COMP396 Design Report

Team 7 Contributors:
Zixuan Zhu,
Yiwang Tian,
Weiyi Liu
Tianyi Ye,
Zhiyu Chen

Contents

Section1 Momentum strategy with double moving average and MACD indicators

Section 1.1 Strategy introduction

This strategy is a type of momentum strategy and is trend-following, meaning that the more trendy the stock market is, the more effective this strategy is and the more it can be used to its maximum effect.

Section 1.2 Data analysis

As this strategy is a market that requires a strong trend, the data analysis and visualisation of the overall 10 stocks can show the market that is suitable for this strategy - i.e. there is a clear trend.

1. volatility analysis

Based on an analysis of the volatility of the overall data opening and closing prices, it is clear that asset number one two eight is the most volatile (risky) and number six seven ten is the least volatile (risky). (This part of the data analysis in python).

open_to_open

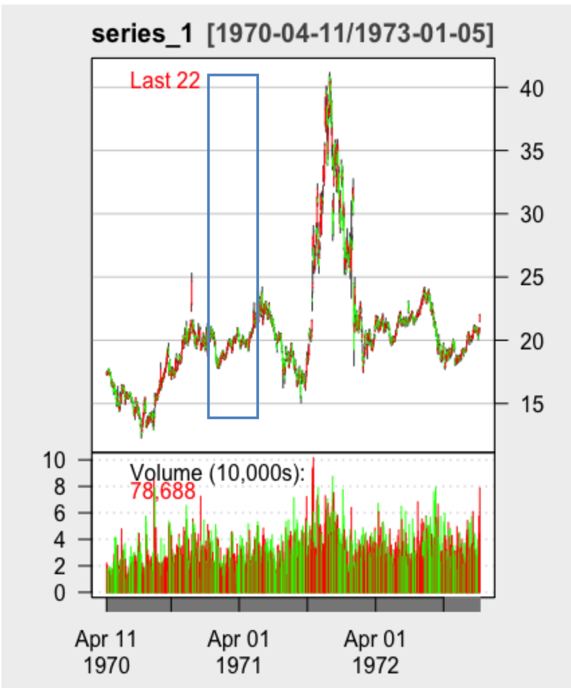
	mean	std	min	25%	50%	75%	max
1	0.000220	0.026972	-0.241745	-0.013458	0.000919	0.013596	0.135427
2	0.000116	0.027669	-0.265703	-0.012202	0.000000	0.013615	0.150823
3	-0.000101	0.017751	-0.120953	-0.009527	0.000000	0.009431	0.099018
4	-0.000009	0.016194	-0.115143	-0.009069	-0.000844	0.008307	0.116433
5	-0.000176	0.018291	-0.096515	-0.008249	0.000000	0.008193	0.134442
6	0.000071	0.003032	-0.014997	-0.001546	0.000344	0.001781	0.016323
7	-0.000055	0.007012	-0.041964	-0.003590	-0.000392	0.003531	0.033572
8	-0.000307	0.020870	-0.145016	-0.010480	0.000000	0.010861	0.091937
9	-0.000281	0.013734	-0.085507	-0.008047	0.000752	0.008368	0.059285
10	-0.000021	0.007508	-0.056541	-0.002825	0.000590	0.003675	0.032374

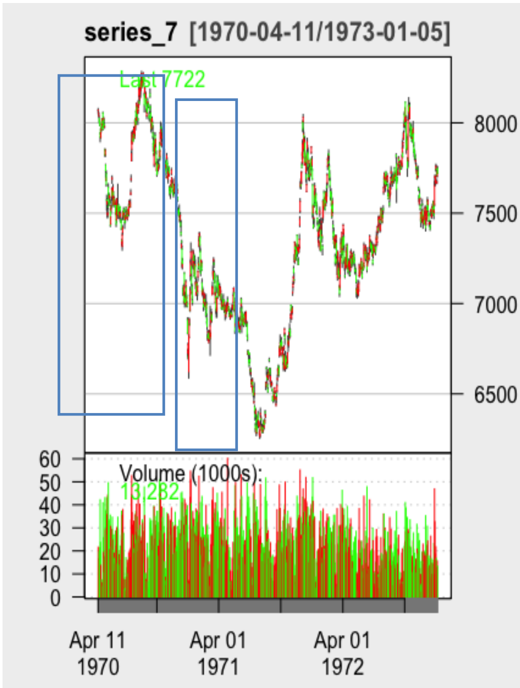
close_to_close

	mean	std	min	25%	50%	75%	max
1	0.000198	0.029452	-0.400478	-0.012132	0.000765	0.012462	0.135724
2	0.000111	0.029703	-0.390942	-0.010572	0.002086	0.013296	0.139942
3	-0.000104	0.016854	-0.123481	-0.008226	0.000392	0.008395	0.072225
4	-0.000008	0.015147	-0.073338	-0.008717	0.000502	0.007673	0.064143
5	-0.000179	0.019207	-0.135403	-0.007468	0.000277	0.006989	0.238970
6	0.000070	0.002888	-0.017042	-0.001279	0.000360	0.001652	0.019916
7	-0.000055	0.007083	-0.041325	-0.003808	-0.000398	0.003603	0.039457
8	-0.000315	0.020250	-0.110105	-0.010047	0.000000	0.010342	0.084023
9	-0.000275	0.013152	-0.056242	-0.006312	0.000000	0.007061	0.054067
10	-0.000026	0.007081	-0.060455	-0.002716	0.000241	0.003129	0.027419

2. k-chart visualisation

I visualised the known 10 series of stocks using k-line charts, such as these three charts below, where a clear trend exists.





3. *Trend detection*

Trends in time series data were analysed using the Mann-Kendall test.

```
Mann_Kendall_Test(trend='increasing',
  h=True, p=0.0, z=20.254244713045708,
  Tau=0.40777400943006037, s=246479.0,
  var_s=148089483.0, slope=0
  .005197889182058045, intercept=16
  .943759894459106)
```

series 1

```
Mann_Kendall_Test(trend='decreasing',
  h=True, p=2.221112183065088e-12,
  z=-7.019845064778539, Tau=-0
  .14133013483331955, s=-85427.0,
  var_s=148089780.33333334, slope=-0
  .3273809523809524, intercept=7664
  .895833333333)
```

series 7

Section 1.3 How the strategy works

1. *Design concept*

This strategy combines two simple moving average lines and MACD indicators to clearly identify buying and selling actions in trending markets.

(1)Double Moving Averaging indicators

Firstly, I considered that if it is a trending and less volatile market, there is a great need for an index of trend indicators. Based on what I learnt in year 2 of COMP226, the first indicator that came to mind was the moving average indicator, which is a digital low-pass filter and therefore gives a visual indication of the general trend of the stock market. Next, I chose the simple moving average (short for SMA) because I thought that all other things being equal, SMA can give a good indication of the general trend of a time series. If a clear entry signal is needed, I think two SMAs would be more convincing, and traders can obtain objective signals that reflect the strength of the market.

However, the averaging indicator also has the disadvantage of tending to lag, and is often only reflected in the averages after a significant pullback in price. In volatile markets, frequent enters and exits might result in losses.

(2)Moving Average Convergence/Divergence (MACD)

It is also a trend-following momentum indicator. It depicts the connection between two exponential moving averages (EMAs) of the price of securities(Dolan, 2022). It is worth noting that the MACD indicator uses the same concept of averaging as the SMA, but the MACD uses exponential decreasing weighting to calculate the mean. According to my knowledge of COMP226, it is weighted by time on the basis of the moving average, with prices closer to the current date having a greater influence on future prices and given greater weight, and being more sensitive to price changes. As a result, adding the MACD indicator to the entry criteria can ease the SMA's latency and eliminate the fault that the SMA occasionally sends out false signals, making the indicator more accurate in detecting medium and long-term trends.

In R, I called the SMA and MACD functions of the TTR package in the following code form

```r
short_ma <- last(SMA(store$cl[startIndexma:store$iter,i],n=params$dmalookbacks$short))
long_ma <- last(SMA(store$cl[startIndexma:store$iter,i],n=params$dmalookbacks$long))
macd <- as.data.frame(last(MACD(store$cl[startIndexmacd:store$iter,i],
                                nFast=params$macdFast, nSlow=params$macdSlow,
                                maType=params$macdMa, percent=TRUE)))
```

It is worth explaining that the result of MACD() is a matrix containing columns, two of which are "macd" and "signal".

The basic logic of the strategy is therefore that when the short SMA crosses the long SMA and macd$macd crosses macd$signal from bottom to top at the same time, this is a clear signal to long and indicates a strong uptrend ahead; when the short SMA crosses the long SMA and macd$macd crosses macd$signal from top to bottom at the same time, this is a clear signal to leave or short and represents the existence of a downtrend and a short market.

What's more, if we analyse indicators, if the macd line diverges from the K line, there will be a possible reversal of the market. The crossover method represents a change in momentum and can determine the point of entry or exit into the market(La Ferla, 2019).
In summary, this strategy requires investors to be patient and wait for trade signals to appear, as this strategy is trend-following. Therefore, the signal may be a long time period and I consider it a long-term trade.

The charts below show these three indicators clearly on the k-chart.

series 1

series 7

2. *Logic and code*

```
#when there is a cross in double moving average and macd indicators
#it is a clear entry point and it also implies the future up-ward trending
if (last((SMA(store$cl[startIndexma:(store$iter-1),i],n=params$dmalookbacks$short)))
    <
    last((SMA(store$cl[startIndexma:(store$iter-1),i],n=params$dmalookbacks$long)))
    &&
    short_ma >= long_ma
    && macd$macd >macd$signal ) {
  pos[params$series[i]] <-  params$posSizes[params$series[i]]
}


#when there is a cross in double moving average and macd indicators
#it is a clear exit point and it also implies the future down-ward trending
else if (last((SMA(store$cl[startIndexma:store$iter-1,i],n=params$dmalookbacks$short)))
         >
         last((SMA(store$cl[startIndexma:store$iter-1,i],n=params$dmalookbacks$long)))
         &&
         short_ma <= long_ma
         && macd$macd < macd$signal
         ) {

  pos[params$series[i]] <- -params$posSizes[params$series[i]]

}
else{
  pos[params$series[i]] <- 0
}
```

Section 1.4 Testing

1. Preliminary testing

1.1 Initial parameter selection and optimization.

(1) parameter selection

From my perspective, in the double SMA strategy, if the periods of the two SMAs are close together, they will be very prone to entanglement, constantly generating buy and sell points and there will be a large number of invalid trades, resulting in high trading costs. If the period of the two averages is wide apart, the trend of the averages with very long trading periods will lag as they are not obvious and there will be a lag when buy and sell points appear, thus potentially resulting in a large loss.I realised that the two parameters chosen should not be too far apart, nor should they be too close together, so for the initial two SMA lookbacks I chose 5 and 10 for the short SMA and 50 and 60 for the long SMA, while the macd parameters were chosen as default values and nFast = 12, nSlow = 26, nSig = 9 and lookback is 50.

Since this strategy is back-tested on series1 and series7, the following results are taken from stock 1 and 7.

1.2 Parameter optimization

First, I use the data from days 1 to 500 as the training set (i.e., in-sample data set), and the data from days 501 to 1000 as the test set (i.e., out-of-sample data set). In the meantime, I have used the position sizing management methodology used in COMP396 Lecture 4 to determine the extent to which individual series stocks contribute to overall profits by using the difference of open price moves. Analysis of the data shows that series7 is the benchmark and the position size for series1 is increased to 323 units. I have shown the trading results of the 10 series with adjusted positions in the position sizing section of the combined strategy section.

Secondly, I used grid search to find all the combinations of parameters and cross-validation to find the best combination of returns in the training set.

```
[1] "Parameters:"
$lookbacks
$lookbacks$short
[1] 5

$lookbacks$long
[1] 50


Profit: 4312.054
```

```
[1] "Parameters:"
$lookbacks
$lookbacks$short
[1] 10

$lookbacks$long
[1] 50


Profit: 654.642
```

```
[1] "Parameters:"
$lookbacks
$lookbacks$short
[1] 5

$lookbacks$long
[1] 60


Profit: 2121.956
```

```
[1] "Parameters:"
$lookbacks
$lookbacks$short
[1] 10

$lookbacks$long
[1] 60


Profit: 2706.102
```

It is clear to see that the combination of parameters (5, 50) performs the best, so I used that in the out-of-sample test. The total result for series 1 and series 7 in out-of-sample is:

```
[1] "Parameters:"
$lookbacks
$lookbacks$short
[1] 5

$lookbacks$long
[1] 50


Profit: 10756.56
```

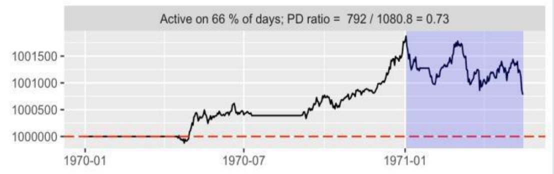
1.3 Backtesting and results analysis

The combination of parameters resulting from the above parameter optimisation is now used for backtesting. The following diagram shows the parameters of the entire strategy.

```
"dmaMACD"=list(dmalookbacks = list(short=as.integer(5),
                                   long=as.integer(50)),
               macdlookback = as.integer(50),
               macdFast=12, macdSlow=26, macdSig=9, macdMa="EMA", series = c(1,7),
               posSizes=as.numeric(list(323,6989,40,374,47,8244,1,6681,14,83))),
```

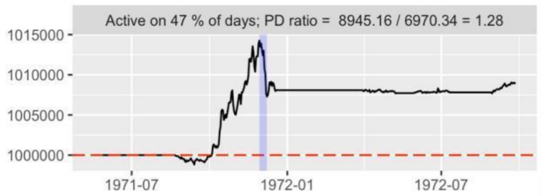
Run main.r, the following pictures are from the in-sample data set.

series 1

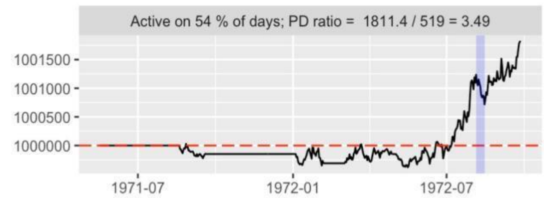


series 7

The following pictures are from out-of-sample data set.



series 1



series 7

Section 1.5:Plan

1. Overall reflection

In the code implementation of this strategy, I first considered the use of market order, not limit order, but if I wanted to make more profit with limited resources, perhaps limit order could help me to do so. However, as there are still some technical and logical problems with "how to build a limit order into the code", the whole group needs to follow up and discuss.

Secondly, as this strategy is suitable for trending markets, it does not evaluate well in mean-reversion markets and oscillating markets, and carries a high degree of risk.

Finally, based on the results of the in-sample, the retracement values are still large, suggesting that the strategy can only provide the most basic trading ideas for the time being.

2. Risk management solutions - Stop-loss settings.

There is no stop loss set in this strategy. Throughout the series (in-sample data), the results of the oscillating market are not objective. The strategy's signals may sometimes be incorrect, and stop-loss settings are needed to avoid larger losses. For example, if in a strong trending market, a buy signal has been generated and has entered the market, but has fallen for three consecutive days and has not yet triggered a short exit signal, then a stop loss is required to clear the position and leave early. But if in a weak trend market, I need to consider a shorter period of time to determine whether there is a stop-loss exit opportunity. Therefore, in the next time, if I continue to follow this strategy, I will carefully study the stop-loss settings in different market conditions to maximize profits as much as possible.

Section2 Resistance Support Relative Strength Strategy

Section 2.1 Strategy introduction

Trading strategy introduction - The concept of support and resistance

The resistance is when the price moves up, the highest point reached before it pulled back. It is the price level indicating where there will be a surplus of sellers and making prices fall.

The support is when the price fall, the lowest point reaches before it starts back, it is the price level where traders believe that there is a surplus of buyers and making prices rebound.

Section 2.2 Data analysis

1. Common strategy of support and resistance levels --- Bollinger bands

Bollinger Bands is one of the common strategy of resistance and support. The strategy will be long (short) whenever the close is below (above) the lower (upper) Bollinger Band. And we tried “bbands_contrarian” in examples given us.

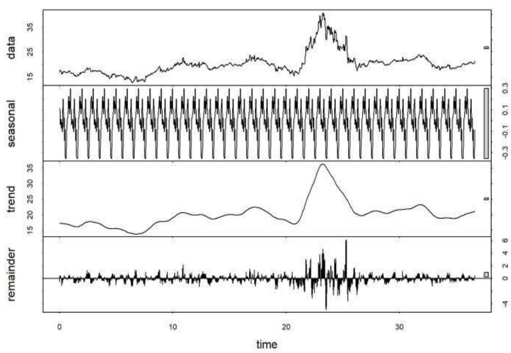
```
#main strategy logic
if (store$iter > params$lookback) {
  startIndex <-  store$iter - params$lookback
  for (i in 1:length(params$series)) {
    cl <- newRowList[[params$series[i]]]$Close
    bbands <- last(BBands(store$cl[startIndex:store$iter,i],
                          n=params$lookback,sd=params$sdParam))
    if (cl < bbands[,"dn"]) {
      # if close is relatively low go long (i.e., contrarian type)
      pos[params$series[i]] <- params$posSizes[params$series[i]]
    }
    else if (cl > bbands[,"up"]) {
      # if close is relatively high go short (again, contrarian type)
      pos[params$series[i]] <- -params$posSizes[params$series[i]]
    }
  }
}
marketOrders <- marketOrders + pos
```

How to find an effective resistance or support level is the key to the effective profit strategy. From the performance of the traditional bollinger bands strategy, it is not difficult to find that the strategy uses the support and resistance levels itself as a threshold.

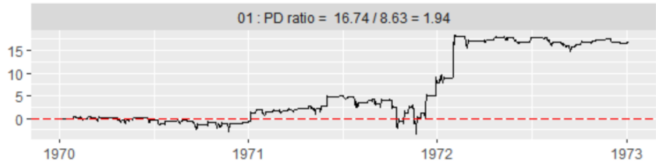
2. Data analysis and compare

Take Stock 1 as exampled, we plotted its K-chart observing its movements and added ATR indicators (**ATR represents the mean of the true fluctuation amplitude, and is a volatility indicator used to measure the price. The higher the ATR index, the greater the chance of the price trend reversal, and vice versa)*. We also plotted The Classical Decomposition Model, through this model, we can clearly display the overall trend and the fluctuation of this stock.

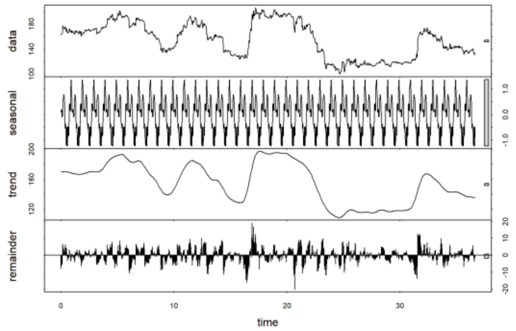
We can see that there is only a clear trend and ATR remains small numbers in stock 01. It experiences an obvious trend in prosperity and downturn in 1972, and stock price remains stable in other observation years.



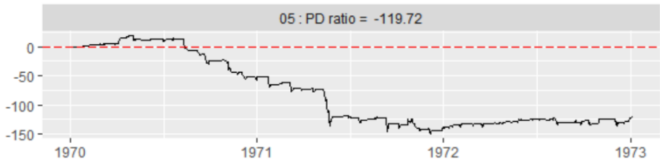
Following the general trend, the type of 'fixed value' strategy is effective and can get returns.



However, at the same time, its disadvantage is also very obvious: time lag. As the middle rail of the BBands is actually a moving average, it cannot respond to multiple changes with fast and sharp trends. In uncertain markets, using this approach means constantly switching between long and short, with little chance to generate profit. This makes it constantly lose money in the volatile market. For example, stock 05 has multiple fluctuations while the ATR is also large overall compared to stock 01.



As shown in the figure, for stocks with frequent volatility and high uncertainty(05), the bollinger bands strategy is in a continuous loss state.



Section 2.3 How the strategy works

1. Inspiration: Resistance Support Relative Strength (RSRS) indicator

Inspired by the bollinger bands, rather than merely interpreting the resistance and support levels as the upper and lower limits of the price range, we investigate another way to use them that focuses on the relative strength between them. In other words, we no longer think of resistance and support as a fixed values. However, they are thought to be variables that reflect what traders perceive to be the top and bottom of the present market situation with good immediacy.

We illustrate the logic of the relative strength of support resistance according to different market states:

(1) when the market is rising:

- support > resistance, the bull market continues and prices accelerate rise
- resistance < support, the bull market may come to an end and prices will peak

(2) when the market is volatile:

- support > resistance, the bull market may be about to start
- resistance < support, a bear market may be about to start

(3) when the market is going down:

- support > resistance, the bear market may be coming to an end and prices will bottom out
- resistance < support, the bear market continues and prices accelerate fall

2.RSRS strategy specify

2.1 Definition of RSRS and quantification of relative intensity

After considering the given data in our project and some existing indicators, we believe that the daily highest and lowest prices are a good way to meet this demand.

Daily highs and lows can be considered as the resistance and support level which is recognized by the trading behavior of all market participants on that day. The daily highs and lows can quickly reflect the nature of recent market attitudes toward resistance and support levels, which is the most important reason we use them.

The relative strength of support and resistance is described by the value of the relative position change, *delta(high)/delta(low)*, i.e., how much the high price moves for every 1 change in the low price.

delta(high)/delta(low) is the slope of *(low[0],high[0])* and *(low[1],high[1])*. However, the slope obtained through the two points contains too much noise due to the presence of noise in the market volume itself. Therefore, we consider building a linear regression model for the last *N(low, high)* data points to obtain the relative degree of change of the highest and lowest prices with a high signal-to-noise ratio.

$$high = alpha + beta*low + epsilon, epsilon \sim N(0,sigma) \quad (1)$$

When the slope is large, it means the support strength is stronger than the resistance strength; otherwise, it indicates the resistance strength is stronger than the support strength.

2.2 RSRS index construction

Calculation of everyday slope index:
(1) Take the highest price sequence and the lowest price sequence of the previous N day.
(2) OLS linear regression of the two lists of data(high & low) according to formula ①.
(3) The fitted beta value was taken as the RSRS slope index value of that day

2.3 The RSRS Index trading strategy

We use threshold trading logic as the trading framework. We go long when the RSRS exceeds a threshold S1, and go short to close positions when the RSRS crosses the threshold.

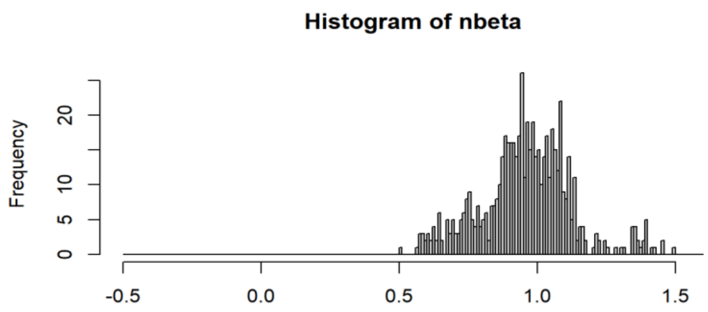
Section 2.4 Testing

1. Preliminary testing

1.1 Initial parameter selection and optimization

The beta is the slope we need. N cannot be taken too little or too large as we want to reflect the current market support and resistance relative strength. If N is too large, the lag will be too high; if N is too small, it won't be able to filter out enough noise. Here we choose N = 18.

Afterwards, we need to determine the upper and lower thresholds. To find a more reasonable threshold value, we calculated the slope for the sample-test, and observed the historical data distribution of the slope (calculated in N=18):



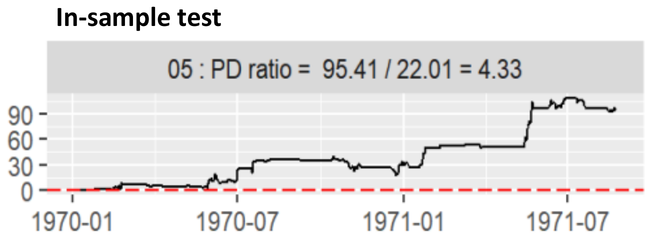
Mean	0.958
Standard Deviation	0.165
Skewness	0.167
Kurtosis	0.680

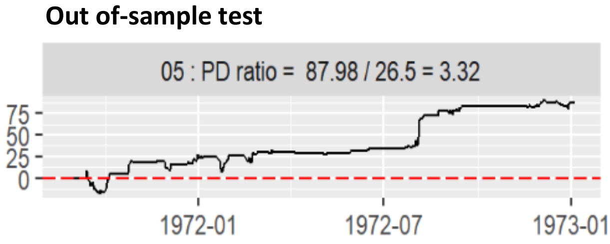
From the data, a more reasonable threshold selection is the mean plus and minus a standard deviation, we take S1=1.1 and S2=0.8.
Then, the RSRS slope trading strategy is:
(1) Calculate the RSRS slope.
(2) If the slope is greater than S1, long and hold.
(3) If the slope is less than S2, short and close the position.

```
#main strategy logic
#Iterate through the series in params$series
for (i in 1:length(params$series)){

  #start from the 19th day
  if(store$iter > params$lookback){
    #Get every stock's high and low price data
    High = store$high[,i]
    Low = store$low[,i]

    startIndex <- store$iter - params$lookback

    #iterate every 18 days
    for (j in startIndex:store$iter){

      HighList <- High[startIndex:(store$iter-1)]
      LowList <- Low[startIndex:(store$iter-1)]

    #OLS linear regression
      fit=lm(HighList~LowList)
      rsrs <- as.numeric(fit$coefficients[2])

    #decide if we should go long/short/flat (returning 1/-1/0)
      if (rsrs < 0.8){
        pos[params$series[i]] <- -1
      }
      else if (rsrs > 1.1){
        pos[params$series[i]] <- 1
      }
      else{
        pos[params$series[i]] <- 0
      }

    }
  }
}
```

2. Backtesting and results analysis

To better test our idea, we divided the data into in-sample (1:600) data and out-sample (601-1100) data, and tested the RSRS strategy on stock 5:





Section 2.5 Plan

Standard score optimization

In the above slope threshold strategy, the upper and lower thresholds are selected with reference to the mean and standard deviation of the historical data of the slope. In practice, however, there is no guarantee that the mean and standard deviation of future data will not change as the market evolves. Traders may also be more interested in where the market is in the recent environment, or how the market will develop over the next period compared to the current one.

If we look at different periods of the market, we can see that the mean value of the slope fluctuates relatively widely. For example, if we use data from the year before the trading day (50) to calculate the mean, we can find that the mean will fluctuate between 0.6 and 1. Therefore, it occurs to us that standardizing the RSRS would provide more flexibility to accommodate the recent overall market fundamental state. By using the RSRS standard score instead of the slope value as the indicator value, we actually increase the freedom of the strategy with the parameters M*(the period used to calculate the standard score)* and S *(the opening and closing threshold)*.

The improvement of the RSRS standard score strategy is done as follows.
The slope is normalized and the standard score is taken as the indicator value.
1) take the slope time series of the previous M days. 2.
2) use this sample to calculate the standard score of the slope of the day.

$$RSRS_{std} = \frac{RSRS - \mu_M}{\sigma_M}$$

Where μM is the mean value of the slope of the previous M days, σM is the standard deviation of the previous M days, and RSRS is the indicator value of the slope of the previous N days.

3) The calculated standard score z is used as the RSRS standard score indicator value for the day.

Section3 Multi-factor strategy

Section 3.1 Strategy introduction

A multi-factor strategy is a stock-picking strategy by combining several factors for analysis. Factors are based on certain yield-related indicators (e.g. High, Close, Volume...). It calculates a value and compares it to a specified threshold to go long or short a certain number of stocks. For example, we calculate the value of the indicator for each day and compare it to a threshold value endowed. We go long when the indicator value is greater than the threshold value and go short when it is less. The key to the multi-factor model is finding the relationship between factors and returns to make favourable trading decisions.

The key factor in our strategy is modified from the original Alpha-006. The basic logic of this factor is to go long when the stock price and volume are negatively correlated. We first calculate the correlation between the close price and the volume in the most recent *n* days, where *n* is the observed time. The higher the correlation, the closer the opening price and trading volume are to "rise-rise". As we need to find the circumstance of “rise-low” or “low-rise”, we multiply the factor by -1, resulting in the following equation:

$$\alpha = -1 * correlation(close, volume)$$

$$where\ observed\ time = most\ recent\ n\ days$$

Section 3.2 Data analysis

Alpha-006 calculates the correlation between the close price and the corresponding volume. Thus, before applying this factor, we must first make sure that there exist some levels of correlation between the close price and volume on a particular day in the given data. In addition, as we introduce the parameter of observed time *n* in factor Alpha-006, we need to analyse how to choose the value of n based on the volatility of the given data.

Part A – Prove of correlation between close price and volume

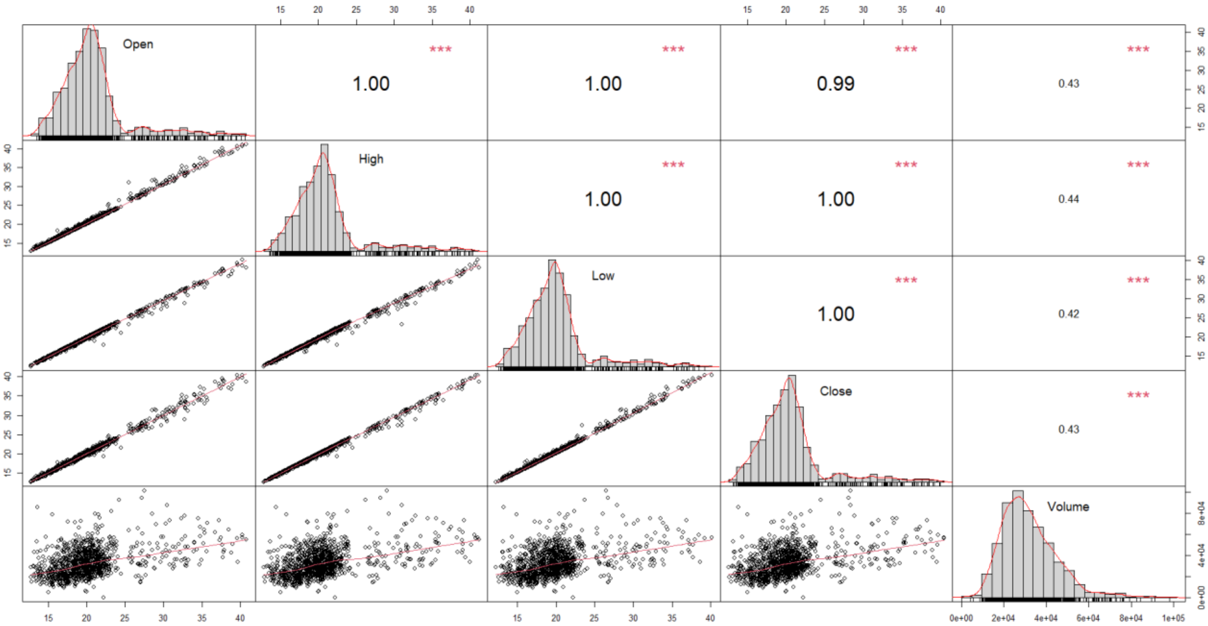
We use the following R code to analyse the correlation between the given indicators in stock. Set stock 01 as an example.

```
#Additional - Correlation Analysis
Stock <- read.csv("10.csv", header=TRUE, sep=",", dec=".", fileEncoding="UTF-8-BOM")
head(Stock)
#Remove the date index
Stock <- Stock[,-1]
head(Stock)
#Use of function cor()
cor(Stock$Close, Stock$Volume, method='pearson')
#Use of function rcorr()
library(Hmisc)
Stock <- as.matrix(Stock)
rcorr(Stock)
#Visualization
library(PerformanceAnalytics)
chart.Correlation(Stock)
```

We get the display of correlation between indicators:

```
> rcorr(Stock)
       Open High  Low Close Volume
Open   1.00 1.00 1.00  0.99   0.43
High   1.00 1.00 1.00  1.00   0.44
Low    1.00 1.00 1.00  1.00   0.42
Close  0.99 1.00 1.00  1.00   0.43
Volume 0.43 0.44 0.42  0.43   1.00
```

And a graph that clearly visualized the correlation:



From the graph displayed, there exist high levels of correlation between indicators of stock 1. By conducting the same operations from stock 01 to stock 10, the correlation index of close price and volume are 0.4284853, 0.4394195, 0.0691163, 0.04081823, 0.1069442, -0.07284416, -0.0242775, 0.07110634, 0.4357325, -0.15527.

Part B – The stocks' volatilities and the observed period setting

In the previous data analysis section in momentum strategy, we used Python to calculate the volatilities of the given ten stocks and made a ranking. The result shows that stock no.1, 2 and 8 have the greatest fluctuation while stock no.6, 7 and 10 have the smallest.

For stocks with relatively stable prices, we tend to set the observed period longer so that the decision on a particular day can reference more previous day's data, resulting in a more reliable trading decision. However, for the stocks with high volatility, the decision using a long-observed period will be greatly affected by fluctuations and become error-prone. Thus, our strategy with a long-observed period input is supposed to work well on stock 6,7,10 and badly on stock 1,2,8, and vice versa.

Section 3.3 How the strategy works

1. Original Factor formula

Original Alpha006: (-1 * correlation(close, volume, 10))

2. Factor function description

For *correlation(x,y,d)* part, the correlation coefficient of two random variables x,y in the past d days, its value range is [-1,1]. For this factor, it can be interpreted as the daily opening price of the stock and the volume of the correlation coefficient in the past d days.

3. Explanation of the formula

First, the correlation between the opening price and the volume of the last 10 trading days is calculated. The higher the correlation, the more the opening price and volume tend to be the same. The final result is then added on the negative sign, which means that the higher the value of the factor, the lower the correlation between the stock's opening price and volume over the past 10 trading days.

4. Formula logic

We observe a phenomenon of securities trading when a new peak in the price of securities occurs, the volume does not increase but begins to fall. This means that the price of the security is not proportional to the volume of the relationship. In other words, the stock's price has fallen to a minimum and is about to end its downtrend and start an uptrend. This is a buy signal and investors may consider going to buy a certain number of shares at this time. Conversely, when the stock price continues to rise in phase, the volume of the stock continues to fall. This means that the stock price has reached its peak and is about to end its uptrend and start a downtrend. This is a sell signal and investors may consider going to sell a certain volume of the stock. When the above two phenomena are not obvious, it means that the stock does not yet have a strong buy or sell signal and investors can temporarily hold off on trading waiting to trade the stock from the next node where the change is obvious (Chen, 2021).

5. Strategy Setting

There is a threshold, which is imported to the strategy function from parameters.

(1) When the factor value of the day is greater than the threshold, it means that the stock price has fallen to the lowest and is about to end the downtrend and open an uptrend. At this point we should buy, so set the position to 1.
(2) When the factor value of the day is less than the threshold, it means that the stock price has reached its peak and is about to end its uptrend and start a downtrend. At this point, we should sell, so set the position to -1.
(3) When the factor value of the day is exactly equal to the threshold, there is not a clear signal to buy or sell. We should not trade at this point, so set the position to 0.

```
getOrders <- function(store, newRowList, currentPos, info, params) {

  #Initializing
  allzero  <- rep(0,length(newRowList)) # used for initializing vectors

  if (is.null(store)) store <- initStore(newRowList,params$series)
  store <- updateStore(store, newRowList, params$series)

  marketOrders <- -currentPos; pos <- allzero

  #Initialize threshold value
  thr <- params$thr

  #Iterate through the series in params$series
  for (i in params$series){

    #Ignore the first day, or it shall occur error
    if(store$iter>1){

      #Get every stock's volume and closed price data
      VOLUME = store$vol[,i]
      CLOSE = store$cl[,i]

      #For the first specified days, store the first day to today's volume and close price
      if(store$iter<=params$obday){
        VOLUMELIST <- VOLUME[0:store$iter]
        CLOSELIST <- CLOSE[0:store$iter]
      }

      #After the specified days, store the most recent n days' data
      #n is the observed day and is passed in through parameter "obday"
      else if(store$iter>params$obday){
        VOLUMELIST <- VOLUME[as.numeric(store$iter-params$obday):store$iter]
        CLOSELIST <- CLOSE[as.numeric(store$iter-params$obday):store$iter]
      }

      #Apply Alpha006 equation
      #Get Everyday's new alpha rate
      alpha = -1*cor(as.vector(CLOSELIST), as.vector(VOLUMELIST), use = "everything", method="pearson")
```

```
      print(paste("day", store$iter))
      print(paste("series no.", i))
      print(paste("alpha006 =",alpha))

      #Change Position
      if (alpha*100 < thr){
        pos[params$series[i]] <- -1
      }
      else if (alpha*100 > thr){
        pos[params$series[i]] <- 1
      }
      else if (alpha*100 == thr){
        pos[params$series[i]] <- 0
      }

    }
  }

  #Update market orders
  marketOrders <- -currentPos + pos
```

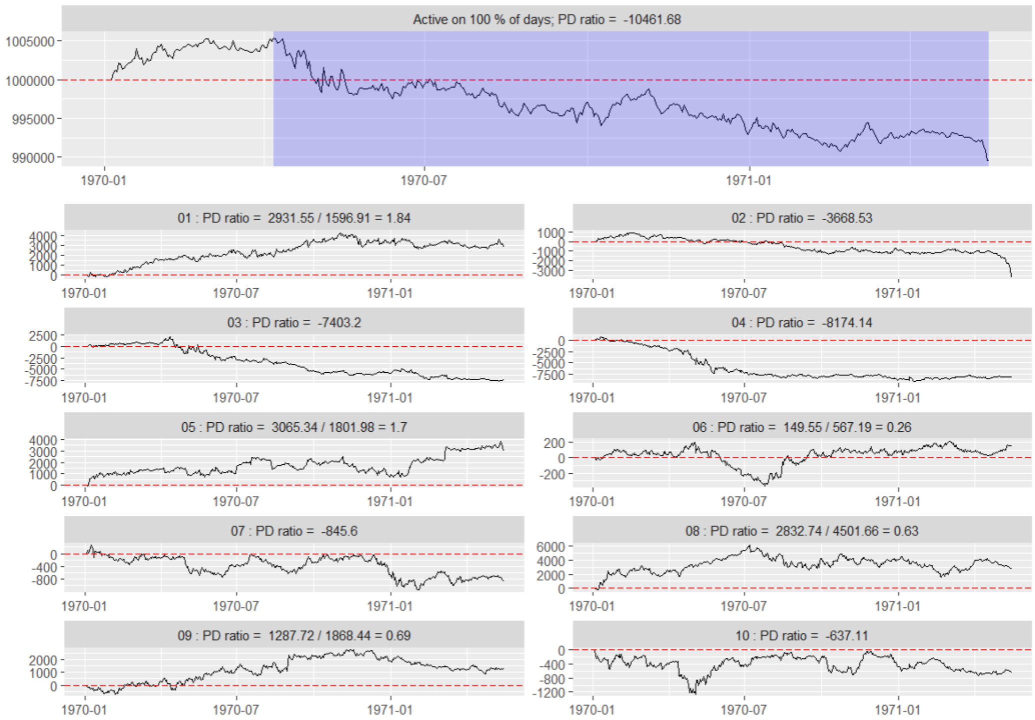
Section 3.4 Testing

1. Strategy Implementation

We choose the first 500 days as an in-sample test

```
numDays <- nrow(dataList[[1]])
inSampDays <- 500

# in-sample period
dataList <- lapply(dataList, function(x) x[1:inSampDays])

# out-of-sample period
dataList <- lapply(dataList, function(x)
                              x[(inSampDays+1):numDays])
```

For example, we set the parameters as thr=0 and obday=35. In addition, we need to guarantee all series contribute a similar amount to the overall result. Thus, we use the knowledge of COMP396 lecture 4 to set parameters to make position sizing. It then calculates the PD-ratio result of the ten stocks.



2. Strategy Optimisation

By the definition of the formula, the calculation should be based on the previous *n* days (observed day) of stock trading data. This means that in order to run the strategy, the stock would have to start trading after *n* days. So we made a slight change to the strategy and ran the logic as follows: the strategy starts on the second day and uses all previous trading data as parameters for the calculation of the factors each day. After the *nth* day, it then changed to use the previous n days of the trading data. This effectively avoids the disadvantage that the strategy cannot be conducted on the first *n* days.

Section 3.5 Plan

We have tried to develop a better strategy. But due to some technical or time issues, we are unable to achieve that in this version of the coding.

1. Inability to determine the optimal threshold

Our first thought for this strategy is to find the optimal threshold. A previous version of the code is to loop the threshold from -100 to 100 and simulate the trading, record the highest PD-ratio and the corresponding threshold, then use the best threshold to conduct the strategy again. However, our code meets infinite recursion issues. We need to call backtest() function inside the getOrders() function to simulate the trading and get the PD-ratio, while inside backtest() function it calls the getOrders() function again to conduct the strategy. Thus, we cannot use a functional approach to optimally select the threshold and have to manually set the threshold as a parameter in the current version. We will try to find a solution to solve the infinite recursion issue and try to functionally find the optimal threshold in the future.

2. The incorrectly choose of observed day on some occasions.

Another parameter of this strategy, besides the threshold, is the observed day. In this version of strategy design, we calculate the overall volatility of a stock and determine whether to set a long *n* or a short *n*. However, if there is one stock whose price is stable for a period, but suddenly turns to fluctuate in another period, The *n* we set based on the above data analysis will be inaccurate. This remains an unsolved weakness in our design. For the future improvement plan, this inaccuracy is assumed to be avoided if we could intelligently and intermittently analyse the volatility of each stock over time and set different *n* to a different period.

3. Maybe more factors?

Although this strategy is named a multi-factor strategy, only one factor, which is the modified version of Alpha006, has been used. We have tried to modify and add other factors in the Alpha101 series, but have met problems. Some factors require additional stock indicators that we couldn't get, some require functions that we cannot express correctly in R code, and some are performing an even worse PD-ratio outcome when we directly add to the strategy. In the future plan, we will try to examine and modify more factors in the strategy to make it a literally "real" multi-factor strategy.

Section 4 The Combined Strategy

Section 4.1 Combination methods

We have developed 3 trading strategies, each of which can be traded individually or in combination with the other. Sometimes the combined use of strategies can help traders to better recognise the market. Firstly, we consider this combined strategy divided into two components, one combined part is a long-term momentum strategy(double moving average with MACD indicators and RSRS timing strategy) and the other is an alpha multi-factor strategy.

In the performance of strategy 1 and strategy 2, we find that the two strategies show a weak correlation, which rarely shows the same profit or loss trend situation. Strategy 1 had better returns in trending markets, while Strategy 2 performed reasonably well in oscillating markets. Therefore, we considered blending the indicators of these two strategies, i.e. the RSRS is used as an auxiliary moving average crossing indicator.

For example, when the short SMA crosses the long SMA and the macd$ crosses the macd$ signal from bottom to top, this is clearly a bullish signal as it implies that the short-term price is higher than the long-term price; also when the RSRS is greater than 1, this means that the bullish crossover signal is confirmed. Therefore, the role of the RSRS here is to filter out some false signals. If the moving average enters a bullish signal but the RSRS is below 1, then it is not appropriate for the trader to open a long position.

Alpha strategies are based on a series of alpha factors, each of which determines a position, and the results of multiple factors are superimposed to determine the final position and size. The source of the alpha may be from industry, style, or others. The Alpha strategy focuses on the selection of stocks and is an active investment; in contrast, the Momentum strategy relies on the selection of the timing of investment and is a passive investment. Therefore, the third strategy is very flexible and we will run it separately.

In summary, the long-term strategy and the multi-factor strategy will work on ten stocks at the same time and will not affect each other, we combine the two positions together from the strategies and get a whole position for this series.

For the allocation of funds, we plan to use an even distribution of the two strategies initially.

Section 4.2 Risk management

We plan to adopt some risk management approaches to each of our three strategies. Our approach contains three parts.

The first part is to analyse the target stock's characteristics on the in-sample testing before conducting the strategies for in-sample trading. Strategies should only be considered to be applied if certain stock features are satisfied. For example, the long-term momentum strategy requires a strong trend from the stock market, we first conduct trend detection, i.e., the Mann-Kendall test on the out-sample testing. If the test does not detect a strong trend in stock, we stop using the long-term momentum strategy to avoid bad performance. Similarly, we abandon the multiple-factor strategy if we cannot find an effective correlation between price and volume.

The second part is to analyse the profit outcome of the in-sample testing. If a strategy performs a negative PD-ratio outcome on a specific stock, we set a lower position size against that stock. Otherwise, we set a higher position size for that stock if it tests well. If the PD ratio in the in-sample testing is below -50, we consider disabling that strategy to the specific stock. In addition, for strategies that have the lowest overall profit in the in-sample testing, we allocate no more than 20% of the total fund to it in the strategy combination session.

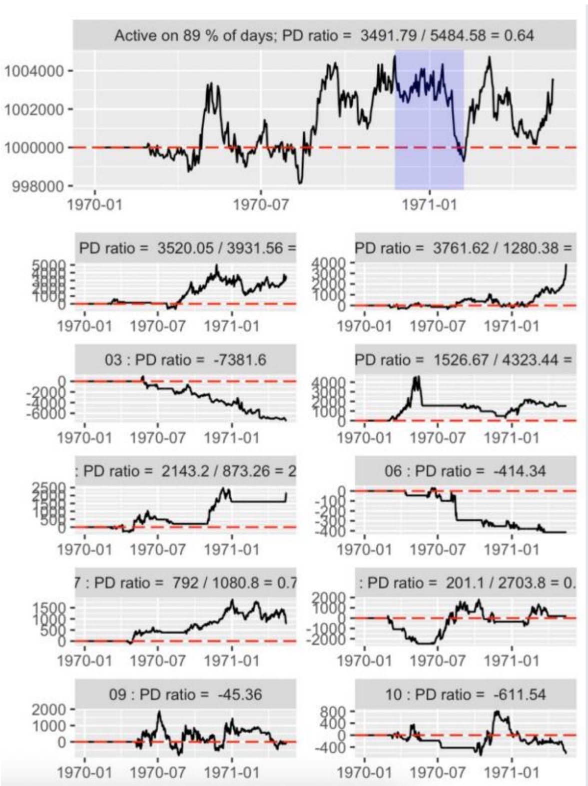
The third part is to supervise the performance during the out-sample trading. We calculate and rank the profits for every strategy on seasonal period bases. For the strategy whose profit ranks 1st in that season, we allocate 60% positions for it in the next seasonal trading period. For the 2nd one, we allocate 40% positions. If a strategy performs negative profit for continuing ten seasonal periods and does not have a warm-up trend, we disable that strategy on the stock. If all three strategies perform negative profit for continuing ten seasonal periods, we stop trading on that stock to stop loss and begin trading on the next stock. What's more, if the overall PD ratio remains negative and keeps decreasing for ten seasonal periods, we also stop trading on that stock to stop loss.

Section 4.3 Position sizing

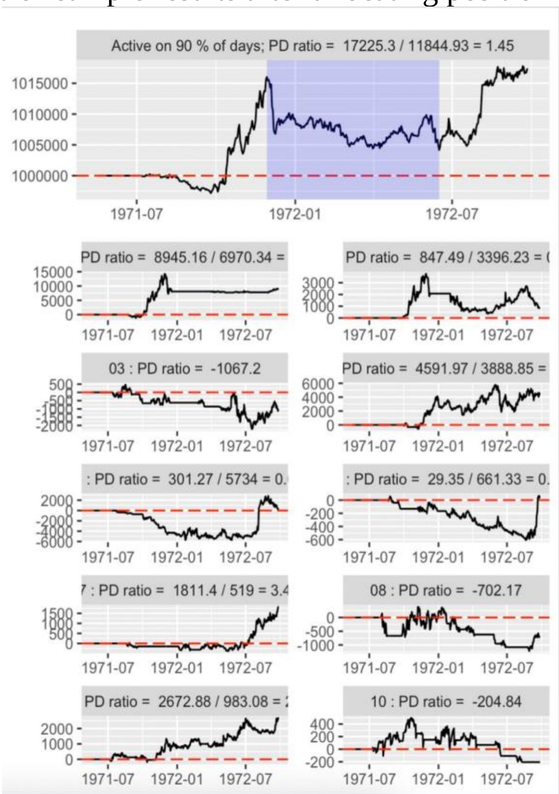
1. Plans and ideas for position sizing for a single strategy

For the first strategy of momentum strategy with double moving average and MACD indicators, I believe that if this strategy is applied to all series, different series will have different price units and will contribute differently to the final total profit. The approach to position sizing is to use the difference in opening price as a multiple of the positions in the different series according to the study of COMP396. In addition, this approach to position sizing is also not applicable to oscillating markets and therefore requires that positions in non-trending markets are sized to maximise profits, taking into account volatility and risk factors.

The charts demonstrate the in-sample results and out-of-sample results after allocating position sizes.



in-sample test



out-of-sample test

The Resistance Support Relative Strength strategy (short for RSRS strategy), which we have already mentioned in the previous section, is able to detect market volatility better and therefore allows for more accurate trading in volatile markets, complementing the first strategy. Therefore, for this strategy, we have considered the possibility of adding the analysis of volatility indicators to dynamically adjust positions. For example, the ATR function in the TTR package can be used to calculate the volatility indicator to obtain a coefficient of judgment. If the higher the volatility, the higher the real risk will be, then the position can be reduced, and vice versa.

For the multi-factor strategy, its nature is very flexible and can adjust the parameters according to the performance of the first two strategies, and the logic of this strategy is to go for stocks with more plus volume divergence. The initial position sizing for this strategy is managed by first testing the correlation between this price and the quantity purchased, and for higher correlations then go for smaller position sizes, and for lower correlations, add a little more to that position. Thus, the strategy is dynamic and flexible in the way it is position sizing.

2. For position sizing discussions across strategies

According to the combined strategies we mentioned earlier, firstly, the combined stock selection regarding the single momentum strategy and the RSRS strategy (we call this combination is long-term momentum strategy) can assist investors in making more accurate buy and sell trades. We consider giving the same position size to both strategies in order to maximise profitability when making accurate judgments. That is, the overall position is designed by analysing the market and trend of a clear series. For example, after the RSRS strategy has completed its position analysis for the volatility indicator, both strategies use price normalisation based on open prices differ for each series. In addition to this, we also need to analyse the spread between the buy and sell points. If the spread is large, this indicates that there is significant scope for arbitrage, and this also requires flexibility in adjusting positions.

In addition, according to the capital allocation of our three strategies, through the study of COMP396 Lec4, we also need to get the capital allocation to maximise the use of capital, so we also consider scaling up our position sizes according to the allocated capital.

Section 4.4 Future plan

In the plan, apart from considering the above-mentioned position sizing methods, we also consider different risk management indicators to develop the position sizing based on the requirements of part2 and data analysis. For example, as mentioned above: stock volatility, price variance, etc. In general, we consider a method called Risk Parity (Chen, 2021), which features a balanced weighting of each strategy based on risk, rather than a simple equal weighting of investment shares. This would allow strategies that perform well in a variety of equity trading environments to make a similar contribution to equity risk. This results in more robust equity returns and the pursuit of better yields.

In contrast, if an allocation of strategies is based solely on equal weights without attention to the risk profile of equity, it would result in a larger loss of equity returns if a greater risk were to arise in a given period. In other words, the Risk Parity strategy allows for more robust investment returns across different economic cycles and business environments, resulting in asset appreciation. It allows stocks to gain market upside in periods of rising markets, control strategy risk, and reduce losses from the strategy in periods of market imbalance. It is therefore a more rational way of spreading the risk.

In addition, we need to use cross-validation (*Cross-validation: evaluating estimator, 2007*) to find an optimal portfolio of strategy allocations and avoid overfitting. This means splitting the entire data into a training set and a test set, and applying the results of the training set to the test set for validation. So we decided to divide the part1 data into four groups, using one of them as the validation set and the rest as the training set for cross-validation. The results were then averaged to get the best allocation and applied to the part2 data for validation. This will make the final results more accurate and valid.

Section 5 Breakdown of the teamwork

Zixuan Zhu

Background research

-Looking through the suggested academic sources and quantitative strategy websites to choose a strategy that would suit our data, the first one chosen was the Triple Moving Average strategy.
-After reading more information and more detailed analysis, I found that taking macd or double averages alone is a loss, so TMA strategy was not a good choice and finally I decided to develop a momentum strategy with double moving average and macd indicators (MACD).

Coding

-In order to show the suitability of momentum strategy to the given stock market， I used python and Rstudio to do volatility analysis, k-chart analysis and trend detection.
-I am the first one who has completed the code part in our team. I not only use the TTR package in Rstudio to call SMA function and MACD funtion to show the strategy, but also complete the preliminary testing and backtesting part.

Strategy ideas

-I have come up with two strategies: TMA strategy and a momentum strategy with double moving average and macd indicators (MACD) with more consideration, I chose the second one.

Design presentation and report

-Presentation:I will give a presentation about a momentum strategy with double moving average and macd indicators (MACD) and answer questions about it.
-Report: All parts about a momentum strategy with double moving average and macd indicators (MACD) and position sizing ideas about the combined strategy.

Yiwang Tian
Background research
-I collected and read reports on quantitative strategies on the Joinquant website and research reports on quantitative strategies from securities firms to know about how Resistance and support work.
- By gathering information, I also thought about how to combine strategies. If we want to reduce risk and make money, it is better to choose individual strategies with a negative correlation between the effects of the strategy returns, i.e. when one strategy fails it is likely that the other strategy will work, and presented it in a weekly group meeting.

Coding
-I finished this part strategy code and wrote the code of the testing part. I used the calculation of N to calculate the slope for the sample test and found the suitable value of S1 and S2. After that, I completed the coding part of backtesting.
- To implement my code, I also learned statistics packages in R.

Strategy ideas
-I have come up with the strategy "Resistance support relative strength(RSRS) strategy" based on the limitations of bollinger bands.
- I ran through and analysed how our three strategies works as well as features of each, and proposed the combined idea based on extensive research.

Design presentation and report
-Presentation: I will explain and answer questions about RSRS strategy.
-Report: I am responsible for all contents related to RSRS strategy and the part of how our combined strategy works.

Weiyi Liu

Background research

-I read the information that professors suggested and searched a lot of strategies that we may develop and introduced their principles to my team members, including the Momentum strategy, Multi-factor strategy and Alpha hedging strategy.

-I discussed with my group members the ideal development of our strategy and presented the ideas to the group in a weekly group meeting.

Coding

-I discussed one example of strategy - bbands_trend_following with my team members and I took it as an example to show my team members how the main file is linked to the strategies file.

Strategy ideas

-Due to we can choose a number of strategies, I analysed how each strategy works, its strengths and weaknesses, and above all, its feasibility and future development. And help other team members to confirm which strategy to develop.

-I proposed the test and stop loss for every strategy, and decide on the logic to finish it.

Design presentation and report

-Presentation: I will introduce the whole content of the report and answer the questions about the ideas of our team.

-Report: I am mainly responsible for the whole part, I collected every idea and completed typesetting and editing. Also, I finished the breakdown of the teamwork part by myself.

Tianyi Ye

Background research

-Based on the information Weiyi Liu had gathered and the advice given on strategies, I selected and found many parameters that required by the strategies that were not given.
-After confirming the strategies I wanted to develop with Zhiyu Chen, I found some problems that needed to be solved, such as we could not get the volume value with the existing parameters, we could not get the whole period earnings value inside GetOrders so I was responsible for seeking advice from my supervisor and lecturer.

Coding

-I worked with Zhiyu Chen on the code for this strategy - Multi-factor strategy α101 in Multi-factor strategy.
-During the writing of the strategy, I would check each method that we were going to implement in a new page to ensure that the method would live up to the intended requirement.
-After we have written the code, I start to compare the returns generated by setting different parameters, e.g. threshold, and size of in-sample. Finally, I compared the return curves generated under each parameter and selected an optimal combination of parameters.

Strategy ideas

-I summarised some of the potential problems with the multi-factor strategy and thought about what changes and upgrades the strategy might need to make in the future, e.g. we only designed one factor at the moment and we would need to consider multiple factors together at a later stage.

Design presentation and report

-Presentation: I will explain and answer questions about alpha101 for multi-factor strategy.
-Report: I am responsible for all contents related to alpha101 for multi-factor strategy.

Zhiyu Chen

Background research

-To support the other team members of the group in the subsequent development of the strategy, I have gathered a lot of information on how to analyse stocks and what specifically to analyse about stocks.
-I researched through reading academic material to learn about how alpha101 for multi-factor strategy among multi-factor strategies work. I choose

Coding

-I used RStudio and PyCharm to do basic data analysis and find the characteristics of each stock separately. The data analysis part includes concrete feature analysis of particular stocks and comparisons between different stocks. Besides, my code shows the results of stock data analysis, with seasonal plots, violin plots, and median line charts.
-I used R code to do stock analysis and finished the data analysis part of the multi-factor strategy.
-I finished the code and comment writing. In addition, I solved bugs and problems (e.g., the infinite recursion problem when we first tried to loop to find the optimal threshold), and iterated the code (e.g., adding the observed day parameter).

Strategy ideas

-I picked alpha006 as the factor to conduct in the code among the 101 listed factors. I optimised the logic of the alpha 006 equation and created the parameter - the observed day “obday”. Original alpha006 can only make decisions based on the most recent 10 days’ data, while our updated strategy allows us to set the number of most recent days that the decision is based on.

Design presentation and report

-Presentation: I will explain and answer questions about alpha101 for multi-factor strategy.
-Report: I am responsible for all contents related to alpha101 for multi-factor strategy.

Section6:Reference list

Chen, J. (2021) *What Is the Volume Price Trend (VPT) Trading Indicator* Available at: https://www.investopedia.com/terms/v/vptindicator.asp (Accessed: 15 November 2022).

Chen, J. (2021) *Risk Parity: Definition, Strategies, Example* Available at: https://www.investopedia.com/terms/r/risk-parity.asp (Accessed: 15 November 2022).

Cross-validation: evaluating estimator (2007) Available at: https://scikit-learn.org/stable/modules/cross_validation.html (Accessed: 15 November 2022).

Dolan, B. (2022). *MACD Indicator Explained, with Formula, Examples, and Limitations.* Available at: https://www.investopedia.com/terms/m/macd.asp (Accessed: 1st Nov 2022)

La Ferla, Michele. (2019). 'Generating Trading Signals using Dual Moving Average Crossover Strategy'. [e-journal] available at: https://www.researchgate.net/publication/334194514_Generating_Trading_Signals_using_Dual_Moving_Average_Crossover_Strategy (Accessed: 1st Nov 2022)