ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

TIK Institut für
Technische Informatik und
Kommunikationsnetze

Make Us Rich!

Semester Thesis
August 29, 2012

Thomas Bürli
tbuerli@ethz.ch

Supervisor:
Philipp Brandes
philipp.brandes@tik.ethz.ch

Professor:
Prof. Dr. Roger Wattenhofer
wattenhofer@tik.ee.ethz.ch

Distributed Computing Group –   DCG
Computer Engineering and Networks Laboratory –   TIK
Department of Information Technology and Electrical Engineering –   ITET
Swiss Federal Institute of Technology –   ETH

Abstract

In the area of algorithmic trading, the main academic and industrial focus lies within the field of high frequency trading (HFT). However, the flash crash in 2006 and Knight Capital Group disaster led to political debates about a financial transaction tax which would destroy the HFT business model. Therefore, normal algorithmic trading (systematic trading) approaches are gaining more attention and acceptance.

This thesis proposes a modular framework for defining and evaluating trading strategies based on historical stock data of the past decade. As a proof of concept, a handful simple, but efficient strategies are defined and evaluated. The best algorithms are successfully applied to new data with a maximal return of up to 11% within half a year. This indicates the benefit of a systematic trading algorithm which is easily realizable with this proposed framework.

Thomas Bürli

Contents

List of Figures

List of Tables

Thomas Bürli

Introduction

The stock market is one of the most important sources for companies to obtain capital resources. In 2010, the global stock market capitalization is estimated to be 54 trillion dollar [Roxburgh et al., 2011]. Hence, both the complexity of the market and the possibilities to make gains have grown over the last decade. Therefore, algorithmic trading has seen enormous growth and has increased the capabilities of cost-effective automated trading. An important subclass of automated trading are high frequency trading (HFT) algorithms. They are involved in more than 70% of all U.S. cash equity trades [Easley et al., 2012]. These algorithms usually hold stocks or certificates only for a brief time, sometimes only for a few milliseconds and gain money by making thousands or tens of thousands trades a day.

1.1. Motivation

These HFT algorithms gained public attention since the flash crash in May 6, 2010 [Lauricella and STRASBURG, 2010] when the Dow Jones nearly dropped 1000 points in a matter of minutes. According to regulators, HFT increases the volatility and therefore caused the flash crash. In the beginning of August 2012, another disaster happened when the Knight Capital Group lost about 400 million dollars in 45 minutes due to a faulty HFT algorithm [Mehta, 2012]. Recently, politics tries to burden HTF with a financial transaction tax [BOWLEY, 2011] which would destroy their business model. For this reason, a return to normal algorithmic trading, also called systematic trading, is required. Systematic trading is a method of creating rules that can execute a trade idea and manage risk with some degree of consistency and predictability over time [Today, 2010]. This thesis aims to analyse the approach of exploiting statistical properties of assets to predict general trends and/or foresee the future value of certain stocks and use those predictions to beat the market.

1.2. Related Work

A lot of research has been done in the field of trading and analysing the stock market. Alexander [2001] explains the market, its mechanism and the model used by investment risk managers and investment analysts. Vince [1992] writes about the mathematics behind systematic trading which includes money management techniques, risk analysis and modern portfolio theory. A good overview of the topic of systematic trading and a insight into the state of the art techniques currently used by the financial traders can be found in Acar and Satchell [2002]. Furthermore, they also show how computer models can be used to gain money on the market and minimize the risk. This covers technical indicators, neural networks and genetic algorithms which are currently used in the area of finance.

Taylor and Allen [1992], TSAY [2010] andAchelis [2000] provides a basic overview of technical analysis. Some research about the algorithmic aspect of systematic trading has been cone by Kannan et al. [2010] and Boyarshinov and Magdon-Ismail [2005]. Kannan et al. [2010] used a combination of different algorithms to get a better prediction of the next day's closing price.

1.3. Contributions

The main contribution of this semester thesis is to design a methodical framework which makes it possible to define strategies, train with historical data and evaluate their performance. The handling is simple and adaptable. In addition, some simple strategies are defined, implemented, and evaluated.

Background

This chapter provides some basic information on stock market to understand this thesis.

2.1. Stock Market

The stock market is part of the capital market. It is a marketplace for trading with company stocks. One must differentiate between only privately traded and exchange traded equity. Stock exchanges such as SIX[1], NYSE[2] or NASDAQ[3] provide service to brokers and traders. Each company has to be listed on a stock exchange before their stock can be traded on this specific exchange. Companies can be listed on more than one stock exchange at the same time. Stocks which are not listed in any stock exchange are traded on secondary markets (private trade equity). These trades are also called Over-The-Counter (OTC) traded share.

2.1.1. Trading

During trading hours, the stock market gives traders the opportunity to trade stocks and connects sellers and buyers. A buyer or seller can put orders to the exchange at which rate he is willing to buy or sell. A broker collects these orders in a so called order book and tries to find a price at which the most trades can be achieved.

---

[1] Swiss Exchange
[2] New York Stock Exchange
[3] An American stock exchange

2.1.2. Information

After close of trading, the stock exchanges provide information about stock prices and traded volume. This includes the following:

Close Price of the stock in the evening after the trading phase.

Open Price of the stock in the morning before the trading phase. This may not be the same as the close price the day before.

Min Minimum price of the stock that day.

Max Maximum price that day.

Volume The total quantity of stock traded during a day.

This information is included in the historical stock data.

2.2. Analysis

There are two different techniques to analyse a stock. The fundamental analysis and the technical analysis.

2.2.1. Fundamental Analysis

The fundamental analysis tries to get a fair price for a stock by looking at a company's financial report, its health, other competitors and the market. The goal of a fundamental analysis is to get a value of the business/stock and find mispricings.

2.2.2. Technical Analysis

The technical analysis only depends on historical stock data. It tries to predict future stock courses using this data. In the historical stock data, price and volume is included. Therefore, the main goal is to find the right time to sell or buy stocks for the return to be maximal. Often technical indicators are used to find trends and signs of momentum.

Implementation

The goal of this thesis is to provide a methodical framework to define and analyse algorithmic trading strategies. For this purpose, a framework is developed which can handle a large data set and allows easy testing and analysing of trading strategies. To evaluate the framework, some simple strategies have to be found and tested. In this chapter, the design and implementation of the framework is described.

3.1. Framework

To have a flexible system, the framework is constructed modularly such that each part can easily be replaced by a new implementation. As shown in Figure 3.1, the framework consists of three main parts: the data handler, the strategy, and the simulator. The framework is implemented in Python 2.6.[1] It depends on the Python library, the Numpy[2] library and the Pandas[3] library. Numpy adds support for large, multi-dimensional arrays and matrices and it has a wide range of mathematical functions. Pandas builds on Numpy and provides high-performance, easy-to-use data structures and data analysis tools. The implemented classes can be found in the appendix (Section A.5).

3.2. Data Handler

The data handler is responsible for providing stock data. All access to data is managed by the handler. Furthermore, it enables to abstract from the underlying data format and thus

---

[1] http://www.python.org/
[2] http://numpy.scipy.org/
[3] http://pandas.pydata.org/
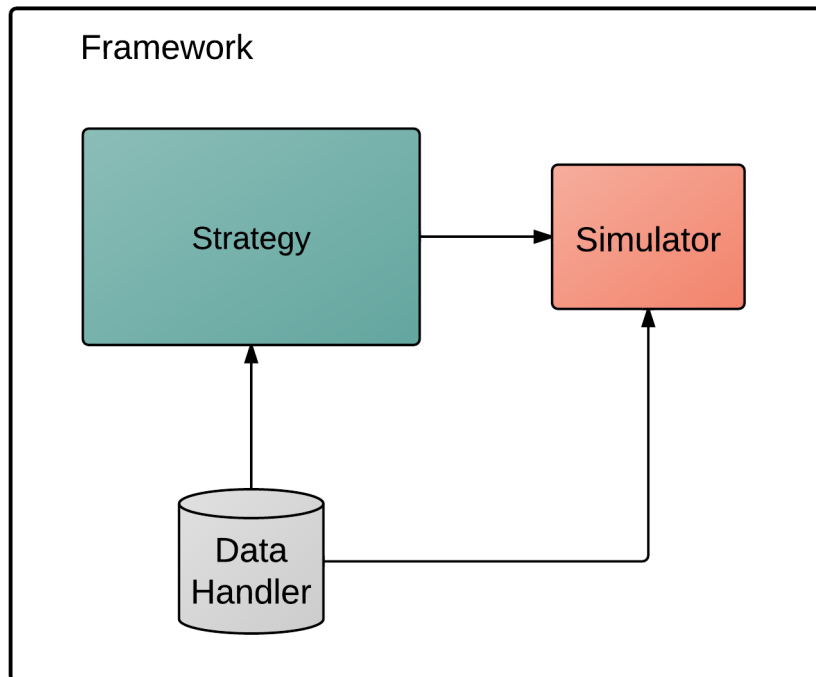
Figure 3.1.: The main components of the framework

makes it possible to combine data from different sources. The Dataset used in this thesis is explained in Section 3.2.1.

Generally, the data is loaded from CSV files. If the data for a stock does not exist locally, it can be loaded from an online data source. For further use inside the framework, the date has to be stored in a structure called data container. The data container defines the pool of stock from which the trading strategies can choose from.

3.2.1. Dataset

For this work, a dataset with historical stock data has been provided including Open, Close, Min and Max price and the daily volume for each stock. The data has a daily frequency and includes more then 30'000 stocks from NYSE[4], AMEX[5] and OTC derivatives markets. The datasets ranges from Mai 1, 1993 to June 1, 2012.

3.3. Strategy

In real life, a traders strategy relies on different metrics and heuristics to decide when to buy or sell stocks. This behaviour is modelled by a strategy within the framework. The strategy decides when to buy or sell depending on different underlying algorithms. In particular, there are the council, the agents and the indicators (Figure 3.2). The main algorithmic

---

[4]New York Stock Exchange
[5]American Stock Exchange
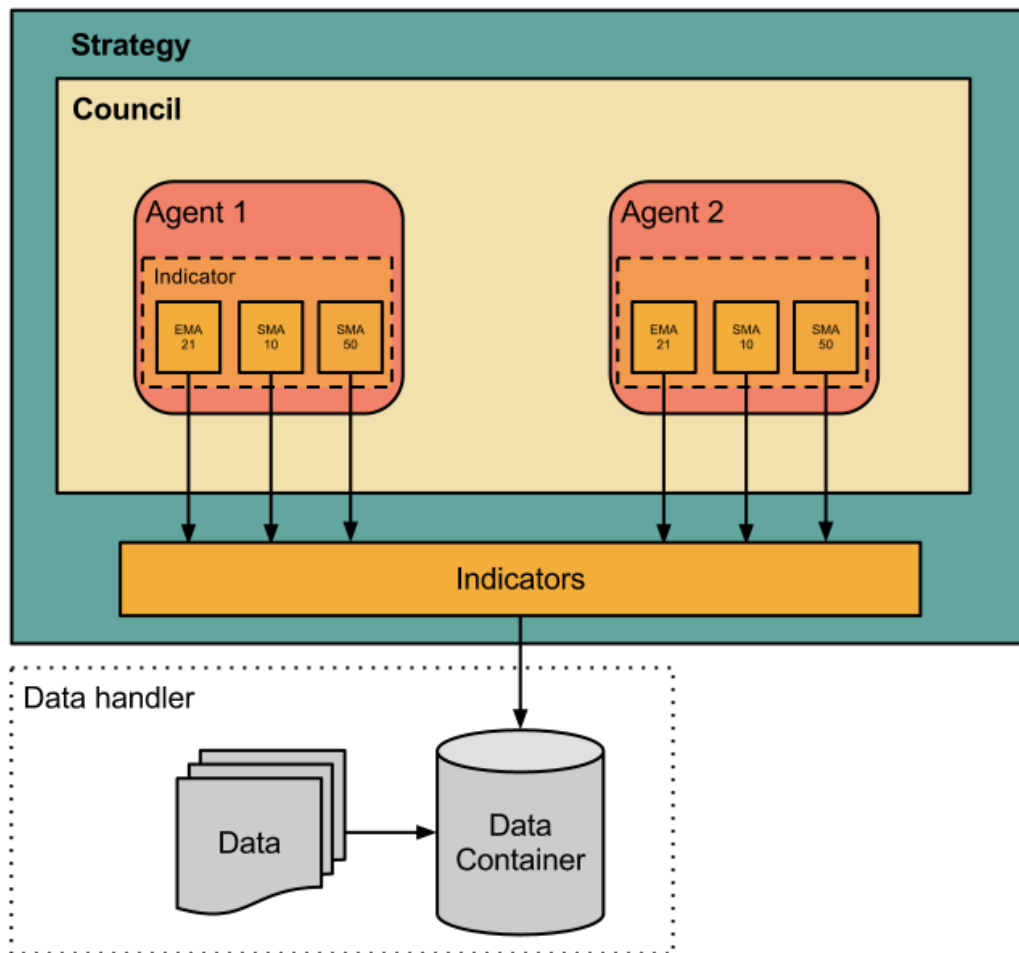
Figure 3.2.: Structure of a strategy

computation is implemented in the indicators. Each layer is a composition of the layer below. Therefore, the strategy functionality is to combine different agents to one big trading strategy.

3.3.1. Indicator

An indicator could be seen as basis of the strategy. It is the lowest element in the hierarchy and has direct access to the data container. The purpose of an indicator is to provide preprocessed data to the upper instance, the agent. Most indicators are just implementations of a technical indicator. A technical indicator is defined as a class of metrics whose value is derived from the historical stock price over a period of time. It tries to predict the future price or shows trends by looking at past patterns. Examples of common technical indicators are Simple Moving Average (SMA), Relative Strength Index, Money Flow Index, Stochastics, MACD and Bollinger Bands [Investopedia, 2012b]. In the following, some of the implemented technical indicators are described here:

OHLC[6] & Volume The OHLC & Volume are the daily data we get from the stock market.

Simple moving average (SMA) Average of the close price over the last n days.

Chaikin Money Flow Indicator Measures buying and selling pressure for a given period of time

Daily Return Calculates the return over the last n days.

The complete description of all indicators can be found in the documentation.

3.3.2. Agents

An agent is a simple implementation of an algorithm which combines different indicators. Therefore, the agents have access to indicators. The purpose of an agent is to advice the strategy regarding a single heuristic or metric. The advice includes a vote when to buy or sell a stock and how confident it is about its decision. A strategy makes its decision based on the advices it gets from all its agents, analogous to politics where a mayor has his advisers to assist the decision making process.

Different Agents are implemented in the framework and some of them are described in the following:

SMAAgent(x)

The SMAAgent uses the SMA indicator over x days (SMAx). It checks if the SMAx is over or under the close price. Therefore, the crossing point of the SMAx and the close is the buy/sell signal. If the close price crosses the SMAx from below to above it is a buy signal because it shows an upward trend. The opposite shows a downward trend and is a sell signal. The chosen length x of the SMA depends on the analytical objectives. 100 days or more show the long-term trend which is better suited for investment strategies. On the other hand, a short period of 20 days is best suited for short-term trends and trading strategies.

SMADoubleAgent(x,y)

The SMADoubleAgent uses two SMA indicators one over x days (SMAx) and one over y days (SMAy). The SMAx should have a smaller period than the SMAy and those $x < y$. The functionality is similar to the SMAAgent but it uses the second SMA and not the close price for the crossover comparison. Therefore, the agent promotes a buy if the SMAx value exceeds the SMAy.

BestAgent(x,y)

The BestAgent uses a daily return over y days to get the x best and x worst stocks each day. For the x best stocks it recommends a buy and for the x worst a sell.

3.3.3. Council

To easily provide agents to the strategy a council is introduced. The council takes an order from the strategy which includes a list of agents the strategy it wants to use. The council will then execute this order and construct all those agents. Furthermore, the council provides an access to the agents for the strategy.

3.3.4. Example Strategies

Some simple strategies are implemented which are based on the agents and indicators described previously. In the following, these simple strategies are briefly discussed.

SingleSMA(x)

The 'Single Simple Moving Average' strategy only uses the SMAAgent(x). It checks if the SMAx goes above or below the close price which gives a buy or sell signal. This strategy represents a common short-term/long-term heuristic which is often used.

Double SMA(x,y)

This strategy is called the 'Double Simple Moving Average' and uses the DoubleSMAAgent(x,y). It is like the SingleSMA strategy but it uses another SMAy instead of the close price.

BuyBest(x,y)

The BuyBest strategy tries to buy the x best stocks according to their return over the last y days. Furthermore, it sells each of the worst x stocks if they are in the portfolio. For this purpose, the BestAgent(x,y) is used. This strategy was presented in the Economist [Economist, 2011].

Half

The Half strategy is a example for a combination of different agents. The strategy checks if the majority of the agents vote for a buy or a sell. The council makes it easy to change the agents used.

3.4. Graphical User Interface

The analysis can be run in the console. However, to increase usability, a simple GUI exists. The functions of the GUI is limited. It lets the user choose stocks, strategy and the time range for simulation. The GUI can be seen in Figure 3.3. It was built with PyGTK[7] and Glade[8]. For fast prototyping Quickly[9], a template generator from Ubuntu, was used.

---

[7]http://www.pygtk.org/
[8]http://glade.gnome.org/
[9]https://wiki.ubuntu.com/Quickly
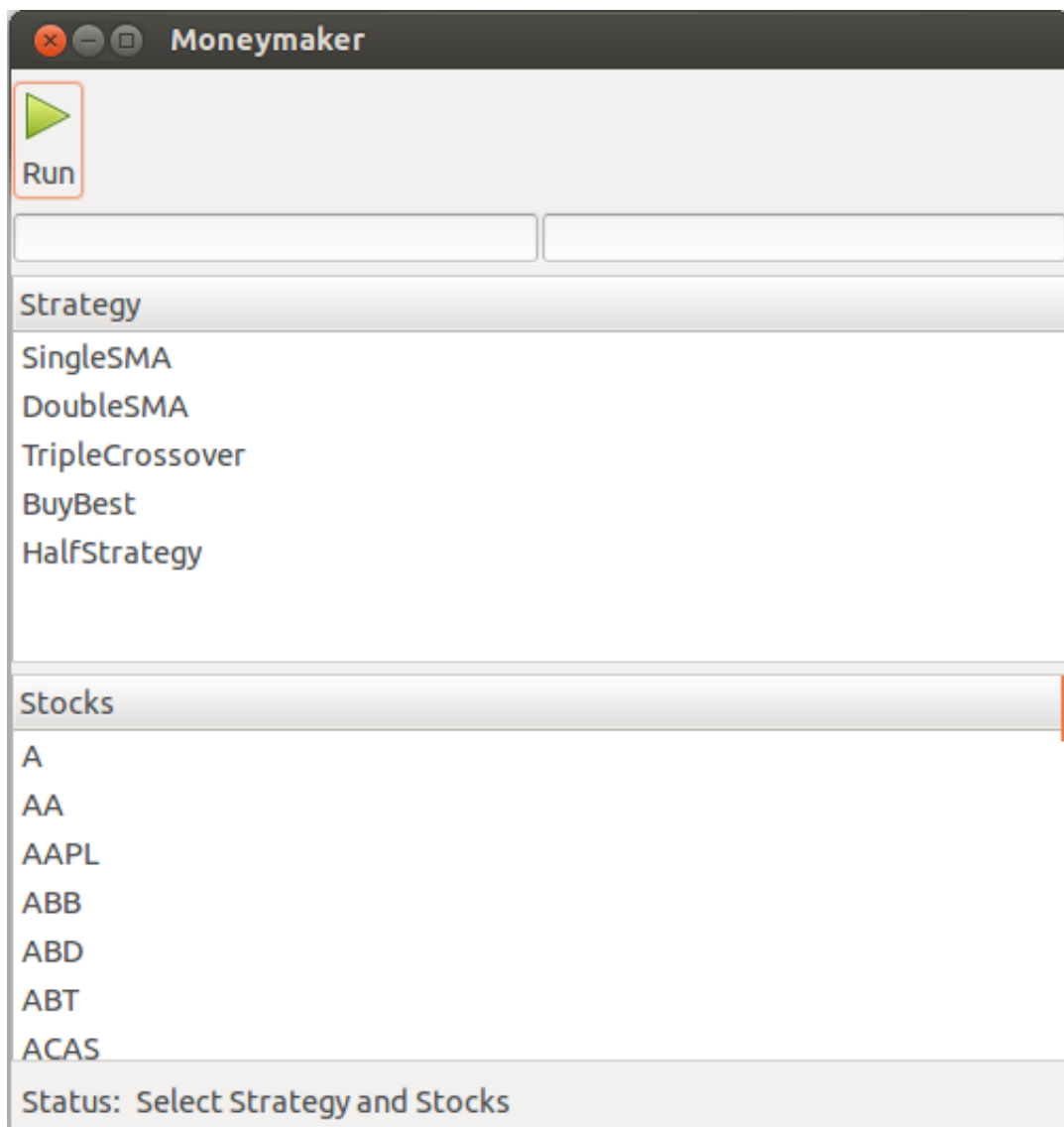
Figure 3.3.: The GUI of the framework

CHAPTER 4

Analysis and Evaluation

The testing of the strategies is done in two steps. First the strategies are simulated and analysed with different parameter on historical data. The best 12 strategies are chosen (4 from each strategy) and are evaluated with new data to find out how the strategies handle the new condition.

4.1. Assumptions

The following assumptions are made for the analysis and evaluation according to Boyarshinov and Magdon-Ismail [2005]:

A1 **Frictionless Exchange** No transaction cost for trades.

A2 **No Market Impact** Trades can be placed without affecting the quoted price.

A3 **Fractional Market** Arbitrary amounts of stock or bond can be bought or sold at any time.

A4 **Long Strategies** We assume that we can only hold long positions in stock or bond.

4.2. Testing Set

For testing the strategy, a set of stocks had to be chosen. Because the result strongly depends on the selection of this set, the stocks from the common Standard & Poor 500 (S&P 500) index were used. The S&P 500 is based on the common stock prices of the 500 top publicly traded American companies. The stocks are chosen according their to market size, liquidity, industry grouping and among other factors [Investopedia, 2012a]. Therefore, all stocks which have ever been in the S&P 500 are included in the testing set which gives
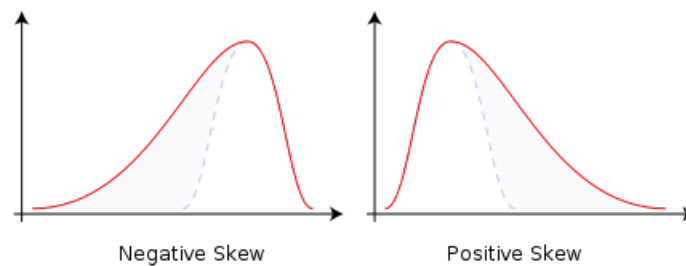
Figure 4.1.: Normal distribution with different skewness [Wikipedia, 2012]

a list of around 900 stocks. The complete list can be found in the appendix A.4.

4.3. Measures

The performance analysis can be categorized into three groups: Mean-variance measures, the profitability measures and the risk measures.

4.3.1. Mean-Variance Measures

The mean-variance measures are a statical analysis of the returns of assets. It is based on the assumption of the Markowitz's mean-variance model (MVM) which states that the returns follow a normal distribution [Markowitz, 1952]. In consequence, MVM can be used as measures for the performance of strategies. The following moments are used: mean, standard deviation, skewness and kurtosis.

Mean

The mean is the first raw moment. It shows whether the strategies has made profit or loss in the analysed period. Therefore, a high mean is preferred.

Standard deviation

The standard deviation, the square root of the variance, is the most common measure of volatility. It shows the uncertainty and the risk of the strategy.

Skewness

Skewness measures the symmetry, respectively the lack of symmetry of the normal distribution. As it can be seen in Figure 4.1, a strategy with a negative skewness has a high probability of achieving high profits but according to Holton [2003] it has also a higher risk. A positive skewness has a lower risk and also a lower probability of high profit.

Kurtosis

Kurtosis describes whether the data is peaked or flat relative to a normal distribution. Thus, it measures the "peakedness" of the strategy. A peak strategy have some extremely but unlikely returns. A standard normal distribution has a kurtosis around 3 [Brown, 2011].

Therefore, the higher the kurtosi's coefficient is relative the "normal level", the more likely returns will be either extremely large or extremely small.

4.3.2. Profitability Measures

The total return, the logarithmic monthly return and the excess return is used for profitability measures. The logarithmic return is calculated with

$$return_{log} = \ln(return)$$

. The excess return is calculated from the monthly return of the asset and the reference Index (S&P 500).

$$excess\_return = asset\_return - index\_return$$

4.3.3. Risk Measures

As a measurement of the risk, the volatility is used which measures dispersion of returns. A high volatility means that the price of the asset can change dramatically over a short time and therefore the risk is high. The volatility is calculated using a monthly and yearly frequency.

4.4. Analysis

Some of the strategies described in Section 3.3 are analysed over a period of twelve years starting on January 1, 2000 and ending on December 31, 2011. Different parameter are set and the outcome is analysed with previously described the measures. The complete results can be found in the appendix A.3.

4.4.1. SingleSMA(x)

The SingleSMA(x) strategy is analysed with different sizes x. A look to 'SingleSMA 84' showed a relative high kurtosis of 892 which means it has small number of extreme returns. Figure 4.2 shows the asset of 'SingleSMA 84' and S&P 500 index. A strange behaviour can be found around September 2010 which is responsible for the high kurtosis. The analysis of the underling data revealed that the stock of BNS Holding Inc. (BNSSA) is responsible for this outbreak. According to PRNewswire [2010] they did a reserve stock split which means that the number of stocks decrease and the price of the stocks rise. Therefore, historical data of BNSSA after the stock split (September 3, 2010) were removed and analysis was repeated with the adjusted dataset.

The detailed results of the analysis can be found in the Table A.1 in the appendix. A good strategy should make money. Therefore, those with a negative return can be neglected. Furthermore, a good strategy should also have a low risk. For this reason, the Risk-Return-
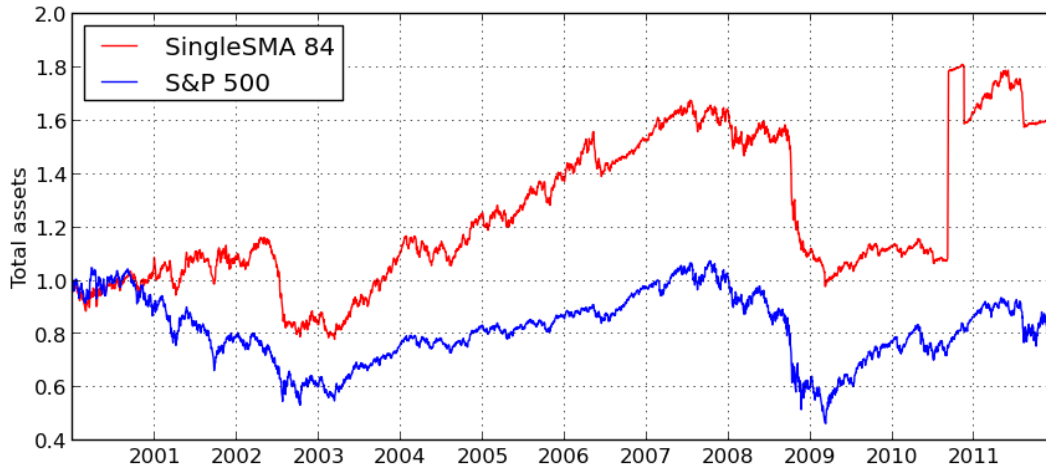
Figure 4.2.: SingleSMA 84: total assets

name	return (%)	$\bar{x}\,(10^{-5})$	$\sigma\,(10^{-3})$	skew.	kur.	vol. yearly	RRR
S&P 500	-13.58	-4.836	13.865	-0.153	7.051	0.210	-0.309
SingleSMA 60	1.68	0.552	9.2	-0.652	4.987	0.22	0.036
SingleSMA 64	8.64	2.746	11.698	0.882	70.529	0.262	0.157
SingleSMA 72	4.66	1.509	8.283	-0.44	3.997	0.213	0.104
SingleSMA 76	36.03	10.195	8.784	-0.601	5.284	0.182	0.94
SingleSMA 78	23.37	6.959	8.438	-0.599	4.736	0.198	0.563
SingleSMA 80	26.52	7.794	10.385	0.308	59.554	0.198	0.639
SingleSMA 84	10.06	3.176	7.921	-0.512	4.538	0.172	0.279
SingleSMA 100	8.49	2.69	8.99	-0.515	6.556	0.214	0.189

Table 4.1.: The best SingleSMA strategies

Ratio (RRR) is the risk relative to the return and is defined as follows:

$$RRR = \frac{1 - \dfrac{return_{max} - return}{return_{max}}}{1 - \dfrac{risk_{min} - risk}{risk_{min}}}$$

. The yearly volatility is chosen for the risk. The optimal strategy would have a RRR of 1 which means it has the maximal return and also the minimal risk of all strategies. In addition, the skewness and kurtosis are considered as well. 'SingleSMA 64' and 'SingleSMA 80' have a high kurtosis in comparison to the others and thus, they have a small numbers of extreme returns. For evaluation, the strategy with the parameters 76,78,80 and 84 are chosen to represent the best SingleSMA strategies.

4.4.2. DoubleSMA(x,y)

The DoubleSMA(x,y) strategy has two parameters one for each SMA. The parameters for the analysis are chosen based on experiences. All result can be found in the Table A.2 in

the appendix. Most DoubleSMA strategies outperform the S&P 500 index in view of return. The best eight strategies according to their Risk-Return-Ratio, are listed in Table 4.2.

name	return (%)	$\bar{x}\ (10^{-5})$	$\sigma\ (10^{-3})$	skew.	kur.	vol. yearly	RRR
S&P 500	-13.58	-4.836	13.865	-0.153	7.051	0.210	-0.02
DoubleSMA 5/100	271.25	43.463	14.915	0.091	8.98	0.194	0.427
DoubleSMA 10/100	251.02	41.606	13.374	-0.433	6.639	0.243	0.316
DoubleSMA 20/100	133.03	28.032	10.779	0.033	11.076	0.145	0.28
DoubleSMA 21/100	225.93	39.149	12.66	-0.824	12.765	0.218	0.317
DoubleSMA 21/200	294.4	45.467	18.349	-0.469	10.187	0.236	0.381
DoubleSMA 25/100	309.38	46.702	13.321	0.439	9.266	0.24	0.394
DoubleSMA 20/250	396.16	53.073	18.531	1.029	16.508	0.334	0.362
DoubleSMA 25/300	255.31	42.008	15.14	-0.291	5.798	0.154	0.506

Table 4.2.: The best DoubleSMA strategies

In Figure 4.3 the assets of the 'DoubleSMA 20/250' strategy and the normed S&P 500 index are shown. In 2008, the value of asset and the index dropped maximally. This can be explained by the finance crisis. For a better comparison of the strategy with the S&P 500 index, the monthly return and excess return are considered and can been seen in Figure 4.4.
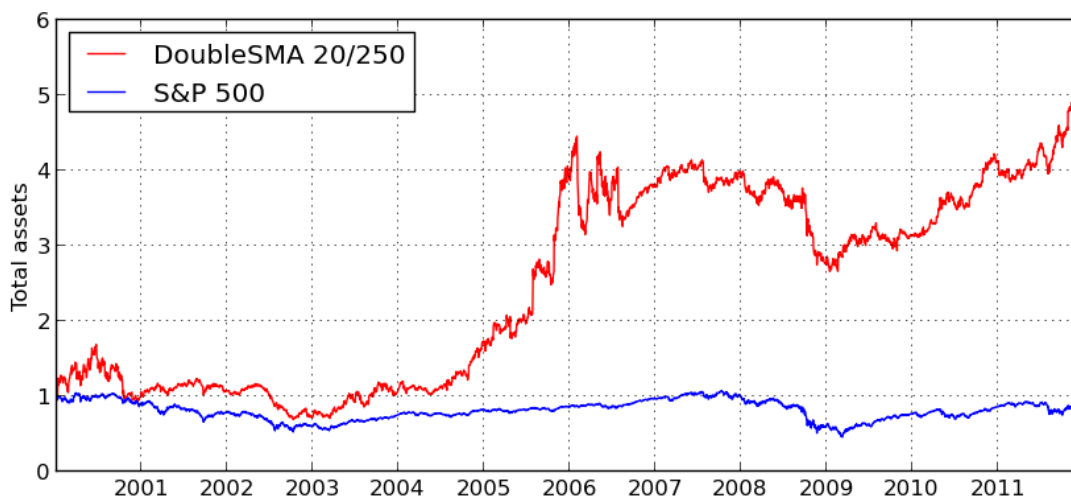
Figure 4.3.: Asset from DoubleSMA 20/250

The following DoubleSMA strategies are selected for further evaluation: 'DoubleSMA 25/300', 'DoubleSMA 5/100', 'DoubleSMA 25/100' and 'DoubleSMA 21/200'.

4.4.3. BuyBest

The BuyBest strategies has also two parameters. The first denotes the number of shares and the second represent the number of days. Table 4.3 shows the eight best BuyBest
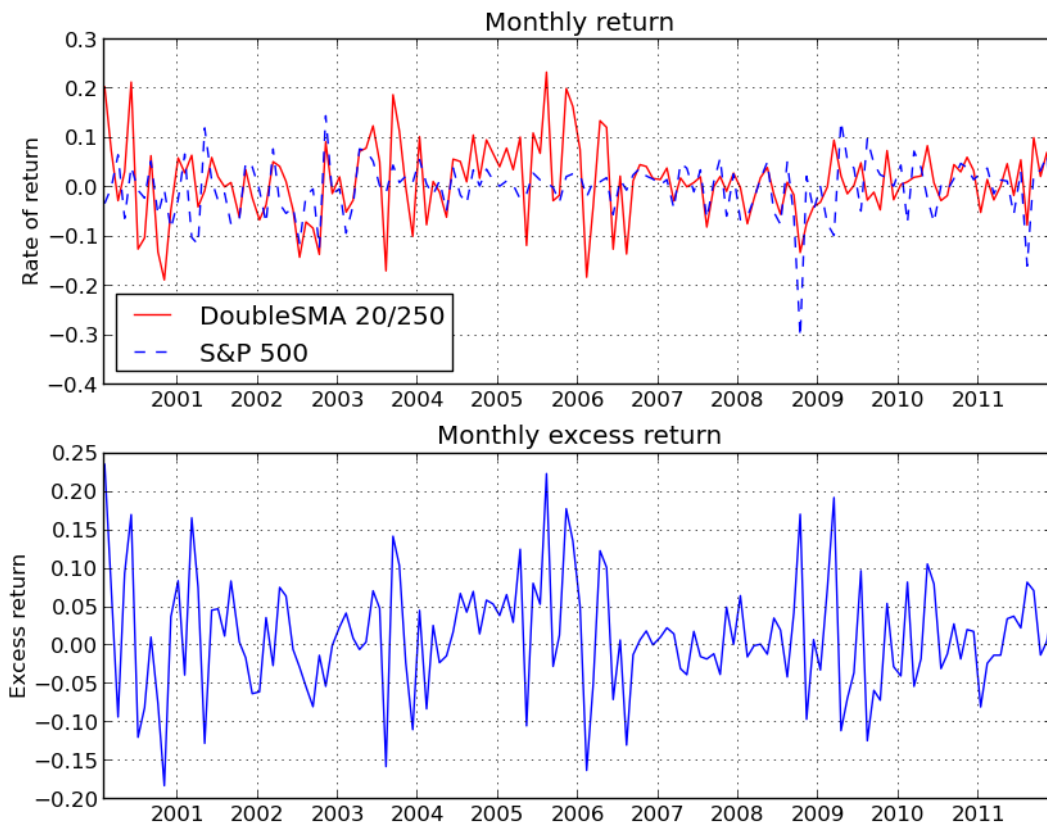
Figure 4.4.: Monthly log and excess return from DoubleSMA 20/250

strategies. The complete table can be found in the appendix (Table A.3). In comparison to the other strategies, the BuyBest outperforms the SingleSMA(x) in terms of return and risk, but it is on the same performance level as DoubleSMA(x,y) strategies. The best four BuyBest strategies are 'BuyBest 5/90', 'BuyBest 20/30', 'BuyBest 25/30' and 'BuyBest 30/30'.

name	return (%)	$\bar{x}\,(10^{-5})$	$\sigma\,(10^{-3})$	skew.	kur.	vol. yearly	RRR
S&P 500	-13.58	-4.836	13.865	-0.153	7.051	0.210	-0.030
BuyBest 5/60	286.92	44.833	26.265	0.403	6.645	0.463	0.288
BuyBest 5/90	488.25	58.714	26.331	-0.009	4.541	0.348	0.652
BuyBest 10/90	225.44	39.099	24.214	-0.069	4.188	0.371	0.283
BuyBest 10/60	252.64	41.759	21.533	-0.277	4.202	0.347	0.338
BuyBest 15/30	207.89	37.263	16.489	-0.506	5.959	0.254	0.381
BuyBest 20/30	228.92	39.451	16.419	-0.469	5.534	0.249	0.427
BuyBest 25/30	229.44	39.504	15.766	-0.403	5.741	0.227	0.470
BuyBest 30/30	224.23	38.975	15.485	-0.37	5.412	0.227	0.459

Table 4.3.: The best BuyBest strategies

4.5. Evaluation

The strategies found in Section 4.4 are now evaluated to see if they can be applied to new data and how well they perform. The period is set to half a year, starting January 1, 2012 and ending June 1, 2012.

The Table 4.4 shows the total return of the strategies. Therefore, if the best strategy had been chosen, a profit of 11 % would have been made. The complete statistics can be found in Table A.4 in the appendix. The measurements are the same as in Section 4.4 exception of the volatility which is measured in a weekly and monthly frequency.

name	return (%)	RRR	name	return (%)	RRR
S&P 500	0.077	0.007	S&P 500	0.077	0.007
SingleSMA 84	11.01	0.444	BuyBest 5/90	8.33	0.148
SingleSMA 80	9.93	0.442	DoubleSMA 25/100	0.65	0.046
SingleSMA 78	10.52	0.425	BuyBest 20/30	-4.64	-0.161
SingleSMA 76	8.49	0.336	BuyBest 30/30	-4.2	-0.166
DoubleSMA 21/200	7.18	0.295	BuyBest 25/30	-5.14	-0.193
DoubleSMA 5/100	6.15	0.235	DoubleSMA 25/300	-7.67	-0.539

Table 4.4.: Total returns from 2012-01-01 until 2012-06-01

BuyBest strategies did not perform as well as before. On the other side, all SingleSMA strategies had gained money. The strategy with the biggest return of 11.02% and Risk-Return-Ratio (RRR) of 0.444 is 'SingleSMA 84'. Figure 4.5 shows the weekly logarithmic return and the excess return of this strategy. The risk the strategy takes can been seen in the plot of the weekly return. Strategy 'SingleSMA84' has extreme returns in both directions (negative and positive). On the other side, the S&P 500 has a flatter curve which means it has less extreme values.

Most strategies outperform the S&P 500 index in rate of return but they have a higher risk (volatility) than the index. However, the 'SingleSMA 84' has roughly 63 times bigger RRR than the S&P 500 which means for every "unit of risk", one gets 63 "units of return". Hence, it is worth to take the additional risk of the 'SingleSMA 84' strategy with respect to its return.
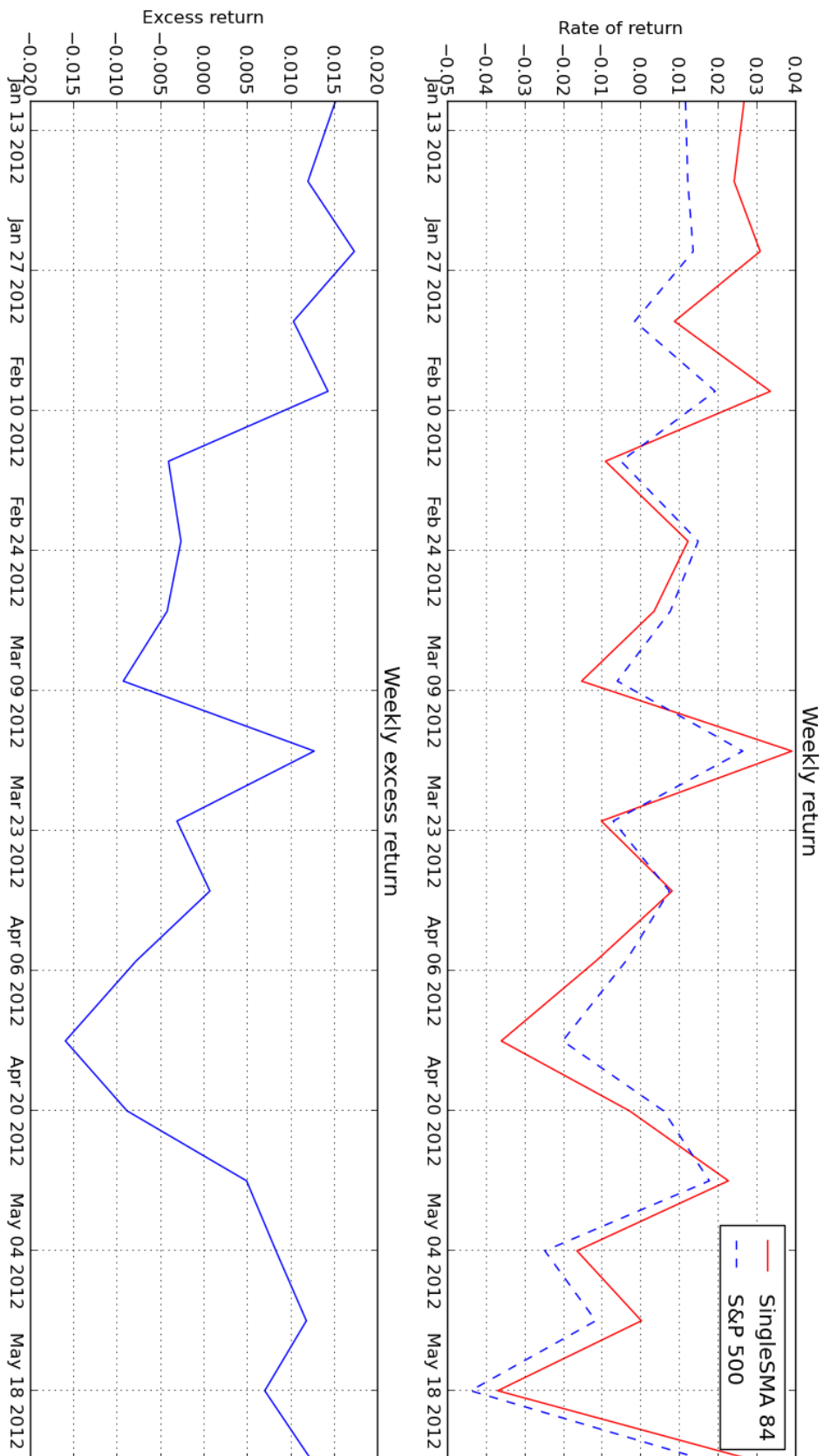
Figure 4.5.: Weekly log and excess return from SingleSMA 84

CHAPTER 5

Conclusion

This thesis addressed two objectives. Firstly, a framework which allows to define, test and analyse a trading strategy has been developed. This framework is easily adaptable to new constraints. The second goal was to find simple strategies and test these with historical data. The analysis shows that these strategies can outperform the S&P 500 and are adaptable to new data. The evaluation showed the higher the return, the higher the risk. The challenge is to find a strategy with a high rate of return and a low risk.

5.1. Outlook

The functionality provided by the framework is useful and allow fast evaluation results of strategies, However, some shortcomings have to be addressed in the near future.

At the moment, the framework can only handle long position. An implementation of short-sell should be consider which would allow to gain profit from falling prices. This would expand the possibility to gain money and create new opportunities for strategies. In the world of trading, large data sets have to be handled and analysed to find patterns and correlations. Data mining, including neural networks, cluster analysis and genetic algorithms, would be a good start to handle this kind of problems. Machine learning algorithms can be used to find new strategies. This could be done on different levels of the framework (indicators, agents or strategies).

Beside the strategies, it is also important to have good measurements for the performance analysis. At the moment, a basic measurement based on the Markowitz model (MVM) is implemented. The standard deviation is not an ideal risk measure since it penalises profits as well as losses. Therefore, it should be considered a different approach/models to measure the performance and the risk of a portfolio/strategy.

Appendix

A.1. Installation

This is a short description how to install MoneyMaker on a Ubuntu 12.04.

A.1.1. Packages

Following packages are needed to run MoneyMaker: python, python-numpy, python-scipy, python-matplotlib and python-datautile $>= 1.5$.

They can be installed using the package-manager.

```
$ sudo apt-get install python python-numpy python-scipy
  python-matplotlib python-datautile
```

Additionally, easy_install , a python package manager, should be installed. It can be found in the 'python-setuptools' package.

```
$ sudo apt-get install python-setuptools
```

Easy_install helps to install packages which could not be found in the Ubuntu repositories. The following command will install pytz and pandas.

```
$ easy_install pytz
$ easy_install pandas
```

All needed packages for running the program are now installed but for ease of use Ipython is recommended.

```
$ sudo apt-get install ipython
```

A.2. How to Use

There are two easy way to run the program, with the GUI and with Ipython. Make sure a internet connection is available.

A.2.1. GUI

The GUI can be installed from Ubuntu package provided. After installation the stock data should be copied to */usr/share/moneymaker/stocks/*. The GUI can now be started with the following command:

```
$ moneymaker
```

If the Ubuntu packages is installed, it can also be started in Ipython console:

```
In [1]: import moneymaker
In [2]: moneymaker.main()
```

A.2.2. Ipython

In the source folder the program can be started with the following commands:

```
$ ipython run.py -i
```

Now their are lot of different option which can be set. Here, is a demonstration of possible setup:

```
In [1]: analysis = run("SingleSMA",["BHI","JNJ"],
"2012-01-1","2012-06-01")
In [2]: analysis.plot_asset()
In [3]: analysis.plot_excess()
In [4]: plt.show()
```

This will run the default SingleSMA strategy from January 1, 2012 until June 1, 2012 with the stocks JNJ and BHI. It also plots the asset return and the excess return.

A.3. Statistics

name	return (%)	mean (10^{-5})	var. (10^{-5})	std. (10^{-3})	skew.	kur.	vol. monthly	vol. yearly	RRR
S&P 500	-13.58	-4.836	19.223	13.865	-0.153	7.051	0.057	0.210	-0.309
SingleSMA 4	-17.12	-6.22	8.706	9.33	-0.15	4.135	0.055	0.24	-0.341
SingleSMA 8	-27.89	-10.837	8.344	9.135	-2.421	34.524	0.054	0.233	-0.571
SingleSMA 12	-17.75	-6.473	7.428	8.619	-0.613	5.983	0.047	0.18	-0.471
SingleSMA 16	-23.93	-9.064	8.667	9.309	-2.59	49.922	0.053	0.193	-0.592
SingleSMA 20	-32.54	-13.045	8.514	9.227	-2.002	27.106	0.054	0.214	-0.726
SingleSMA 24	-22.93	-8.632	8.574	9.26	-1.235	9.894	0.051	0.229	-0.478
SingleSMA 28	-23	-8.66	8.042	8.968	-1.188	9.501	0.05	0.221	-0.497
SingleSMA 32	-21.86	-8.174	7.551	8.69	-1.02	7.515	0.051	0.233	-0.448
SingleSMA 36	-9.37	-3.261	8.023	8.957	-0.637	4.344	0.048	0.212	-0.211
SingleSMA 40	-15.81	-5.702	8.163	9.035	-1.185	9.258	0.05	0.246	-0.307
SingleSMA 44	-0.21	-0.069	7.092	8.421	-0.881	8.46	0.045	0.214	-0.005
SingleSMA 48	-8.61	-2.984	8.51	9.225	-1.134	11.492	0.052	0.238	-0.173
SingleSMA 52	-22.75	-8.551	8.995	9.484	-0.878	11.883	0.052	0.236	-0.460
SingleSMA 56	-6.14	-2.1	7.804	8.834	-0.771	5.835	0.049	0.224	-0.131
SingleSMA 60	1.68	0.552	8.464	9.2	-0.652	4.987	0.053	0.22	0.036
SingleSMA 64	8.64	2.746	13.684	11.698	0.882	70.529	0.053	0.262	0.157
SingleSMA 68	-24.51	-9.317	9.781	9.89	-0.922	9.319	0.057	0.277	-0.422
SingleSMA 72	4.66	1.509	6.861	8.283	-0.44	3.997	0.045	0.213	0.104
SingleSMA 76	36.03	10.195	7.715	8.784	-0.601	5.284	0.046	0.182	0.945
SingleSMA 80	26.52	7.794	10.784	10.385	0.308	59.554	0.051	0.198	0.639
SingleSMA 84	10.06	3.176	6.274	7.921	-0.512	4.538	0.041	0.172	0.279
SingleSMA 88	-12.93	-4.588	6.284	7.927	-1.17	12.376	0.043	0.192	-0.321
SingleSMA 92	-12.6	-4.464	6.678	8.172	-0.372	3.875	0.043	0.189	-0.318
SingleSMA 96	-15.37	-5.53	8.546	9.245	-0.833	5.829	0.05	0.245	-0.299
SingleSMA 100	8.49	2.69	8.09	8.99	-0.515	6.556	0.044	0.214	0.189

Table A.1.: Analysis of SingleSMA from 2000-01-01 until 2011-12-31

Thomas Bürli

name	return (%)	mean (10^{-5})	var. (10^{-5})	std. (10^{-3})	skew.	kur.	vol. monthly	vol. yearly	RRR
S&P 500	-13.58	-4.836	19.223	13.865	-0.153	7.051	0.057	0.210	-0.020
DoubleSMA 2/50	-35.18	-14.367	9.266	9.626	-1.254	13.74	0.052	0.255	-0.042
DoubleSMA 2/300	94.73	22.082	10.417	10.206	-0.41	4.479	0.049	0.222	0.13
DoubleSMA 2/250	-4.7	-1.595	8.719	9.337	-0.34	8.02	0.043	0.218	-0.007
DoubleSMA 2/200	86.72	20.69	13.063	11.429	-0.474	7.909	0.046	0.239	0.111
DoubleSMA 2/100	65.76	16.745	18.271	13.517	2.104	44.025	0.075	0.264	0.076
DoubleSMA 5/50	2.83	0.924	8.561	9.253	-0.801	8.516	0.047	0.195	0.004
DoubleSMA 5/300	34.8	9.895	13.545	11.638	-0.06	4.76	0.052	0.21	0.051
DoubleSMA 5/250	136.53	28.525	11.033	10.504	-0.715	15.422	0.049	0.175	0.238
DoubleSMA 5/200	7.69	2.453	10.327	10.162	-0.801	6.611	0.054	0.26	0.009
DoubleSMA 5/100	271.25	43.463	22.246	14.915	0.091	8.98	0.066	0.194	0.427
DoubleSMA 10/50	79.15	19.319	7.645	8.744	0.392	12.605	0.038	0.188	0.129
DoubleSMA 10/300	34.42	9.802	6.595	8.121	1.062	23.931	0.029	0.121	0.087
DoubleSMA 10/250	122.96	26.568	25.402	15.938	-0.099	15.288	0.056	0.262	0.143
DoubleSMA 10/200	22.76	6.794	17.617	13.273	-0.436	3.83	0.059	0.23	0.03
DoubleSMA 10/100	251.02	41.606	17.885	13.374	-0.433	6.639	0.063	0.243	0.316
DoubleSMA 20/50	41.74	11.557	10.958	10.468	-0.568	10.269	0.048	0.243	0.052
DoubleSMA 20/300	57.57	15.067	14.724	12.134	-0.066	9.121	0.051	0.182	0.097
DoubleSMA 20/250	396.16	53.073	34.34	18.531	1.029	16.508	0.076	0.334	0.362
DoubleSMA 20/200	91.22	21.48	14.286	11.952	-0.449	6.705	0.056	0.191	0.146
DoubleSMA 20/100	133.03	28.032	11.619	10.779	0.033	11.076	0.04	0.145	0.28
DoubleSMA 21/50	107.3	24.155	15.482	12.443	-0.629	18.118	0.05	0.249	0.132
DoubleSMA 21/300	-18.2	-6.657	26.816	16.375	-0.148	9.244	0.071	0.259	-0.021
DoubleSMA 21/250	106.54	24.033	22.12	14.873	0.8	21.316	0.062	0.25	0.13
DoubleSMA 21/200	294.4	45.467	33.668	18.349	-0.469	10.187	0.077	0.236	0.381
DoubleSMA 21/100	225.93	39.149	16.027	12.66	-0.824	12.765	0.059	0.218	0.317
DoubleSMA 25/50	91.55	21.536	10.739	10.363	-0.481	9.025	0.049	0.179	0.156
DoubleSMA 25/300	255.31	42.008	22.922	15.14	-0.291	5.798	0.062	0.154	0.506
DoubleSMA 25/250	119.51	26.051	37.955	19.482	-0.438	13.926	0.077	0.305	0.12
DoubleSMA 25/200	161.46	31.847	20.136	14.19	0.315	10.273	0.061	0.248	0.199
DoubleSMA 25/100	309.38	46.702	17.746	13.321	0.439	9.266	0.064	0.24	0.394

Table A.2.: Analysis of DoubleSMA from 2000-01-01 until 2011-12-31

name	return (%)	mean (10^{-5})	var. (10^{-5})	std. (10^{-3})	skew.	kur.	vol. monthly	vol. yearly	RRR
S&P 500	-13.58	-4.836	19.223	13.865	-0.153	7.051	0.057	0.210	-0.030
BuyBest 5/30	104.96	23.778	58.387	24.164	-0.156	6.983	0.112	0.259	0.188
BuyBest 5/60	286.92	44.833	68.983	26.265	0.403	6.645	0.112	0.463	0.288
BuyBest 5/90	488.25	58.714	69.33	26.331	-0.009	4.541	0.113	0.348	0.652
BuyBest 5/210	-72.19	-42.405	80.233	28.325	-0.026	6.773	0.135	0.488	-0.069
BuyBest 5/300	-80.91	-54.873	105.203	32.435	0.029	6.263	0.152	0.532	-0.071
BuyBest 10/30	125.14	26.891	30.282	17.402	-0.418	6.126	0.083	0.261	0.223
BuyBest 10/60	252.64	41.759	46.367	21.533	-0.277	4.202	0.097	0.347	0.338
BuyBest 10/90	225.44	39.099	58.632	24.214	-0.069	4.188	0.106	0.371	0.283
BuyBest 10/210	52.08	13.891	78.896	28.088	0.07	4.921	0.13	0.342	0.071
BuyBest 10/300	67.92	17.174	81.54	28.555	-0.028	4.75	0.136	0.416	0.076
BuyBest 15/30	207.89	37.263	27.188	16.489	-0.506	5.959	0.079	0.254	0.381
BuyBest 15/60	160.16	31.681	40.991	20.246	-0.243	4.993	0.092	0.34	0.219
BuyBest 15/90	115.45	25.433	44.696	21.141	-0.091	6.832	0.098	0.314	0.171
BuyBest 15/210	51.25	13.71	60.404	24.577	0.014	6.163	0.123	0.331	0.072
BuyBest 15/300	27.19	7.968	73.321	27.078	0.133	4.841	0.125	0.371	0.034
BuyBest 20/30	228.92	39.451	26.957	16.419	-0.469	5.534	0.077	0.249	0.427
BuyBest 20/60	104.99	23.784	36.153	19.014	-0.338	6.308	0.086	0.288	0.169
BuyBest 20/90	114.22	25.243	39.882	19.97	-0.195	6.781	0.092	0.29	0.183
BuyBest 20/210	42.7	11.782	54.513	23.348	0.063	6.218	0.114	0.303	0.066
BuyBest 20/300	33.07	9.466	64.896	25.475	0.1	4.834	0.121	0.343	0.045
BuyBest 25/30	229.44	39.504	24.856	15.766	-0.403	5.741	0.071	0.227	0.470
BuyBest 25/60	122.68	26.527	35.639	18.878	-0.381	6.098	0.085	0.289	0.197
BuyBest 25/90	121.59	26.363	38.377	19.59	-0.166	6.991	0.093	0.278	0.203
BuyBest 25/210	38.7	10.841	49.434	22.234	0.106	5.968	0.107	0.279	0.064
BuyBest 25/300	28.35	8.27	54.657	23.379	0.054	5.446	0.111	0.333	0.040
BuyBest 30/30	224.23	38.975	23.977	15.485	-0.37	5.412	0.071	0.227	0.459
BuyBest 30/60	128.84	27.431	35.353	18.802	-0.332	6.852	0.087	0.284	0.211
BuyBest 30/90	160.41	31.713	37.996	19.492	-0.169	6.174	0.092	0.279	0.267
BuyBest 30/210	35.95	10.175	45.624	21.36	0.046	5.475	0.102	0.275	0.061
BuyBest 30/300	75.88	18.708	50.176	22.4	-0.034	5.212	0.109	0.328	0.108

Table A.3.: Analysis of BuyBest from 2000-01-01 until 2011-12-31

name	return (%)	mean (10^{-5})	var. (10^{-5})	std. (10^{-3})	skew.	kur.	vol. weekly	vol. monthly	RRR
S&P 500	0.077	0.738	5.756	7.587	-0.256	0.820	0.017	0.024	0.007
SingleSMA 76	8.49	78.336	10.929	10.454	-0.168	0.08	0.028	0.055	0.336
SingleSMA 78	10.52	96.213	10.364	10.18	-0.149	0.261	0.025	0.054	0.425
SingleSMA 80	9.93	91.045	9.689	9.843	-0.132	0.285	0.024	0.049	0.442
SingleSMA 84	11.01	100.414	9.102	9.541	-0.09	0.291	0.022	0.054	0.444
DoubleSMA 5/100	6.15	57.366	12.324	11.102	-0.386	0.686	0.024	0.057	0.235
DoubleSMA 21/200	7.18	66.654	14.604	12.085	-0.482	1.039	0.024	0.053	0.295
DoubleSMA 25/100	0.65	6.259	10.003	10.001	0.059	-0.272	0.022	0.031	0.046
DoubleSMA 25/300	-7.67	-76.708	14.032	11.846	0.024	0.467	0.023	0.031	-0.539
BuyBest 5/90	8.33	76.947	69.215	26.309	0.022	0.84	0.06	0.123	0.148
BuyBest 20/30	-4.64	-45.725	18.56	13.623	-0.072	0.036	0.03	0.063	-0.161
BuyBest 25/30	-5.14	-50.697	17.146	13.094	-0.185	0.179	0.029	0.058	-0.193
BuyBest 30/30	-4.2	-41.269	14.819	12.173	-0.22	0.273	0.027	0.055	-0.166

Table A.4.: Evaluation of Strategies from 2012-01-01 until 2012-06-01

A.4. S&P 500

List of stock symbols which ever where part the Standard & Poor 500 Index:

AAMRQ, AAPL, AA, ABC, ABKFQ, ABS, ABT, ABX, ACAS, ACE, ACN, ACS, ACV, ACY_1, ADBE, ADCT, ADI, ADM, ADP, ADSK, AEE, AEP, AES, AET, AFL, AFS, AGC_1, AGN, AHM_1, AIG, AIT_1, AIV, AIZ, AKAM, AKS, AL_1, ALL, ALTR, ALXN, ALY, AMAT, AMCC, AMD, AMGN, AMH_1, AMI_1, AMO_1, AMP, AMT_1, AM, AMT, AMX_2, AMZN, ANDW, ANF, ANR, AN, AON, APA, APCC, APC, APD, APH, APOL, ARG, AR, ASA, ASH, ASN, ASO_1, AS, A, ATI_1, ATI, AT, AVB, AVP, AVY, AW, AXP, AYE, AZO, BAC, BA, BAX, BBBY, BBT, BBY, BCO, BCR, BC, BDK, BDX, BEAM, BEC, BEN, BEV, BF.B, BFI, BGEN, BGG, BHI, BIG, BIIB, BJS, BKB, BK, BLK, BLL, BLS, BLY, BMC, BMS, BMY, BNL_1, BNSSA, BOAT, BOL, BRCM, BRK.B, BRL, BSC_1, BSET, BSX, BTU, BVSN, BV, BWA, BWN_1, BWS, BWY_1, BXP, CAG, CAH, CAM, CAR, CA, CAT, CBB_1, CBE, CBG, CBS_1, CBS_2, CBSS, CBS, CB, CCB_2, CCE, CCI, CCK, CCL, CCMO, CCTYQ, CE_2, CEG, CELG, CEPH, CERN, CFL_1, CFN, CF, CHIR, CHK, CHRS, CHRW, CIEN, CINF, CI, CIT, CJML, CKL, CLF, CL, CLU_2, CLX, CMA, CMB, CMCSA, CME, CMG, CMI, CMS, CMVT, CMX, CNA, CNO, CNP, CNW, CNX, CNXT, COF, COG, COH, COL, COMS, COP, COST, COV, COX, CPB, CPN, CPWR, CRK_1, CRM, CR, CSCO, CSC, CSR_1, CSX, CTAS, CTB, CTCO, C, CTLE, CTL, CTSH, CTXS, CTX, CVC, CVG, CVH, CVS, CVX, CYM_1, CYR_1, CZR, DAL, DAN, DDR, DDS, DD, DELL, DE, DFS, DF, DGN, DG, DGX, DHI, DHR, DIGI_1, DISCA, DIS, DI, DJ, DLPH, DLTR, DLX, DNA, DNB, DNR, DO, DOV, DOW, DPS, DRI, D, DTE, DTV, DUK, DVA, DVN, DV, DYN, E_2, EA, EBAY, ECH_1, ECK_1, ECL, EDS, ED, EFU_1, EFX, EIX, EKDKQ, EL, EMC, EME, EMN, EMR, ENRNQ, EOG, EOP, EP, EQR, EQ, EQT, ESRX, ESY_1, ETFC, ETN, ETR, ETS, EW, EXC, EXPD, EXPE, FAST, FBF, FBN, FBO, FCF, FCN_1, FCX, FDO, FDX, FE, FFB, FFIV, FHN, FII, FIS, FISV, FITB, FJCC, FJ, FLC_2, FLIR, FLMIQ, FLR, FLS, FL, FLTWQ, FMC, FMY_1, FNB, FNP, FORT, FOSL, FPA, FRO_1, FRX, FSLR, F, FTI, FTR, FWLT, G_1, GAPTQ, GAS, GCI, GCO, GD, GDT, GDW, GEN, GENZ, GE, GGP, GIC, GIDL, GILD, GIS, GLBC, GLK_1, GLW, GME, GM, GNW, GOOG, GOSHA, GPC, GPS, GPU, GQ, GRA, GRH_1, GRN_1, GR, GSF_2, GS, GTE_1, GT, GTW, GWF, GWOW, GWW, H_2, HAL, HAR, HAS, HBAN, HCA_1, HCA, HCBK, HCN, HCP, HCR, HDLM, HD, HES, HFS, HIG, HLS, HLT, HMA, HM, HNH, HNZ, HOG, HON, HOT, HPC, HPQ, HP, HRB, HRL, HRS, HSP, HST, HSY, HTMXQ, HT, HUM, IACI, IBM, IBP, ICE, ID, IFF, IGT, IKN, IK, IMNX, INCLF, INGR, INTC, INTU, IPG, IP, IRM, IR, ISRG, I, ITT_1, ITT, ITW, IVZ, JAVA, JBL, JCI, JCP, JDSU, JEC, JHF, JH, JM, JNJ, JNPR, JNS, JNY, JOL, JOY_1, JOY, JPM_1, JP, JWN, KBH, KEY, KFT, KG, KIM, KLAC, KMB, KMI, KMX, KOE, KO, KRB, KRI, KR, KSE, KSS, KSU, K, KWP, LCE, LDG, LEG, LEHMQ, LEN, LH, LIFE, LINB, LIT_1, LLL, LLTC, LLX, LLY, LM, LMT, LNC, LO, LOTS, LOW, LPX, LRCX, LSI, LSWY, LTD, L, LTVCQ, LUB, LUK, LU, LUV, LXK, MAI_1, MAR, MAS, MA, MAT, MAY_2, MBI, MCAWA, MCD, MCHP, MCIC, MCIP, MCK, MCO, MD_1, MDP, MDR, MDT, MEA_1, MEDI, MEE, MERQ, MET_2, MET, MFE, MGI, MGM_1, MHP, MHS, MII, MIL, MIR_1, MI, MJN, MKC_1, MKC, MKG, MMC, MMI, MMM, MNR_1, MOH, MOLX, MON, MOS, MO, MPC, MRK, MRO, MSFT, MSI, MS, MST, MTB, M, MTG, MTW, MUR, MU, MWI, MWV, MWW, MXIM, MXP, MXS, MYG_1, MYL, MZIAQ, NAV, NBL, NBR, NCC, NCR, NC, NDAQ, NEC, NEE, NEM, NES_2, NE, NEU, NFB, NFLX, NFX, NGCO, NGH, NI, NKE, NL, NLT, NMK, NOC, NOVL, NOV, NRG, NRTLQ, NSC, NTAP, NTRS, NUE, NU, NVDA, NVLS, NWA, NWL, NWSA, NXTL, NYN, NYT, NYX, OAT, OC, ODP, OI, OKE, OK, OMC, OMX, ONE_1, ORCL, ORLY, ORX, OXY, PALM, PAS, PAYX, PBCT, PBG, PBI, PBY, PCAR, PCG, PCH, PCI, PCLB_1, PCLN, PCL, PCP, PCS_1, PCS, PDCO, PDG, PD, PEG, PEL_1, PEP, PFE, PFG, PGN, PGR, PG, PHA_1, PHB_1, PHM, PH, PKI, PLD, PLL, PMCS, PMTC, PM, PNC, PNU_1, PNW, POM, PPG, PPL, PRGO, PRI, PRU, PSA, PSFT, PSX, PT_1, PTV, PVN, PVT, PWER, PWJ, PWR, PXD, PX, PZE_1, QCOM, QEP, QLGC, Q, RAD, RAI, RAL, RATL, RBD, RBK, RDC, RDMN, RDPL, RFDS, RF, RHI, RHT, RIG, RL, RNGR, ROH, ROK, ROP, ROST, RRC, RRD, RSG, RSH, R, RTN, RYAN, RYC, RYI, S_1, SAA_1, SAF, SAI_1, SAI, SANM, SAPE, SBUX, SCG, SCHW, SCI, SEBL, SED_1, SEE, SE, SFA_1, SFN, SFS, SGICQ, SGP, SHLD, SHW, SIAL, SII, SJM, SK_1, SKY, SLB, SLE, SLM, SLR, SLZ, SMI_1, SNA, SNC_1, SNDK, SNI, SNT_1, SNV, SO, SOTR, SPG, SPLS, SPW, SRCL, SRE, SRR, SSP, S, STF, STI, STJ, STK_1, STR, STT, STY_1, STZ, SUN, SVU, SWK, SWN, SWY, SXCL, SYK, SYMC, SYN_1, SYY, TA_1, TAP, TCOMA, TDC, TDM, TEG, TEL, TEN, TER, TE, TEX, TFB, TFD, TGT, THC, THY, TIE, TIF, TIN, TJX, TKR, TLAB, TMC, TMK, TMO, TNB, TOS, TRB, TRIP, TROW, TRV, TRW_1, TSN, TSO, TSS, T, TTRR, TUP, TWC, TWX_1, TWX, TXN, TXT, TYC, UAL, UCL, UH, UIS, UNH, UNM, UNP, UN, UPC, UPS, URBN, USBC, USB, USG, USHC, UST_1, UTX, VAR, VAT, VC, VFC, VIAB, VLO, VMC, VNO, VRSN, VRTS_1, V, VTR, VTSS, VZ, WAG, WANG, WAT, WB_1, WB, WDC, WEC, WFC, WFM, WFR, WFT, WHR, WINN, WIN, WLA, WLB, WLL_1, WLP_1, WLP, WMB, WMIH, WM, WMT, WMX, WOR, WPI, WPM, WPO, WPX, WU, WWY,

WYE, WYNN, WYN, WY, XEL, XLNX, XL, XOM, XRAY, XRX, X, XTO, XYL, YHOO, YRCW, YUM, ZION, ZMH, ZRN
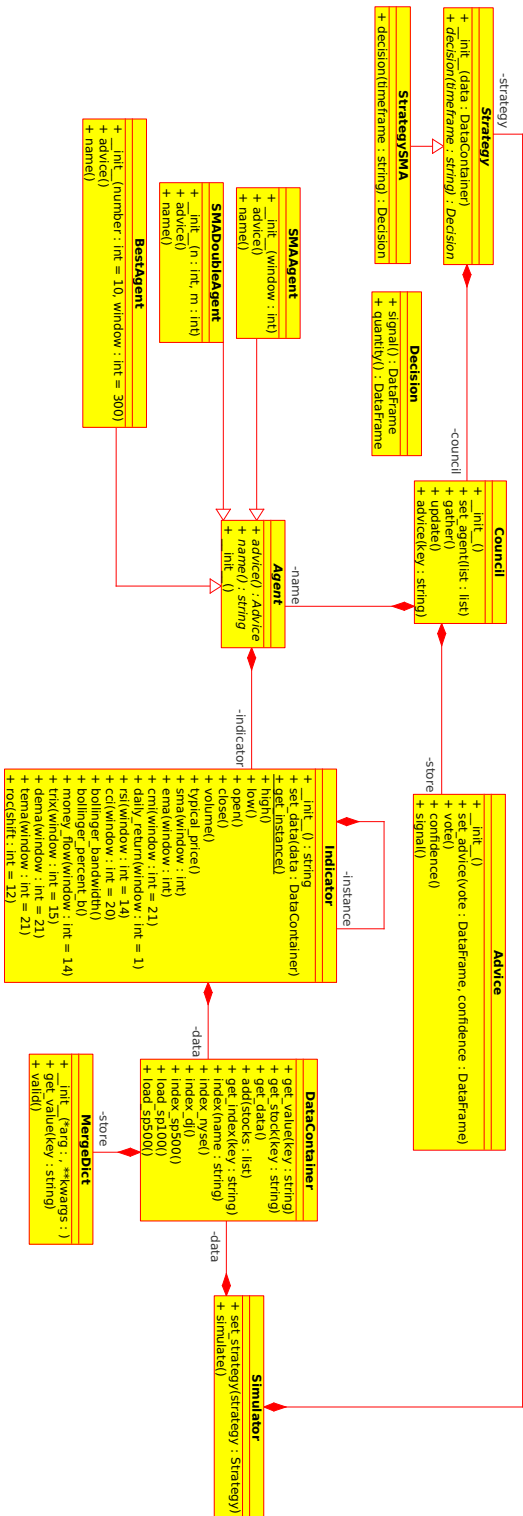
A.5. Class

Figure A.1.: UML Diagram

A.6. Original Problem

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Distributed Computing

Prof. R. Wattenhofer

Lab/BA/SA/Group:
Make Us Rich!

Motivation and Informal Description

In the recent past the algorithmic trading has seen enormous growth and a good place to make lots of money. It is now responsible for more than half the trades in the US which have a total volume of \$27.4 billion a day[1]. A very important subclass are the high frequency trading (HFT) algorithms. These algorithms usually hold stocks or certificates only for a brief time, sometimes only for a few seconds and earn money by making thousands of trades a day.

But since these algorithms increase the volatility of the market, they might be the target of a financial transaction tax which would destroy that business model.

Therefore, we want to return to normal algorithmic trading to get rich. We want to exploit statistical properties of the assets to predict general trends and/or foresee the future value of certain stocks and use those predictions to beat the market.

Your job is to analyze large data sets to be able to make those predictions and implement an algorithm which automatically trades.

Requirements

Good programming skills (preferably in Java) and a genuine interest in the financial markets. Knowledge of data mining is an advantage. The student(s) should be able to work independently on this topic!

Interested? Please contact us for more details!

Contact

- Philipp Brandes: philipp.brandes@tik.ee.ethz.ch, ETZ G64.2

---

[1] value from November 2011

Bibliography

Acar, E. and Satchell, S. (2002). *Advanced Trading Rules.* Elsevier Ltd, second edition.

Achelis, S. (2000). *Technical Analysis from A to Z.* McGraw-Hill, second edition.

Alexander, C. (2001). *Market Models: A Guide to Financial Data Analysis.* Wiley.

BOWLEY, G. (2011). Clamping down on rapid trades in stock market. http://www.nytimes.com/2011/10/09/business/clamping-down-on-rapid-trades-in-stock-market.html.

Boyarshinov, V. and Magdon-Ismail, M. (2005). Efficient computation of optimal trading strategies.

Brown, S. (2011). Measures of shape: Skewness and kurtosis. http://www.tc3.edu/instruct/sbrown/stat/shape.htm.

Easley, D., Lopez de Prado, M. M., and O'Hara, M. (2012). The volume clock: Insights into the high frequency paradigm. *Journal of Portfolio Management*, page 11.

Economist (2011). Momentum in financial markets. `http://www.economist.com/node/17848665`.

Holton, G. A. (2003). Negatively skewed trading strategies. *Derivatives Week*.

Investopedia (2012a). Standard & poor's 500 index - s&p 500. `http://www.investopedia.com/terms/s/sp500.asp#axzz1xWXVkcE1`.

Investopedia (2012b). Technical indicator. `http://www.investopedia.com/terms/t/technicalindicator.asp#axzz23nEPdqDd`.

Kannan, K. S., Sekar, P. S., Sathik, M., and Arumugam, P. (2010). Financial stock market forecast using data mining techniques.

Lauricella, T. and STRASBURG, J. (2010). "how a trading algorithm went awry". *The Wall Street Journal*.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7:77–91.

Mehta, N. (2012). Knight $440 million loss sealed by rules on canceling trades. http://www.bloomberg.com/news/2012-08-14/knight-440-million-loss-sealed-by-new-rules-on-canceling-trades.html.

PRNewswire (2010). Bns holding inc. announces financial results for the quarters ended july 31, 2010 and april 30, 2010. http://www.prnewswire.com/news-releases/bns-holding-inc-announces-financial-results-for-the-quarters-ended-july-31-2010-and-april-30-2010-106333408.html.

Roxburgh, C., Lund, S., and Piotrowski, J. (2011). Mapping global capital markets 2011. Technical report, McKinsey Global Institute.

Taylor, M. P. and Allen, H. (1992). The use of technical analysis in the foreign exchange market. *Journal of International Money and Finance*, 11(3):304 – 314.

Today, M. F. (2010). Systematic trading, systematic risk control. http://www.managedfuturestodaymag.com/systematic-trading-systematic-risk-control.

TSAY, R. S. (2010). *Analysis of Financial Time Series*. A JOHN WILEY & SONS, INC, third edition.

Vince, R. (1992). *Mathematics of Money Management: Risk Analysis Techniques for Traders*. John Wiley & Sons, Inc.

Wikipedia (2012). Skewness statistics. `http://en.wikipedia.org/wiki/File:Skewness_Statistics.svg`.