reasons that may lead to drop of portfolio's expected return and volatility. First, most of these corporations are in high-tech related and financial services industries, such as AMZN, TTWO, FB, PGR and CME. These stocks may have a very high correlation among each other, which may not help the portfolio diversified, but even make it worse. Secondly, these industries, although performing well these years, are not as stable as the those in 10-stock-portfolio. They are sensitive to financial crisis, technology progress and political events. For example, the stock price of Facebook has dropped dramatically after the exposure of the Privacy Scandal; whenever Apple launch new products, its stock price will change according to the expectation of customers towards the new product.

Evaluate our optimal tangency portfolio

According to the standard that a Sharpe ratio higher than 2 is rated as very good, the performance of all our six portfolios is great (table 8). The portfolio with 50 stocks and obtained under mean-variance optimization methodology has the highest Sharpe ratio 2.7642, thus we choose it to assess and then mix it with risk-free asset to find the best investment portfolio for different people.

Besides Sharpe Ratio, Treynor Ratio and Jensen's Alpha are also used to calculate risk adjusted return. The Treynor Ratio of our optimal tangency portfolio is 0.429, which means the investment has good sensitivity to market movements and to gauge risk. Meanwhile, our portfolio satisfies that Jensen's alpha is larger than zero. 0.301 alpha proves the active return on the investment is well-performed. Generally, in terms of the measures of risk-adjusted return (table 9), the performance of the optimal tangency portfolio we choose is great.

Optimal Portfolio Construction

In real world, people with different age, job, and personality have different risk tolerance, thus, have different investment strategy. For example, seniors may be considered as conservative investors; working professionals who just enter the working fields may be considered as aggressive investors and working professionals who are in their middle age may be considered as neutral investors. In order to find the best mix of risky and risk-free asset for different communities, we introduce risk aversion to calculate the weight (w^*) on risky asset and $1 - w^*$ for risk-free asset. We use Treasury bonds as our risk-free asset.

$$w^* = \frac{E(r_1) - r_f}{A \times \sigma_1^2}$$

where E(r_1) is the expected return of risky portfolio, r_f is risk-free rate, σ_1^2 is the variance of risky portfolio and A is risk aversion, which is the behavior of consumers and investors, when exposed to uncertainty, in attempting to lower that uncertainty.

We need to choose A because investors differ in the amount of risk they are willing to take for a given return, and investors who are risk averse require a greater return for a given amount of risk than risk lovers.

$$E(r_p) = r_f + w^* \times E[r_1 - r_f] \qquad \sigma_p = |w^*| \times \sigma_1$$

where E(r_p) is the expected return of the portfolio and σ_p is the volatility of the portfolio, Er_1 and σ_1 are the annual expected return and volatility of the risky portfolio (the portfolio with 50 stocks and obtained under mean-variance optimization methodology).

Because risk aversion is not an objectively measurable quantity, there is no unique equation that would yield such a quantity, but we can select, not for its absolute measure, but for its comparative measure of risk tolerance. According to rule of thumbs, we choose A to be a range of numbers, say, 5.00, 17.50, 25.46, 35.59 and 40.02. Based on the equations, we find that as the weight (w^*) on the risky asset decreases, the risk aversion increases, the expected return of portfolio decreases and the volatility of portfolio decreases(table 10). Since we assume no short sell and no leverage, we eliminate the case when w^* is greater than 1, which means we leave the cases when A=5.00 and 17.50 and only use A=25.46, 35.59 and 40.02 to construct our portfolio.

When A=25.46, we have w^* =0.7046435, $E(r_p)$=0.296 and σ_p=0.105, which indicates that investors could invest 70.4% into risky asset and 29.6% into Treasury bonds. We consider this to be a good suggestion for aggressive investors, such as working professionals who just enter the working fields. Similarly, when A=35.59, we have w^* = 0.5040805 , $E(r_p)$=0.2145 and σ_p=0.07579, which indicates that investors could invest 50.4% into risky asset and 49.6% into Treasury bonds. We consider this to be a good suggestion for neutral investors, ie working professionals who are in their middle age. When A= 40.02, we have w^* = 0.4485056 , $E(r_p)$= 0.19198and σ_p = 0.06743, which indicates that investors could invest 44.9% into risky asset and 55.1% into Treasury bonds. We consider this to be a good suggestion for conservative investors, such as seniors. Figure 4 to Figure 6 show the efficient frontier for the portfolio. The optimal portfolio for different communities are lined out and it changes as the value of A changes.

Risk Management

In market investing, increased potential returns on investment usually imply a positive relationship with increased risk. There are different types of risks such as industry-specific risk, international risk, and market risk. Since we want to find the optimal portfolios that fit different types of investors the best within the risk they can take, here we introduced the definition of value at the risk(VaR) to evaluate the portfolio. Value at risk measures the level of financial risk within a firm or investment portfolio over a specific time frame. VaR is the loss level that will not be exceeded with a specified probability. It is the main cause of uncertainty in any organization. We use VaR to test if our portfolio has a lower risk, in the meantime, also give a desirable return.

We use three methods to calculate VaR, empirical method, T-distribution method, and time series model. With empirical method, we put Columbia Tuition into each asset by historical data. We calculate four groups of data, which are value at risk according to historical data at 95% confidence interval, the best stock performer, mean value and our related asset portfolio. Similarly, we do the same task on t-distribution and time series model. By comparing three ways of evaluating VaR, t-distribution gives us a positive smaller value on VaR. We find that t-distribution model will fit our portfolio best and gets a better result in reducing risk. Additionally, to make sure that our result is correct, we also do the back test to confirm it. We choose 500 training data to test data under three models and find which value is close to 5%. The back test also gives us the desirable result (table 11) as we can see that true distribution sets between t-distribution and time series. T-distribution gives us a prudent prediction while time series analysis gives us a bold prediction.

By measuring VaR in another way to further understanding how our portfolio works well, I have compared the VaR of three best performed stocks portfolio, five random stocks portfolio and our diversified portfolio under 95% confidence interval. The output is shown in (table 12). Taking three stocks portfolio and our portfolio as an compared example. The table shows that there is 5% chance that three stocks portfolio will fall in value by more than 0.0155 when we invest $1, which in comparison to that there is 5% chance that our portfolio will fall in value by more than 0.0146. Thus, we recommend investors to choose our diversified portfolio as their optimal portfolio by giving the smallest VaR.

Conclusion

Our project starts with the goal to help different types of investors find the portfolios they truly want. To do so, we have first utilized mean variance analysis and risk parity to find the optimal weights of our tangency portfolio, and mixed the resulting portfolio with risk-free assets to get the best portfolio for three types of investor, conservative, aggressive, and neutral. Finally, the value at risk test confirms our goal that our portfolio gives investors a relative high return with less risk. Conducting this project, we have learned new asset allocation and risk management techniques. This project also makes us realize the importance of diversifying the portfolio and mixing in risk-free asset when we invest. It has truly motivated us to study more about portfolio construction and optimization in the future.

Acknowledgement

We want to express our sincere gratitude to Professor Ying and Teaching Assistants Brian and George. Thanks for their patience, help, and guidance on this project.

Reference

- Markowitz, H.M. (1959). *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons.
- Chaves, Hsu, Li,and Shakernia (2011). *Risk Parity Portfolio vs. Other Asset Allocation Heuristic Portfolios*
- Molenaar (2014) *"The best of two worlds"-Alternating Between Mean Variance and Risk Parity*

Table

(Table 1)

NOC	NVDA	LMT	DXC	RTN	STZ	CTAS	MSCI	HRS	HII
0.043607	1.34E-19	0.249401	0.052152	2.43E-17	0.230343	-8.75E-18	-3.27E-17	-1.23E-17	-1.38E-18
UNH	**GD**	**AVGO**	**FISV**	**BA**	**ADBE**	**SPGI**	**GPN**	**ALGN**	**EA**
7.71E-17	1.08E-16	0.072548	-2.70E-17	-6.86E-17	3.31E-17	-1.10E-16	0.069956	-4.84E-18	0.075391
CME	**ETFC**	**PGR**	**NFLX**	**NDAQ**	**AMZN**	**TTWO**	**FB**	**IDXX**	**TSS.UN**
-7.04E-17	7.43E-18	1.01E-16	0.008022	1.35E-17	-2.16E-17	-4.84E-19	0.056783	2.45E-18	-4.33E-17
WM.UN	**MA.UN**	**AOS**	**ITW**	**MAR**	**MMC**	**CBOE**	**HUM**	**SYK**	**ANTM**
2.13E-16	-9.59E-17	2.20E-17	-2.74E-17	2.00E-16	-1.74E-16	0.023832	0.058666	-1.75E-16	1.35E-17
TDG	**LRCX**	**RSG**	**MSFT**	**V**	**BSX**	**BDX**	**DPS**	**LUV**	**ATVI**
-4.23E-18	1.15E-16	1.46E-16	1.60E-16	3.85E-17	-4.97E-19	1.58E-16	1.15E-18	0.059297	-1.42E-17

(Table 2)

NOC	NVDA	LMT	DXC	RTN	STZ
0.065675	0.002545	0.275578	0.057231	-1.45E-17	0.252632
CTAS	**MSCI**	**HRS**	**HII**	**UNH**	**GD**
1.72E-17	-2.30E-17	4.70E-17	-2.26E-18	0	-6.85E-17

AVGO	FISV	BA	ADBE	SPGI	GPN
0.089535	-8.97E-17	-2.89E-17	-2.00E-17	-3.19E-17	0.089623
ALGN	**EA**	**CME**	**ETFC**	**PGR**	**NFLX**
-2.18E-18	0.0907	3.22E-17	-2.37E-18	4.05E-17	0.011962
NDAQ	**AMZN**	**TTWO**	**FB**	**IDXX**	**TSS.UN**
3.28E-18	5.63E-18	5.59E-18	0.064519	6.86E-19	1.32E-16

(Table 3)

NOC	NVDA	LMT	DXC	RTN
0.1438155	0.0865634	0.2968410	0.0683920	-2.31E-20
STZ	**CTAS**	**MSCI**	**HRS**	**HII**
0.3328861	-2.56E-18	-4.57E-19	-2.99E-18	0.0715019

(Table 4)

NOC	NVDA	LMT	DXC	RTN	STZ	CTAS	MSCI	HRS	HII
0.019974	0.016833	0.024774	0.018495	0.022623	0.022985	0.022367	0.019521	0.020078	0.018093
UNH	**GD**	**AVGO**	**FISV**	**BA**	**ADBE**	**SPGI**	**GPN**	**ALGN**	**EA**
0.019844	0.020256	0.013701	0.019076	0.018903	0.016744	0.01742	0.017965	0.016011	0.01655
CME	**ETFC**	**PGR**	**NFLX**	**NDAQ**	**AMZN**	**TTWO**	**FB**	**IDXX**	**TSS.UN**
0.022703	0.012626	0.025824	0.014214	0.020915	0.015895	0.017224	0.014603	0.022777	0.017915
WM.UN	**MA.UN**	**AOS**	**ITW**	**MAR**	**MMC**	**CBOE**	**HUM**	**SYK**	**ANTM**
0.029764	0.016383	0.017322	0.021713	0.018503	0.022056	0.028898	0.022814	0.023392	0.021264
TDG	**LRCX**	**RSG**	**MSFT**	**V**	**BSX**	**BDX**	**DPS**	**LUV**	**ATVI**
0.024149	0.015627	0.02947	0.018696	0.017906	0.017732	0.024423	0.029261	0.016627	0.017094

(Table 5)

NOC	NVDA	LMT	DXC	RTN	STZ
0.035321	0.0295	0.044167	0.030292	0.039983	0.041046

CTAS	**MSCI**	**HRS**	**HII**	**UNH**	**GD**
0.040011	0.034101	0.035256	0.03166	0.037155	0.036192
AVGO	**FISV**	**BA**	**ADBE**	**SPGI**	**GPN**
0.023839	0.034336	0.033462	0.029627	0.03105	0.031912
ALGN	**EA**	**CME**	**ETFC**	**PGR**	**NFLX**
0.027807	0.028871	0.03982	0.022088	0.047064	0.023708
NDAQ	**AMZN**	**TTWO**	**FB**	**IDXX**	**TSS.UN**
0.037011	0.027983	0.029617	0.024979	0.040253	0.031889

(Table 6)

NOC	**NVDA**	**LMT**	**DXC**	**RTN**
0.097721	0.084074	0.119568	0.070196	0.105798
STZ	**CTAS**	**MSCI**	**HRS**	**HII**
0.115454	0.117245	0.101141	0.100727	0.088076

(Table 7)

Mean-Variance	**Expected Return**	**Volatility**
10 stocks	0.426	0.156
30 stocks	0.424	0.155
50 stocks	0.415	0.150
Risk Parity	**Expected Return**	**Volatility**
10 stocks	0.343	0.146
30 stocks	0.328	0.148
50 stocks	0.296	0.139

(Table 8)

	10 stocks	**30 stocks**	**50 stocks**
Mean-Variance	2.491912	2.669507	2.697261
Risk Parity	2.281646	2.145401	2.058638

(Table 9)

Risk Ratio	**Value**
Sharpe Ratio	2.697261

Treynor Ratio	0.429248
Jensen's Alpha	0.301409

(Table 10)

A	w^*	$E(r_p)$	σ_p
5.00	3.588045	1.465143	0.53945
17.50	1.025156	0.42583	0.15413
25.46	0.7046435	0.296	0.105
35.59	0.5040805	0.2145	0.07579
40.02	0.4485056	0.19198	0.06743

(Table 11)

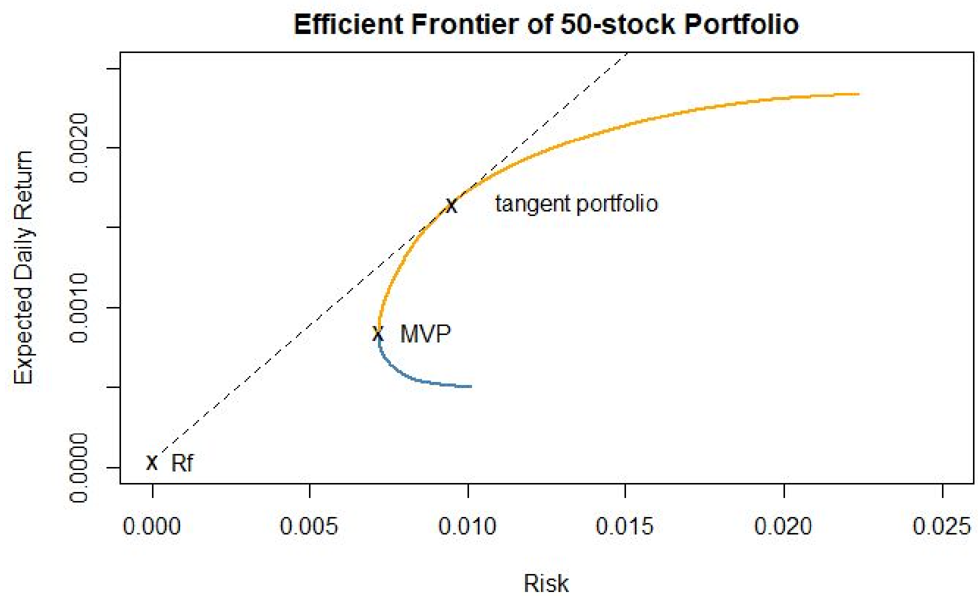
	empirical	**t-distribution**	**time series**
backtest value	0.02994	0.03792	0.08583

(Table 12)

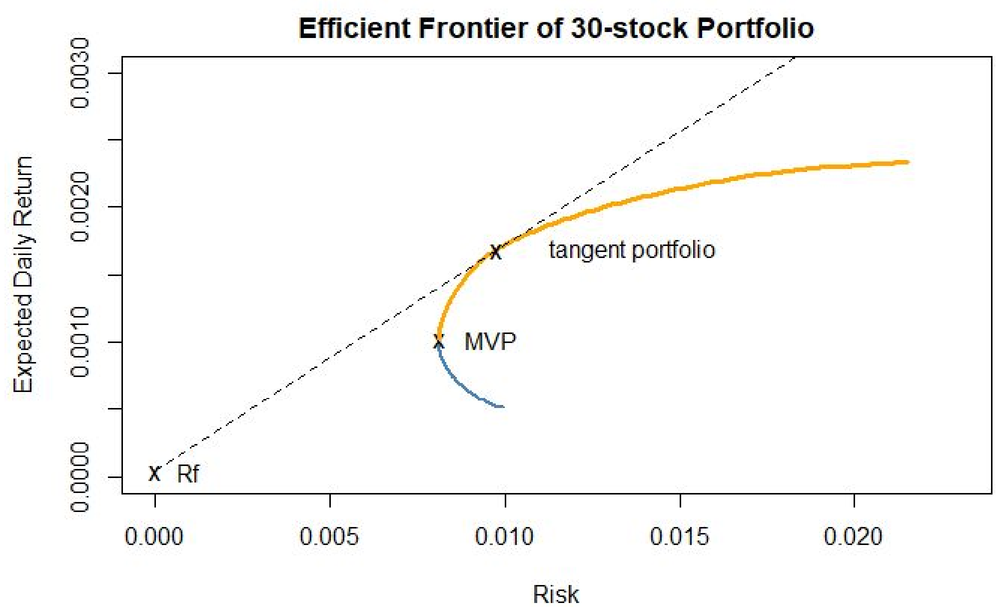
	3 stocks portfolio	**5 stocks portfolio**	**diversified portfolio**
VaR at 95%	-0.01548	-0.01614	-0.01461

Figures

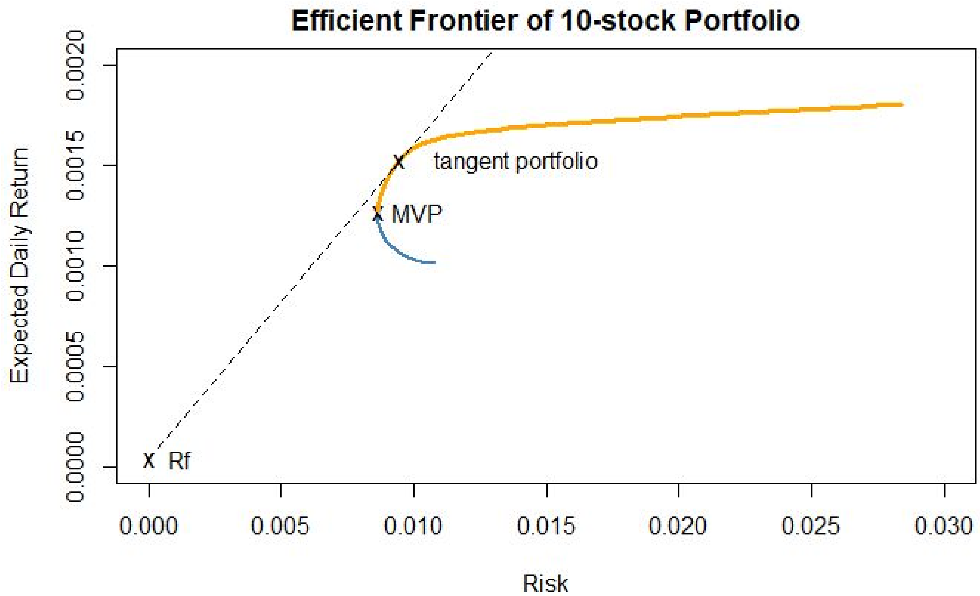
(Figure 1)



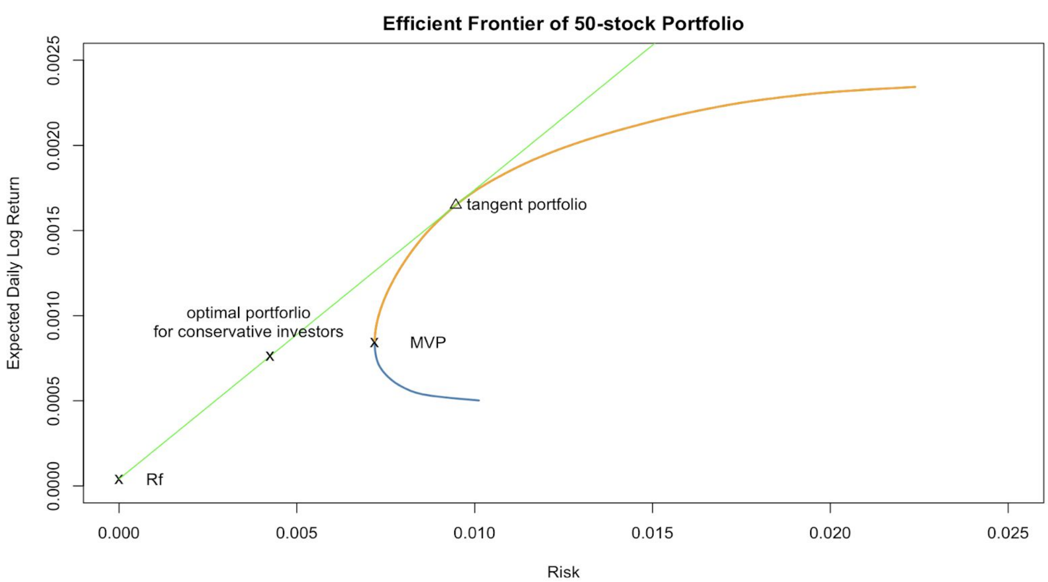
(Figure 2)



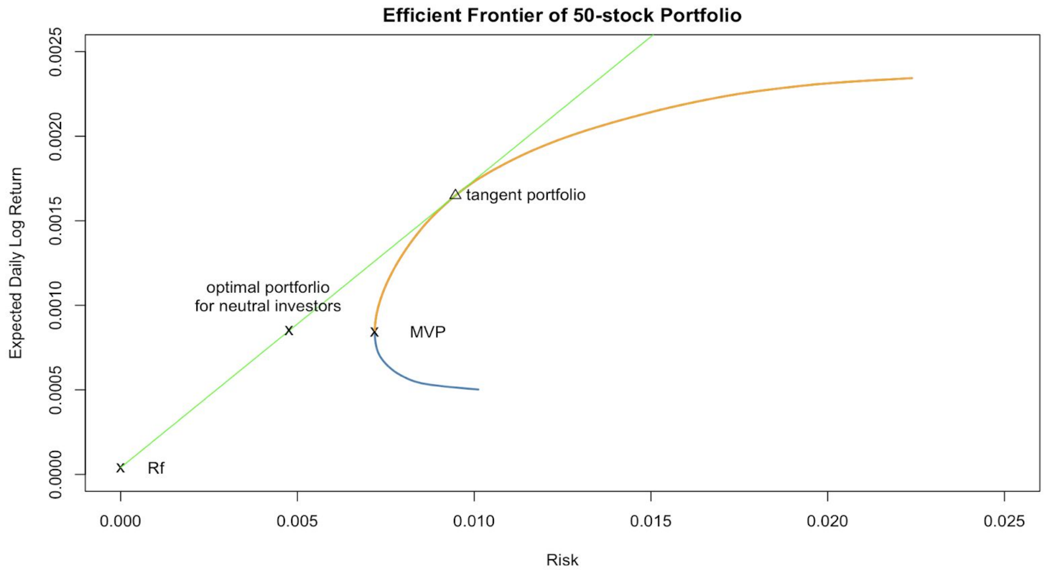
(Figure 3)



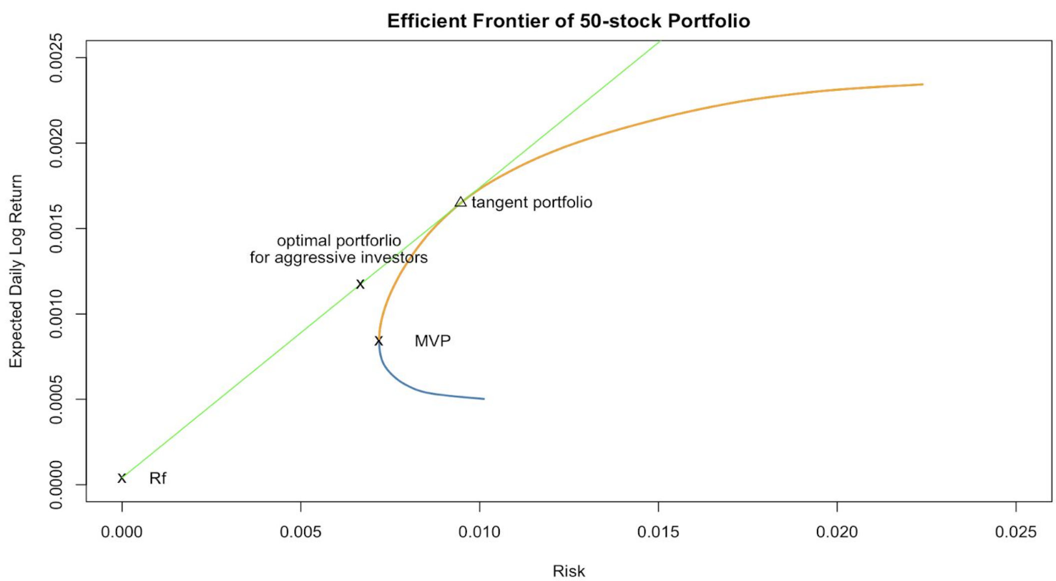
(Figure 4)



(Figure 5)



(Figure 6)



Attachment

(1) Mean-variance Analysis

```
#use 20130423 to 20160422 daily price
rty<-read.csv("spx50_20130423to20160422.csv",header = T)

ret<-log(rty[-1,]/rty[-nrow(rty),])#daily log return
ret<-exp(ret)-1#convert into daily net return

#Asset Allocation
library(MASS)
library(quadprog)

rf<-0.0101/252

#50 stocks' portfolio

ret50<-ret

mu50 = colMeans(ret50)
sigma50 = cov(ret50)


muP50 = seq(min(mu50)+0.000001,max(mu50),length=200)  # target portfolio return
sdP50 = muP50 # sd of portfolio return

weight50 = matrix(0,nrow=200,ncol=50) # storage for portfolio weights
for (i in 1:length(muP50))  # find the optimal portfolios
{
  result = solve.QP(Dmat=2*sigma50,dvec=rep(0,50),
                   Amat=cbind(rep(1,50),mu50,diag(1,50)),
                   bvec=c(1,muP50[i],rep(0,50)),meq=2)
  sdP50[i] = sqrt(result$value)
  weight50[i,] = result$solution
}

par(mfrow = c(1,1))# draw efficient frontier
plot(sdP50,muP50,type="l",xlim=c(0,0.025),ylim=c(0,0.0025),lwd=2,col="steelblue",
     main = "Efficient Frontier of 50-stock Portfolio",
     ylab = "Expected Daily Return",
     xlab = "Risk")

ind1_50 = (sdP50 == min(sdP50))
points(sdP50[ind1_50],muP50[ind1_50],cex=1.1,pch="x")#MVP
text(sdP50[ind1_50]+0.0015,muP50[ind1_50],"MVP",cex=1)

points(0,rf,cex=1.1,pch="x") # show riskfree asset
text(0.001,rf,"Rf",cex=1)

ind2_50 = (muP50 > muP50[ind1_50])
lines(sdP50[ind2_50],muP50[ind2_50],type="l",xlim=c(0,.05),
      ylim=c(0,.0015),lwd=2,col="orange")#efficient frontier

sharpe50 =(muP50-rf)/sdP50 # Sharpe ratio
ind3_50 = (sharpe50 == max(sharpe50)) # find maximum Sharpe ratio
lines(c(0,1),rf+c(0,1)*sharpe50[ind3_50],lwd=1,lty=2,col="black") # line of optimal portfolios
```

```
points(sdP50[ind3_50],muP50[ind3_50],cex=1.1,pch="x")#tangent portfolio
text(sdP50[ind3_50]+0.004,muP50[ind3_50],"tangent portfolio",cex=1)

w50<-weight50[which.max(sharpe50),]#weight of 50 stocks
names(w50)<-colnames(ret50)

#30 stocks' portfolio

ret30<-ret[,1:30]

mu30 = colMeans(ret30)
sigma30 = cov(ret30)


muP30 = seq(min(mu30)+0.00001,max(mu30)-0.00001,length=200)
sdP30 = muP30

weight30 = matrix(0,nrow=200,ncol=30)
for (i in 1:length(muP30))
{
  result = solve.QP(Dmat=2*sigma30,dvec=rep(0,30),
                    Amat=cbind(rep(1,30),mu30,diag(1,30)),
                    bvec=c(1,muP30[i],rep(0,30)),meq=2)
  sdP30[i] = sqrt(result$value)
  weight30[i,] = result$solution
}
par(mfrow = c(1,1))
plot(sdP30,muP30,type="l",xlim=c(0,0.023),ylim=c(0,0.003),lwd=2,col="steelblue",
     main = "Efficient Frontier of 30-stock Portfolio",
     ylab = "Expected Daily Return",
     xlab = "Risk")
ind1_30 = (sdP30 == min(sdP30))

points(sdP30[ind1_30],muP30[ind1_30],cex=1.1,pch="x")
text(sdP30[ind1_30]+0.0015,muP30[ind1_30],"MVP",cex=1)

ind2_30 = (muP30 > muP30[ind1_30])
lines(sdP30[ind2_30],muP30[ind2_30],type="l",xlim=c(0,.05),
      ylim=c(0,.0015),lwd=3,col="orange")

points(0,rf,cex=1.1,pch="x")
text(0.001,rf,"Rf",cex=1)

sharpe30 =(muP30-rf)/sdP30
ind3_30 = (sharpe30 == max(sharpe30))
lines(c(0,1),rf+c(0,1)*sharpe30[ind3_30],lwd=1,lty=2,col="black")
points(sdP30[ind3_30],muP30[ind3_30],cex=1.1,pch="x")
text(sdP50[ind3_30]+0.004,muP50[ind3_30],"tangent portfolio",cex=1)

w30<-weight30[which.max(sharpe30),]#weight of 30 stocks
names(w30)<-colnames(ret30)

#10 stocks' portfolio

ret10<-ret[,1:10]

mu10 = colMeans(ret10)
sigma10 = cov(ret10)

muP10 = seq(min(mu10)+0.00001,max(mu10)-0.00001,length=200)
sdP10 = muP10
```

```
weight10 = matrix(0,nrow=200,ncol=10)
for (i in 1:length(muP10))
{
  result = solve.QP(Dmat=2*sigma10,dvec=rep(0,10),
                    Amat=cbind(rep(1,10),mu10,diag(1,10)),
                    bvec=c(1,muP10[i],rep(0,10)),meq=2)
  sdP10[i] = sqrt(result$value)
  weight10[i,] = result$solution
}
par(mfrow = c(1,1))
plot(sdP10,muP10,type="l",xlim=c(0,0.03),ylim=c(0,0.002),lwd=2,col="steelblue",
     main = "Efficient Frontier of 10-stock Portfolio",
     ylab = "Expected Daily Return",
     xlab = "Risk")
ind1_10 = (sdP10 == min(sdP10))
points(sdP10[ind1_10],muP10[ind1_10],cex=1.1,pch="x")
text(sdP10[ind1_10]+0.0015,muP10[ind1_10],"MVP",cex=1)
text(0.0012,rf,"Rf",cex=1)

ind2_10 = (muP10 > muP10[ind1_10])
lines(sdP10[ind2_10],muP10[ind2_10],type="l",xlim=c(0,.05),
      ylim=c(0,.0015),lwd=3,col="orange")

points(0,rf,cex=1.1,pch="x")
sharpe10 =(muP10-rf)/sdP10
ind3_10 = (sharpe10 == max(sharpe10))
lines(c(0,1),rf+c(0,1)*sharpe10[ind3_10],lwd=1,lty=2,col="black")
points(sdP10[ind3_10],muP10[ind3_10],cex=1.1,pch="x")
text(sdP10[ind3_10]+0.0045,muP10[ind3_10],"tangent portfolio",cex=1)

w10<-weight10[which.max(sharpe10),]
names(w10)<-colnames(ret10)

#summary

expect.rtn10<-sum(w10*mu10)*252
expect.rtn30<-sum(w30*mu30)*252
expect.rtn50<-sum(w50*mu50)*252

risk10<-sdP10[which.max(sharpe10)]*sqrt(252)
risk30<-sdP30[which.max(sharpe30)]*sqrt(252)
risk50<-sdP50[which.max(sharpe50)]*sqrt(252)

summary.port<-data.frame(expected.annual.rtn=c(expect.rtn10,expect.rtn30,expect.rtn50),
                         Volatility=c(risk10,risk30,risk50))
rownames(summary.port)<-c("10-stock","30-stock","50-stock")

summary.port

#compute sharpe ratio

rf<-0.0101
sharpe.ratio10<-(expect.rtn10-rf)/risk10
sharpe.ratio30<-(expect.rtn30-rf)/risk30
sharpe.ratio50<-(expect.rtn50-rf)/risk50
c(sharpe.ratio10,sharpe.ratio30,sharpe.ratio50)
```

(2) Risk Parity

```r
z<-read.csv("spx50_20130423to20160422.csv",header = T)

library(PortfolioAnalytics)

z.logrtn <- apply(log(z),2,diff)#50 stocks
z.rtn <- exp(z.logrtn)-1#turn logreturn into net return
std <- apply(z.rtn,2,StdDev)
cov.mat <- cov(z.rtn)
mean.rtn<-apply(z.rtn,2,mean)#daily mean return

z.rtn30<-z.rtn[,1:30]#30 stocks
std30<-std[1:30]
cov.mat30 <- cov(z.rtn30)
mean.rtn30<-apply(z.rtn30,2,mean)

z.rtn10<-z.rtn[,1:10]#10 stocks
std10<-std[1:10]
cov.mat10 <- cov(z.rtn10)
mean.rtn10<-apply(z.rtn10,2,mean)

x0 <- (1/std)/sum(1/std) #starting weights
x0_30<-(1/std30)/sum(1/std30)
x0_10<-(1/std10)/sum(1/std10)

# objective function
eval_f <- function(w,cov.mat,vol.target) {
  vol <- sqrt(as.numeric(t(w) %*% cov.mat %*% w))#sigma(w)
  marginal.contribution <- cov.mat %*% w / vol#wi
  return( sum((vol/length(w) - w * marginal.contribution)^2) )
}

# numerical gradient approximation for solver
eval_grad_f <- function(w,cov.mat,vol.target) {
  out <- w
  for (i in 0:length(w)) {
    up <- dn <- w
    up[i] <- up[i]+.0001
    dn[i] <- dn[i]-.0001
    out[i] = (eval_f(up,cov.mat=cov.mat,vol.target=vol.target)
          - eval_f(dn,cov.mat=cov.mat,vol.target=vol.target))/.0002
  }
  return(out)
}

#do optimization
library(nloptr)

#50 stocks
res50 <- nloptr(x0=x0,
            eval_f=eval_f,
            eval_grad_f=eval_grad_f,
            eval_g_eq=function(w,cov.mat,vol.target) { sum(w) - 1 },
            eval_jac_g_eq=function(w,cov.mat,vol.target) { rep(1,length(std)) },
            lb=rep(0,length(std)),ub=rep(1,length(std)),
            opts = list("algorithm"="NLOPT_LD_SLSQP","print_level" = 3,
                        "xtol_rel"=1.0e-8,"maxeval" = 1000),
            cov.mat = cov.mat,vol.target=.02 )

weight50<-res50$solution#weights
expect.rtn50<-sum(weight50*mean.rtn)#daily return
expect.rtn50<-expect.rtn50*252# turn daily return into annual return
```

```
risk50<-sqrt((res50$solution %*% cov.mat %*% res50$solution)*252)#annual

#30 stocks
res30<-nloptr(x0=x0_30,
            eval_f=eval_f,
            eval_grad_f=eval_grad_f,
            eval_g_eq=function(w,cov.mat,vol.target) { sum(w) - 1 },
            eval_jac_g_eq=function(w,cov.mat,vol.target) { rep(1,length(std30)) },
            lb=rep(0,length(std30)),ub=rep(1,length(std30)),
            opts = list("algorithm"="NLOPT_LD_SLSQP","print_level" = 3,
                        "xtol_rel"=1.0e-8,"maxeval" = 1000),
            cov.mat = cov.mat30,vol.target=.02 )
weight30<-res30$solution#weights
expect.rtn30<-sum(weight30*mean.rtn30)
expect.rtn30<-expect.rtn30*252
risk30<-sqrt((res30$solution %*% cov.mat30 %*% res30$solution)*252)

res10<-nloptr(x0=x0_10,
            eval_f=eval_f,
            eval_grad_f=eval_grad_f,
            eval_g_eq=function(w,cov.mat,vol.target) { sum(w) - 1 },
            eval_jac_g_eq=function(w,cov.mat,vol.target) { rep(1,length(std10)) },
            lb=rep(0,length(std10)),ub=rep(1,length(std10)),
            opts = list("algorithm"="NLOPT_LD_SLSQP","print_level" = 3,
                       "xtol_rel"=1.0e-8,"maxeval" = 1000),
            cov.mat = cov.mat10,vol.target=.02 )
weight10<-res10$solution#weights
expect.rtn10<-sum(weight10*mean.rtn10)
expect.rtn10<-expect.rtn10*252
risk10<-sqrt((res10$solution %*% cov.mat10 %*% res10$solution)*252)

summary.port<-data.frame(expected.annual.rtn=c(expect.rtn10,expect.rtn30,expect.rtn50),
                      Volatility=c(risk10,risk30,risk50)) ##volatility is portfolio std

rownames(summary.port)<-c("10-stock","30-stock","50-stock")

summary.port

#10-stock portfolio has max expected logreturn; 50-stock portfolio has min risk

rf<-0.0101
# choose the "best" portfolio with biggest Sharpe Ratio

sharpe.ratio10<-(expect.rtn10-rf)/risk10
sharpe.ratio30<-(expect.rtn30-rf)/risk30
sharpe.ratio50<-(expect.rtn50-rf)/risk50
c(sharpe.ratio10,sharpe.ratio30,sharpe.ratio50)


##calculate the weight (w*) on risky asset and 1-w* for risk-free asset
E_risky<-0.4156253 #risky portfolio
rf<-0.0101
A<-c(5.00, 17.50, 25.46, 35.59, 40.02)
sigma<-0.1503471
w<-rep(0,5)
for (i in A){
w<-(E_risky-rf)/(i*sigma^2)
}
```

(3) Criteria

```
### 1. Sharpe Ratio
sharpe.rp<-sharpe.ratio50
sharpe.rp

### 2. Treynor Ratio

mrk<-read.table("F-F_Research_Data_Factors_daily.txt",header = T)
mrk<-mrk[22959:23714,]#choose the corresponding period data
mrk$Rf<-as.numeric(levels(mrk$RF))[mrk$RF]
mrk$RM<-(mrk$Mkt.RF+mrk$Rf)*0.01#market daily return
mrk.var<-var(mrk$RM,na.rm = T)
std.M_with_all<-apply(ret50,2,cov,mrk$RM)

beta.i<-std.M_with_all/mrk.var
beta.p<-sum(w50*beta.i)

treynor.ratio<-(expect.rtn50-rf)/beta.p
treynor.ratio

### 3.Jensen's alpha
RM<-mean(mrk$RM)*252
jensen.alpha<-expect.rtn50-rf-beta.p*(RM-rf)
jensen.alpha


### 4. Fama French Model's alpha

FF<-read.table("F-F_Research_Data_Factors_daily.txt",header = T)
FF<-FF[which(rownames(FF)=="20130424"):which(rownames(FF)=="20160422"),]

port50.rtn<-apply(ret50,1,function(x){return(sum(w50*x))})# Rp
Rf<-as.numeric(levels(FF$RF))[FF$RF]
excess.rtn<-port50.rtn-Rf#Rp-Rf

fit1 <- lm(excess.rtn~FF$Mkt.RF+FF$SMB+FF$HML)
fit1$coefficients[1]
```

(4) Value at risk

```
### 1. Empirical Distribution
library(ggplot2)
w50=as.matrix(w50)
ret50.mat=as.matrix(ret50)
portfolio=ret50.mat%*%w50  # the same as port50.rtn


average=apply(ret50,1,sum)/ncol(ret50)
alpha=0.05
q_nonp = as.numeric(quantile(portfolio, alpha))
q_nonp #value at risk according to historical data
q_nonp_best_performer=as.numeric(quantile(ret50[,1], alpha))
q_nonp_equal_weight=as.numeric(quantile(average, alpha))
q_nonp_mediocre_performer=as.numeric(quantile(ret50[,floor(ncol(ret50)/2)], alpha))
y=-67466*cbind(q_nonp,q_nonp_equal_weight,q_nonp_best_performer,q_nonp_mediocre_performer)
ggplot(data= NULL, aes(x =colnames(y) , y = as.numeric(y)))+
 geom_bar(stat="identity",width=0.5,position=position_dodge(0.6))+

 labs(x = "Portfolio by type", y = "Dollars at risk", title = "Value at risk if you put your Columbia Tuition into each asset ")

### 2. T Distribution model
```

```
fit_t=fitdistr(portfolio,'t')
mean = as.numeric(fit_t$estimate)[1]
df = as.numeric(fit_t$estimate)[3]
sd = as.numeric(fit_t$estimate)[2] * sqrt((df) / (df - 2))
lambda = as.numeric(fit_t$estimate)[2]
q_t_alpha = qt(alpha, df = df)
#q_t_alpha
VaR_t_port = (mean + lambda * q_t_alpha)


###3.  VaR according to time series model
portseries=ts(portfolio)
plot(portseries)
acf(portfolio)
#acf(portfolio,main='',xlab='Lag (a)',ylab='ACF',las=1) #
pacf(portfolio)

averseries=ts(average)
plot(averseries)
acf(averseries)
#acf(portfolio,main='',xlab='Lag (a)',ylab='ACF',las=1) #
pacf(averseries)


library(forecast)
arima_model=auto.arima(portfolio,ic="bic")
arima_model
#it is supposed that does not follow Arima model

library(rugarch)
library(fGarch)
garch.t = ugarchspec(mean.model=list(armaOrder=c(0,0)),#variance.model=list(garchOrder=c(1,2)),
distribution.model="std")
sp.garch.t = ugarchfit(data=portfolio, spec=garch.t)
show(sp.garch.t)
alpha = 0.05
nu = as.numeric(coef(sp.garch.t)[5])

pred = ugarchforecast(sp.garch.t, data=portfolio, n.ahead=1)
time_seris_VAR = as.numeric(qstd(alpha, mean=fitted(pred), sd=sigma(pred), nu=nu))

z=-67466*cbind(q_nonp,VaR_t_port,time_seris_VAR)
ggplot(data= NULL, aes(x =colnames(z) , y = as.numeric(z)))+
 geom_bar(stat="identity",width=0.5,position=position_dodge(0.6))+
 labs(x = "Method by type", y = "Dollars at risk", title = "Value at risk if you put your Columbia Tuition into each asset ")

### 4. Backtest

r_test<-read.csv("spx_50.csv",header = F,as.is=T)
rmm=which(r_test=="4/22/2016") #rmm=505
r_test=r_test[1:505-1,]
test=r_test[-2,-1]
colnames(test)=test[1,]
rty_test=as.data.frame(sapply(test[-1,], as.numeric))

rettest<-log(rty_test[-1,]/rty_test[-nrow(rty_test),])#daily log return
rettest<-exp(rettest)-1#convert into net return


w50=as.matrix(w50)
ret50_test.mat=as.matrix(rettest)
```

```
portfolio_test=ret50_test.mat%*%w50

backtest_nonp=sum(portfolio_test<q_nonp)/length(portfolio_test)
backtest_nonp

backtest_t=sum(portfolio_test<VaR_t_port)/length(portfolio_test)
backtest_t

backtest_ts=sum(portfolio_test<time_seris_VAR)/length(portfolio_test)
backtest_ts
```