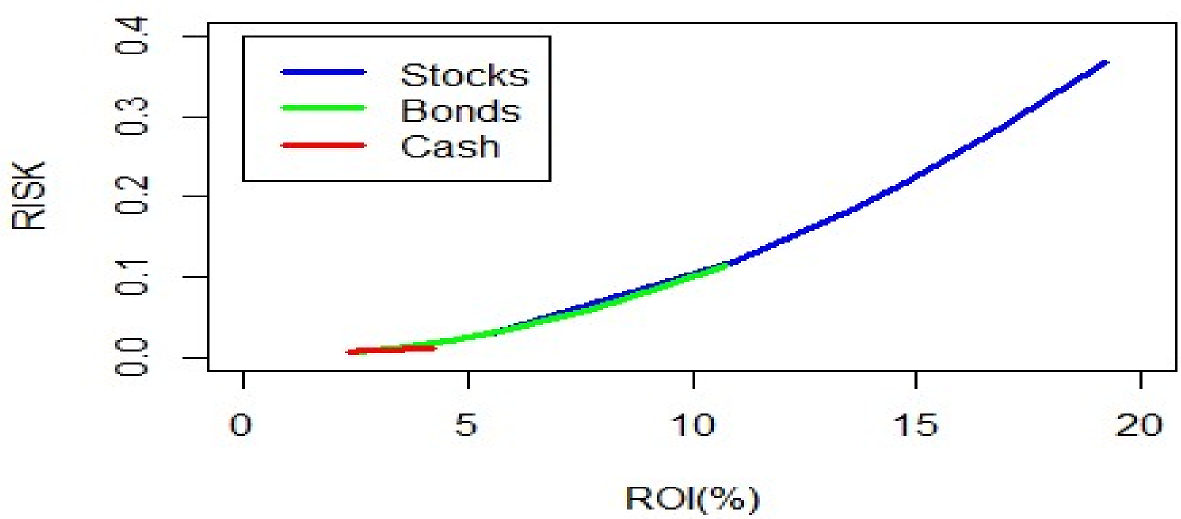
INVESTMENT PORTFOLIO MANAGEMENT

1. Visualise and discuss the different ROI and Risk associated with Stocks, Bonds and Cash.

The R code for the visualisation is as follows:

```
stock<-subset(invest, invest$Type==1 )

bond<-subset(invest, invest$Type==2 )

cash<-subset(invest, invest$Type==3 )

stock<-stock[order(stock$ROI),]

bond<-bond[order(bond$ROI),]

cash<-cash[order(cash$ROI),]

plot(stock$ROI,stock$Risk,col="blue",xlim=c(0,20),ylim=c(0,0.4),xlab="ROI(%)",ylab="RISK",lwd=3,ty
pe="l")

lines(bond$ROI,bond$Risk,col="green",lwd=3)

lines(cash$ROI,cash$Risk,col="red",lwd=3)

legend(0,0.4, legend = c("Stocks", "Bonds","Cash"),col=c("blue",
"green","red"),lty=c(1,1,1),lwd=c(3,3,3))
```

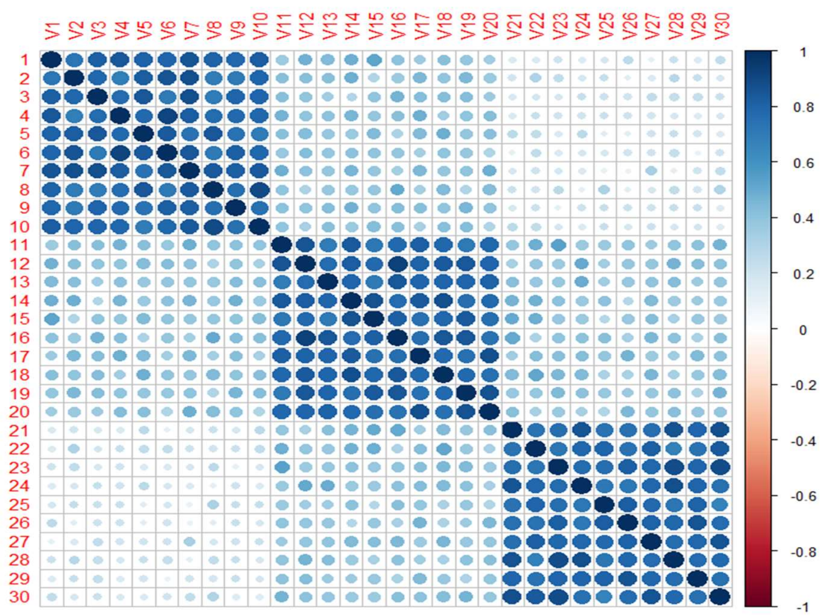


From the screenshot above, we can see that different investment products will get different benefits and take different risks. Stocks offer the highest investment return rate and the highest risk. Bonds offer a medium investment return rate and a medium risk, while cash has the lowest investment return rate and the lowest risk. Specifically, it can be seen from the above plot that the ROI of stocks is between 5% and 20%, the ROI of bonds is between 3% and 11%, and the ROI of cash is between 2% and 4%. From the perspective of risk, cash has almost no risk, the highest risk of bonds is about 0.1, and the highest risk of stocks reaches 0.35.

In addition, load, visualise and give a simple interpretation for the correlation table.

The R code for this task is as follows:

```
corr <- read.table("corr.tab")

corr<-as.matrix(corr)

library(corrplot)

corrplot(corr)
```



From the above screenshot, we can see that there's a strong correlation between stocks, there's a strong correlation between bonds, there's a strong correlation between cash, there's a moderate correlation between stocks and bonds, there's a moderate correlation between bonds and cash, there's a weak correlation between stocks and cash.

2, Run the "invest.R" script. This script does a multi-objective criteria analysis to determine the best mix of stocks, bonds and cash over a range of tradeoffs. Describe in words what the "invest.R" script is doing. Ensure that you address how a solution is represented and interpreted, and the relationship between the solution space, the objective space and the constraints.

From the "invest.R" script, we can find the following:

1, There are 3 investment products involved, they are stocks, bonds, and cash. The relevant investment data need to be analysed based on the ROI and risk. The ROI and risk for each of the investment products are strongly correlated.

2, The “invest.R” script is used to build a multi-objective constrained model for an investment portfolio based on the stocks data, bonds data, and cash data.

3, From the “invest.R” script and the question, we find that it is a multi-objective problem. We need to maximize ROI and minimize Risk in such a problem. So we can use the Pareto front solution to trade off the ROI and the risk for each of the investment products.

4, In the solution from the “invest.R” script, the number of different investment products is the number of rows in the table. The minimum amount we can invest for each product is 5%, and the maximum amount we can invest for each product is 20%.

5, The constraints will be penalized when doing nsga2 if it is negative, but there will be no penalization when the constraint’s result is 0 or positive. From the “invest.R” script, we can see the following constraints need to be considered:

- The total percentage of selected investments (over 5%) should be between 95% and 100%.
- The investment amount should be between 5% and 20%.
- The total number of investments should be between 8 and 12.

```
> portfolio$par[1,]
 [1] 0.1984536298 0.0905666727 0.0310658259 0.0105674983 0.0378783180 0.0632429781 0.1949534812 0.0482629278 0.1009133106
[10] 0.1161412928 0.0167702136 0.0176228231 0.0545115230 0.0083351129 0.1249962379 0.0006724907 0.0449707856 0.0495240453
[19] 0.0538291245 0.0498099208 0.0379567462 0.0336067958 0.0147091280 0.0375574082 0.0360480496 0.0004211961 0.0234502988
[28] 0.0441770722 0.0410984461 0.0420149096
```

6, The screenshot above represents one of the x.

Return on Investment (ROI) measures the gain or loss of an investment, as a percentage, relative to its cost. From the “invest.R” script, we can find that the ROI is equal to the investments in x when it is large than the mini amount of 0.05, which is the proportion of the investments that we selected from the table, which is used to be multiplied by the corresponding ROI. Because the ROI needs to be maximised, so the ROI needs to take the negative of it as the nsga2 we are using always try to minimise the objective.

7, From the “invest.R” script, we can find that the risk of the portfolio, is equal to the investments in x when it is large than the mini amount of 0.05, which is the proportion of the investments that we selected from the table, which is used to be multiplied by the corresponding risk. Because the risk needs to be minimized, so the risk needs to take the positive of it as the nsga2 we are using always tries to minimise the objective.

8, Nsga2 is used as the none dominated GA to do the actual Pareto front. From the nsga2, we can find that there are some parameters, we set the dimension as the number of investments, there are 2 objectives separately are ROI and risk, and there are 3 constraints mentioned. We can learn the following information from the nsga2, The Pareto front solution is held in Portfolio$par. We have 52 population in nsga2 and 30 investments, so the Portfolio$par will have 52 rows to represent each individual in the population, and 30 columns to represent each investment.


```
> portfolio$par[1,]
 [1] 0.1984536298 0.0905666727 0.0310658259 0.0105674983 0.0378783180 0.0632429781 0.1949534812 0.0482629278 0.1009133106
[10] 0.1161412928 0.0167702136 0.0176228231 0.0545115230 0.0083351129 0.1249962379 0.0006724907 0.0449707856 0.0495240453
[19] 0.0538291245 0.0498099208 0.0379567462 0.0336067958 0.0147091280 0.0375574082 0.0360480496 0.0004211961 0.0234502988
[28] 0.0441770722 0.0410984461 0.0420149096
```

The screenshot above is the first row in Portfolio$par, that is an individual in the population that propose a solution, and that is one of the x in the ROI<-function(x). The above screenshot solution can be interpreted as any bit that is over 0.05 in the vector, which will be selected, and the number of the bit that is over 0.05 is the proportion of the investment in a portfolio. The Portfolio$value stores all the objective values. Portfolio$value is the interpretation of each Pareto front solution in the population Portfolio$par.

9, From the nsga2 plot, we can find that the risk increases when the ROI increases. There is a positive correlation between them.

3, Using the result of the nsga2 model you have previously run, examine and present the blend of stocks, bonds, and cash for a low risk, moderate risk, and high-risk investment blend (just pick one from each general category).

First, we need to define the range for low-risk portfolio, moderate risk portfolio, and high-risk portfolio separately.

Low-risk portfolio <= 0.05

Moderate risk portfolio >0.05 & <= 0.13

High risk portfolio > 0.13

The R code for the low-risk portfolio:

lowriskrow <- which(portfolio$value[,2]==min(portfolio$value[,2]))

panswerl <- portfolio$par[lowriskrow[1],]

lowriskp<-invest[which(panswerl >= 0.05),]

```
> lowriskp
             ROI        Risk Type
Bond 4  2.526935 0.006385402    2
Bond 8  3.635072 0.013213749    2
Cash 2  2.841369 0.008118196    3
Cash 4  3.823933 0.010925524    3
Cash 6  3.071053 0.008774437    3
Cash 7  3.013850 0.008610999    3
Cash 9  2.351452 0.006718435    3
Cash 10 4.183141 0.011951832    3
```

The R code for the moderate risk portfolio:

subset1<-subset(portfolio$value[,2], portfolio$value[,2] > 0.05)

subset2<-subset(subset1, subset1<=0.13)

mriskrow <- which(portfolio$value[,2]>median(portfolio$value[,2])-0.01&portfolio$value[,2]<median(portfolio$value[,2])+0.01)

panswerm <- portfolio$par[mriskrow[1],]

mriskp<-invest[which(panswerm >= 0.05),]

```
> mriskp
               ROI       Risk Type
Stock 5 13.843274 0.19163622    1
Stock 8 13.761788 0.18938680    1
Stock 9 11.600344 0.13456798    1
Bond 2   7.426795 0.05515729    2
Bond 5   8.310481 0.06906410    2
Bond 6   6.377948 0.04067823    2
Bond 7  10.724566 0.11501632    2
Bond 9   7.630622 0.05822639    2
Cash 4   3.823933 0.01092552    3
```

The R code for the high risk portfolio:

highriskrow <- which(portfolio$value[,2]==max(portfolio$value[,2]))

panswerh <- portfolio$par[highriskrow[1],]

highriskp<-invest[which(panswerh >= 0.05),]

```
> highriskp
               ROI       Risk Type
Stock 1 14.719142 0.21665314    1
Stock 2 16.795006 0.28207223    1
Stock 5 13.843274 0.19163622    1
Stock 6 19.187734 0.36816913    1
Stock 8 13.761788 0.18938680    1
Stock 9 11.600344 0.13456798    1
Bond 2   7.426795 0.05515729    2
Bond 5   8.310481 0.06906410    2
Bond 7  10.724566 0.11501632    2
Bond 9   7.630622 0.05822639    2
```

The R code for the new data frame with stock, bond, and cash percentage:

pv<-as.data.frame(portfolio$value)

pv[,1]<-pv[,1]*-1

colnames(pv)<-c("ROI(%)","RISK")

newpar<-portfolio$par

newpar[which(newpar<=0.05)]<-0

pv$stockperc<-apply(newpar[,1:10],1,sum)*100

pv$bondperc<-apply(newpar[,11:20],1,sum)*100

pv$cashperc<-apply(newpar[,21:30],1,sum)*100

```
> pv$stockperc
 [1] 63.146125  0.000000 34.744085 33.986488 17.928512 56.034532  8.711677 47.517539 51.066223
[10]  8.061075  6.426092  5.992170  8.671502 52.967717 32.040112 48.782653 21.847746  9.150057
[19]  7.864735  6.008228 56.494427  0.000000 45.226996 44.830287 20.928400 29.795044 34.040986
[28] 43.609581 60.783017 31.888046 56.723948  0.000000  0.000000 43.462158  0.000000  0.000000
[37]  6.426092 20.804456 60.206107 35.870201 20.782082  5.079144 20.889454 61.821549  8.189358
[46] 22.734228 32.352729 57.399778 33.938591  5.079144 57.638684  0.000000
> pv$bondperc
 [1] 36.57591 22.94282 53.95816 52.21913 66.74819 40.21653 66.55093 48.74593 48.42246 28.96099
[11] 34.47222 60.00610 77.33324 46.82007 57.17323 48.35442 47.60568 75.81072 58.26334 22.76753
[21] 42.66390 28.09591 53.13158 54.40320 68.50949 62.41348 58.35804 54.48651 37.16512 59.24027
[31] 41.61912 69.36583 22.94282 55.50703 61.70084 48.23367 40.16292 70.51730 38.01227 50.34864
[41] 61.92832 49.43298 69.41969 38.01227 79.79619 61.10474 58.00538 41.61912 62.75862 50.82181
[51] 41.84663 22.67565
> pv$cashperc
 [1]  0.000000 72.197606  6.451401 10.660824 15.255020  0.000000 20.535918  0.000000  0.000000
[10] 62.834219 59.019627 32.008931 11.212171  0.000000  6.702004  0.000000 29.662986 12.542426
[19] 31.106708 70.920046  0.000000 70.571304  0.000000  0.000000  9.058068  7.011598  6.397447
[28]  0.000000  0.000000  6.397447  0.000000 27.077929 76.798029  0.000000 37.507888 49.658886
[37] 51.923827  8.617647  0.000000 10.599624 14.570910 41.709035  7.384985  0.000000  9.104066
[46] 15.035064  6.849369  0.000000  0.000000 41.787023  0.000000 75.502713
```

Discuss, in relation to Table 1, the level of risk that seems to be taken by the brokerage houses and whether the one-year return performance is related to the associated risk of the brokerage house.

From the bh table,

```
> bh<-rbind(pv[1,],pv[44,],pv[40,])
> bh
      ROI(%)       RISK stockperc bondperc cashperc
1  12.269570 0.16370526  63.14612 36.57591  0.00000
44 12.217839 0.16258276  61.82155 38.01227  0.00000
40  8.751731 0.08685114  35.87020 50.34864 10.59962
```

I find that the row 1(63%,36%,0%),row44(61%,38%,0%),row40(35%,50%,10%) investment blend separately close to Lehman Brothers (65%,35%,0%),Everen (62%,33%,5%), Merrill Lynch(40%,50%,10%),related to the stocks,bonds,cash in sequence.

The R code is as follows:

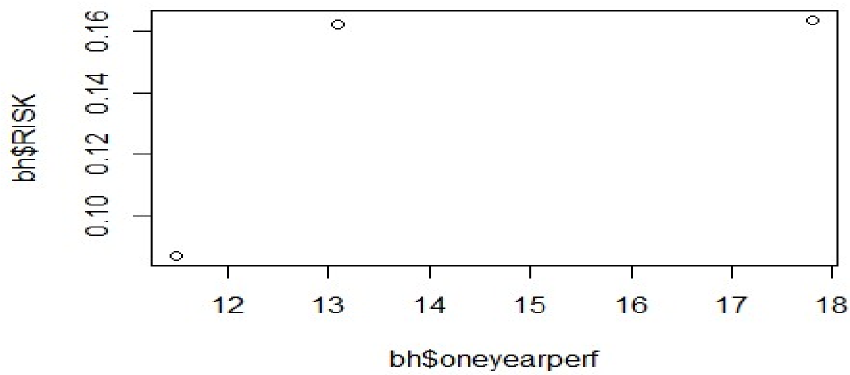
bh<-rbind(pv[1,],pv[44,],pv[40,])

bh$oneyearperf<-c(17.8,13.1,11.5)

Then add the one-year performance of the 3 different companies “Lehman Brothers, Everen, Merrill Lynch” into

bh <- bh [order(bh$oneyearperf),]

plot(bh$oneyearperf,bh$RISK)



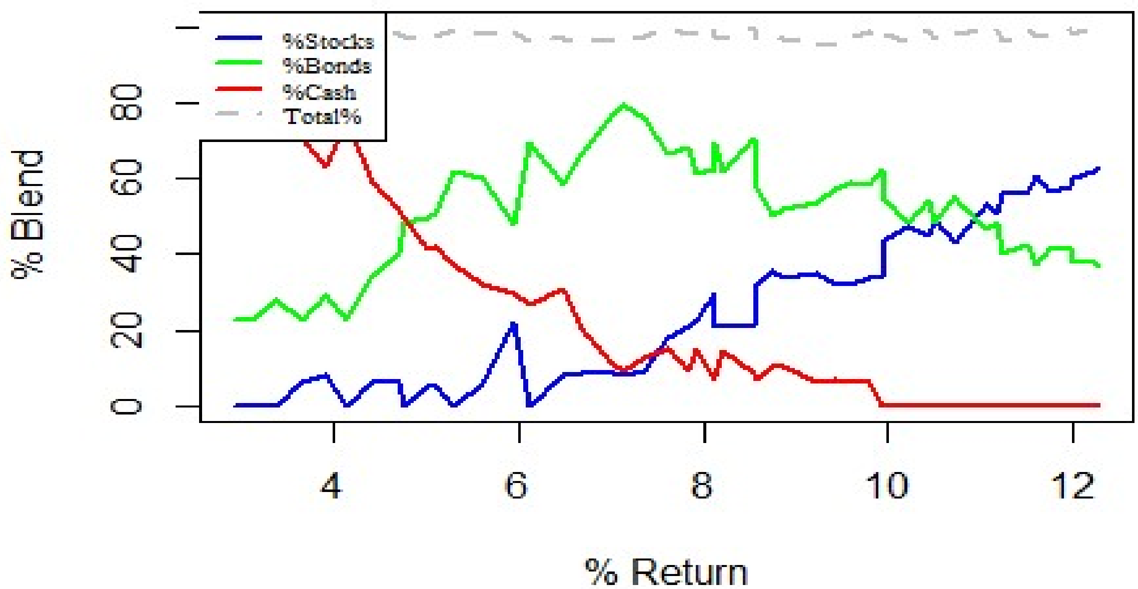
From the screenshot above, I find that the one-year return performance is not necessarily related to the associated risk of the brokerage house. Because the one-year return performance can be affected by many factors, not only be affected by the risk. For example, in the plot above, from the left side to the right side, the second point and the third point have a similar risk, but their one-year return performance has a very big difference. Maybe there are other factors that also affect the one-year return performance besides the risk.

4, Examine the plot shown in Figure 3. This shows how the percentage of bonds, stocks, and cash vary as you move along the Pareto front from the least to the greatest percentage return. Outline the approach (set of steps, algorithm, pseudo code, or R code…) that would be required to produce this figure given the output from nsga2.

The R code for this task is as follows:

```
newp<-portfolio$par

newp[which(newp<0.05)]<-0

sumstockp<-apply(newp[,1:10],1,sum)*100

sumbondp<-apply(newp[,11:20],1,sum)*100

sumcashp<-apply(newp[,21:30],1,sum)*100

sump<-sumstockp+sumbondp+sumcashp


newpv<-cbind(portfolio$value,sumstockp,sumbondp,sumcashp,sump)

newpv<-as.data.frame(newpv)

colnames(newpv)<-c("ROI","RISK","stockperc","bondperc","cashperc","totalperc")

newpv$ROI<-newpv$ROI*-1

newpv <- newpv [order(newpv$ROI),]
```

```
plot(newpv$ROI,newpv$totalperc,col="grey",ylim=c(0,100),xlab = "% Return",ylab ="% Blend",lwd=2,lty=2, type="l")

lines(newpv$ROI,newpv$stockperc,col="blue",lwd=2)

lines(newpv$ROI,newpv$bondperc,col="green",lwd=2)

lines(newpv$ROI,newpv$cashperc,col="red",lwd=2)

legend (2.5,104, legend = c("%Stocks", "%Bonds","%Cash","Total%"),col=c("blue","green","red","grey"),lty=c(1,1,1,2),lwd=c(2,2,2,2),text.font=6,cex=0.6)
```



5, Extend the model in invest.R to have an additional objective which is to minimise the correlation for the investment blend. Describe in words the approach you have chosen and explain the rational for the model. Include your R code for the function that calculates this objective. Produce 2 plots: 1. plot(portfolio) – this will give the 3 objectives as a pairwise plot; and 2. A plot showing the objectives correlation versus return (x-axis correlation, y-axis return) for the Pareto solutions, and discuss what type of return is likely to be best for minimizing correlation (and probably therefore the safest combination to consider).

Describe the approach and explain the rational for the model:

Based on the question require to minimise the correlation for the investment blend, explain the approach step by step as follows:

1, Find out the investments that are selected from an individual in the population (>= minAMOUNT).

2, For all the investments being selected, find out their correlation numbers with each other by doing a for loop:

- Find all the columns that contain the selected investments, combine all these columns, and build a new matrix named matrix 1.
- From matrix1, select all the rows that contain the selected investments, combine all these rows, and build a new matrix named matrix2.

3, Now matrix 2 includes all the correlation numbers of the selected investments, add all of them together to get a final correlation number. Because the nsga2 takes the minimise, we can take the positive value of the correlation number in the function.

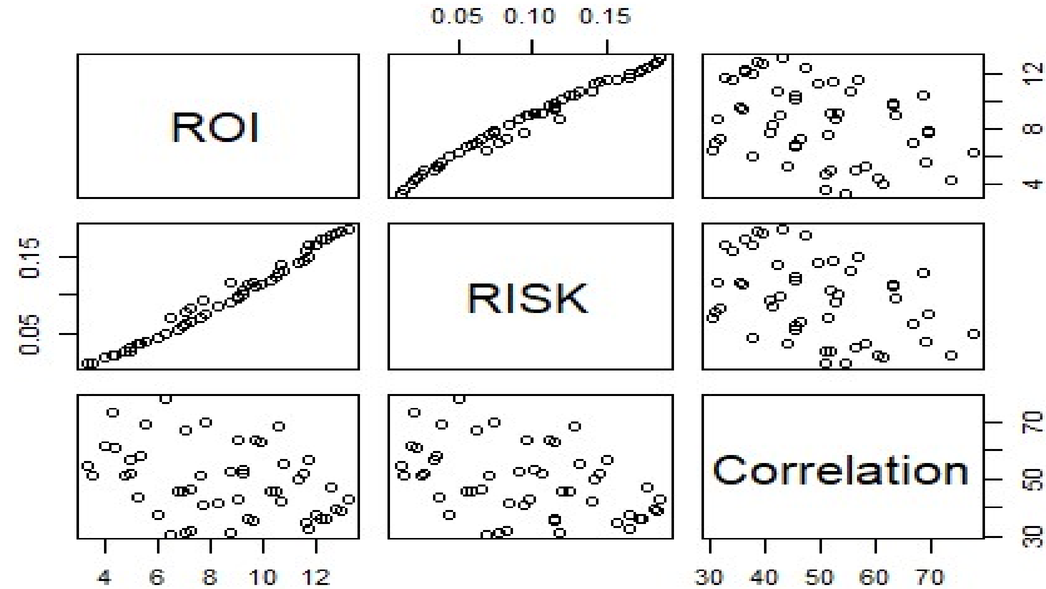
4, Make another function (for example funs3) to take maximum the ROI, minimise the risk, and minimise the correlation all into account.

5, Run the nsga2, this time needs to change to use funs3, also change the objective dimensions to 3 because now the objectives are ROI, RISK, and correlation.

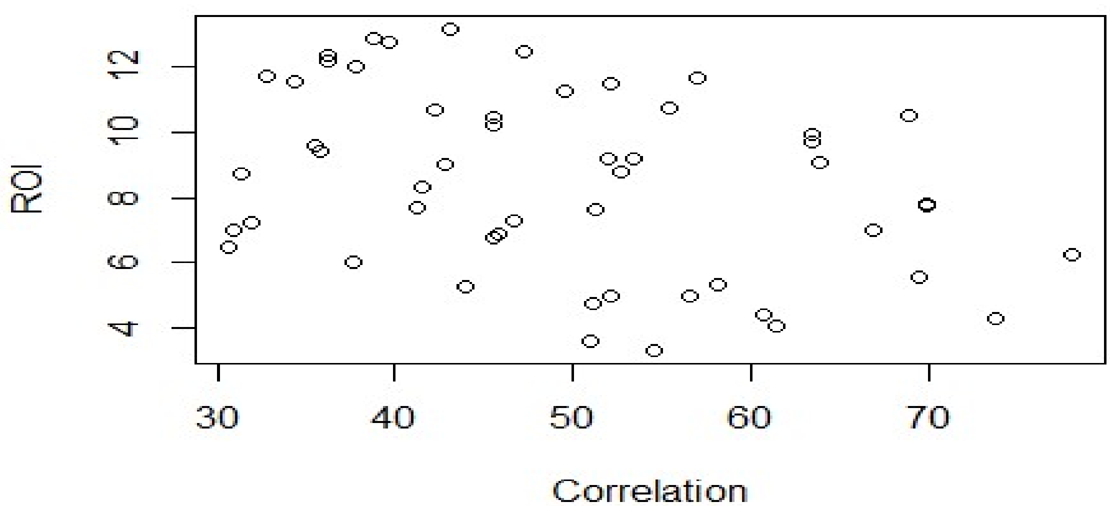
The R code for this task is as follows:

```
corr <- read.table("corr.tab")

resultcol<-matrix(nrow=30, ncol=30)

correlation <- function(x)

{

 selected <- which(x >= minAMOUNT)

 for (c in 1:length(selected))

 {

  resultcol[,c]<-corr[,selected[c]]

  newresultcol<-resultcol[,1:length(selected)]

 }

  table<-matrix(nrow=30, ncol=length(selected))

  for (r in 1:length(selected))

 {

 table[r,]<-newresultcol[selected[r],]

 newcorrtable<-table[1:length(selected),]

 }

 corrsum<-sum(newcorrtable)

  return(corrsum)

}
```

```
funs3 <- function(x)

{

  return(c(ROI(x), RISK(x), correlation(x)))

}


portfolio3 <- nsga2(funs3,

              idim=numberOptions,

              odim=3,

              popsize=52,

              generations=500,

              lower.bounds=lower,

              upper.bounds=upper,

              constraints = constraintFNS,

              cdim=3)

p3v<-portfolio3$value

p3v[,1]<-p3v[,1]*-1


colnames(p3v) <- c("ROI","RISK","Correlation")

pairs(p3v)
```



plot(p3v[,3],p3v[,1],xlab="Correlation",ylab="ROI")



The type of return is likely to be best for minimising correlation:

mincorr<-min(p3v[,3])

minrows<-which(p3v[,3]==mincorr)

mincorrreturn<-p3v[minrows[1],1]

```
> mincorrreturn
     ROI
6.481301
```