DEPARTMENT OF COMPUTER SCIENCE

COMPUTATIONAL EXAMINATION OF CHARTISTS' STRATEGIES AND FUNDAMENTALISTS' STRATEGIES AND THEIR USE OF PORTFOLIO OPTIMISATION AND PROJECTION

JOHN O'AKUWANU

MAY 5, 2021

Supervisor: Dimitar Kazakov

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

B.ENG. COMPUTER SCIENCE (WITH A YEAR IN INDUSTRY)

Contents

1 | Executive Summary

1.1 Motivation

There are several motivations behind this research. Firstly, there is a surge in individual investing, especially with the rise of trading platforms, such as FreeTrade and Robinhood. However, many individual investors are overwhelmed by the research aspect of deciding what assets to purchase or sell - the dissertation aims to provide insight to show effective investing methods.

Secondly, there is a continuous debate between fundamentalists and chartists regarding fundamental vs technical analysis. The debate also calls into question how to decide what stocks and assets to purchase and determine when to buy or sell a stock.

Other motivations include identifying the best technical indicator and the best parameters for each technical indicator. Various parameters can take arbitrary values. For instance, if a technical indicator uses a moving average as part of its method, what should the number of days for which the moving average is calculated?

The majority of preceding research papers fixate on one set value and then evaluated how profitable that indicator is. The problem with this approach is that it leads to data dredging, where Sullivan [1] shows that the data dredging leads to selection bias whereby the fixed parameters are bound to work even on a table of random numbers. Logical inference due to common cognitive dissonance where humans tend to alter the results to match their beliefs.

This dissertation attempts to minimise this effect by collecting trading rules that include various combinations of the parameters. Another motivation is to gauge how well a portfolio's performance can be predicted as the dissertation is tailored for individual investors and accurate prediction of portfolio performance helps plan according to one's personal goals.

1.2 Hypothesis

The hypothesis under assessment is whether there will be a substantial similarity between the best-performing stocks per sector for each form of analysis. The fact that that fundamental analysis is the household view industry-wide will outperform technical analysis. However, a composite analysis taking both forms of analysis will provide the best allocation of stocks for profitability and minimised risk. For portfolio optimisation, I believe Black Litterman will be more beneficial for a user to incorporate an individual's views. Nevertheless, the capital asset price model portfolio will achieve the highest returns for overall profitability since it is the original method for portfolio optimisation.

1.3 Method

Backtesting will occur with various known technical indicators and its performance in terms of risk, profits and volatility are measured. Backtesting will also occur with a selection of fundamentally strong stocks with and without undergoing different portfolio optimisation strategies and the same metrics will be measured. Statistical tests will be implemented to identify statistical significance between different strategies.

1.4 Results

Fundamental analysis still has a prominent role to play in financial markets. Fundamental analysis proves to be superior to Technical analysis in the long run regarding profitability and risk to reward ratio; however, technical analysis is better in risk aversion.

Passive investing is better than active investing in the long run as the cumulative returns aspect of passive investing profits investors more than small gains through repeatedly buying and selling stocks at the best entry and exit points.

Portfolio optimisation helps improve investor returns, and in theory, they can beat the S&P 500 index and funds in a bull market, but conversely, in a bearish market, institutions and the S&P 500 index are far superior. Portfolio projection techniques still have a relevant role in gauging a portfolio's potential. Individual investors should opt for mutual funds as fundamental analysis is time exhaustive so the time efficient method of investing is long term investing through mutual funds.

2 | Introduction

The project aims to study a problem in algorithmic trading with deciding the best trading discipline (technical or fundamental analysis) and explore the benefits of portfolio optimisation and forecasting. The rise of retail investors has had a remarkable effect on the stock market, such as the recent short squeeze by many individual investors that actively bought fundamentally weak stocks that hedge funds shorted. Similar events have occurred in the past, for example, the Dotcom bubble, where there was irrational purchasing of stock(s) due to group thinking that disrupted the market trend. Many retail investors are looking towards more short term forms of analysis, such as technical analysis.

There are two leading schools of thought deciding what asset or stock is worthwhile to be included in a portfolio. They are popularly referred to as Fundamental Analysis and Technical Analysis. Various methods are used to assess the assets' valuation, potential profitability and risk of owning the asset - the research is necessary to evaluate whether a stock or asset should be included in a portfolio.

The project performs both technical and fundamental trading disciplines, implements a back-testing process to evaluate their performance, portfolio optimisation to mitigate risk and hopefully improve performance, and financial forecasting to evaluate predictive capabilities. This work will act as the foundation for potential democratised machine learning technologies for participants in the financial market in the context of discovering investment opportunities. To judge my findings' performance, risk and profit metrics will be taken into account, and the self-made portfolios will be compared to mutual funds and the S&P 500 index. Furthermore, financial forecasting will take place with the self-made portfolios to see their viability in predictive analysis.

3 | Literature Review

3.1 Fundamental analysis and stocks

Fundamental analysis aims at analysing the macroeconomic qualities of a company or asset. Benjamin Graham initially proposed this train of thought in his book, Security Analysis [2]. Since then, vital research has been done in this domain to find the best metrics of companies' qualities and attributes to determine the future price and profitability of a stock or asset. This analysis' underlying belief maintains that markets may incorrectly price an asset in the short run but the market is self-correcting and the price will eventually reach its "true value". Fundamentalists believe that if an asset's current value is underestimated compared to its "true value", it can be bought to make a profit. Vice versa, if the current value is overestimated compared to its "true value", it can be shorted to gain profit. However, it does not warrant that the price reaches its "true value" anytime soon, so the fundamental analysis is preferable in predicting long term trends rather than short-term trading.

3.2 Technical Analysis

Technical Analysis is predicting future price movements of stocks based on the examination of past price movements. Analysts have manufactured numerous technical indicators to predict future price movements of stocks accurately.

Technical analysis relies on past data, mainly price and volume, to make future predictions. The underlying school of thought is that historical patterns tend to repeat themselves.

The gains from technical analysis stem mainly from entering the market and exits to make a profit, for example, buying low and selling high.

Murphy [3] describes three premises on which technical analysis is based:

1. Market action discounts everything.
2. Prices move in trends.

3. History repeats itself.

Academics often reprimand technical analysis for its lack of scientific and statistical rigour and validation [3]. In response, technical analysts (chartists) often argue that technical analysis is a practical discipline, primarily interested in what works rather than existing theory.

3.3 Backtest strategy optimisation

Model validation is the process of assessing whether a model is sufficient to achieve a set goal or perform a strategy. One of the tools of conducting validation is by backtesting. Therefore, appropriate backtesting techniques should be used to verify the capabilities of current strategies and predictive analysis models [4].

Backtesting is the practice of utilising historical financial data to test trading strategies or analytical models to see how well the strategy or models would perform [5]. Backtesting acts upon the concept of Laplace's demon, where the present is the effect of the past and the cause of the future. To put it simply, if the strategy had worked in the past, it would be considered likely to garner profit in future, and conversely, if the strategy had failed previously, it would be considered unlikely to perform well in the future. A well-formed backtested strategy should be viable to make relatively accurate predictions of the future.

With the help of technology, programming languages and historical trading data, backtesting is easier to do and makes it possible for investors to test and improve strategies before engaging in any risk in trading in a real-world setting [6].

Various statistical indicators are used for the evaluation of backtesting; for example, Sharpe ratio is used to measure the performance of a trading strategy after adjusting for its risk, dividing the excess return (profits) one may receive by the extra volatility that one endures for performing that strategy (risk) [7].

Previous studies also show that one of the significant problems of backtesting is overfitting. Overfitting is primarily a machine learning problem where a model tunes its coefficient and parameters too closely to the inputs (training data) to the point that it is unable to describe the general structure. This phenomenon is problematic with predictive analysis. Overfitting cannot be wholly eliminated, but its influence can be reduced using regularisation techniques such as Ridge Regression and reducing the number of parameters or testing the same strategy to various historical data types [7].

3.3.1 Portfolio Optimisation

The formal approach towards making investment decisions for obtaining an optimal portfolio with a specific objective requires a mathematical formulation for the problem. Objectives can vary depending on the investor's primary desires. For example, a portfolio composed of companies with a tremendous social impact - evaluated using sentiment analysis and ESG value - or having a risk-averse portfolio or a portfolio to maximise profits.

Major strives in computational capabilities and the growing volumes of data, computational models and algorithms are often developed to validate the mathematical models with empirically available data.

Fundamental breakthrough in asset allocation and portfolio optimisation is often dated to Markowitz's Modern Portfolio Theory [8]. In simple terms, the theory assumes the existence of both risky and risk-free assets, where in most cases, the risk-free rate is lower than the expected return on the risky asset. The logic behind the theory rests on the assumptions that the issuer of the risky asset has to offer investors the expectation of a higher return to persuade them to forgo the risk-free asset else, every investor will opt for the more profitable and safer option. The main flaw with this form of portfolio optimisation is the assumption of operating in a market without any taxes or transaction costs.

An alternative is Sharpe's Capital Asset Pricing Model (CAPM) [9] - this takes into account the asset's sensitivity to any risk that affects a large segment of society while it is being added to a heterogeneous portfolio. It considers the importance of the returns' covariance structure, the variance of the portfolio and the market premium (i.e. the difference between the expected return on the asset from the market and the risk-free rate of return on the asset). Moreover, the model assumes that the investors are rational and risk-averse while ensuring that the portfolio is diversified across various investments and it cannot influence the assets' prices.

Continually, a Bayesian perspective seems to be gaining popularity due to the rapid rise in social media presence. Black-Litterman's model [10] considers including the investors' views as prior information into the estimates of the expected asset returns and the covariance structure. It provides a mechanism to model user-specified confidence levels based on the investors and financial experts' arbitrary opinions, either in totality or partial knowledge.

Lastly, the final form of portfolio optimisation I will be looking into is Fama French 3 Factor and 5 Factor models. Fama and French [11] proposed that three primary factors play an essential role - market premium, a size factor and a B/M factor. Fama and French [12] tested the variation in stock prices with size and B/M ratios, and in 1996 they discovered that the

CAPM model's anomalies marginalised by using the 3 factor Fama model. Their finding was later backed by Gökgöz [13], using the Istanbul Stock Exchange sectors indexes.

The theoretical advances and computational techniques in the literature on financial portfolio optimisation reflect the ongoing and progressive work in this field. The plethora of information available cites the relevance of asset allocation and the optimisation problem in financial studies. This work may not be able to do complete justice in exploring the literature as exhaustive and comprehensive it could be, but it is believed to have covered the broader spectrum of available literature. This work focuses on contributing to the pool of knowledge by looking at the problem from a computational perspective and delving into its applicability to trading disciplines and financial forecasting, despite the challenging nature of market dynamics and growing complexities.

3.4 Portfolio Forecasting

Rather than an analytical approach, the Monte Carlo method uses a series of simulations to approximate the outcome of a model. It does so by generating large quantities of pseudo-random numbers of certain probability distributions. By running the model extensively, the outcomes of the model will converge to a particular result. The law of large numbers causes this convergence.

The initial use of the Monte Carlo model in mathematical finance was by Boyle [14]. In his paper, Boyle shows that Monte Carlo simulation can be used as an alternative approach to option valuation. Boyle highlighted significant advantages the Monte Carlo model had over its predecessors, such as the Black-Scholes Model. For example, state that Monte Carlo can use different distributions, making it helpful in associating probabilistic likelihood to forecasts and potential price movements. Potentially can be used to express the probability density function of underlying returns. Furthermore, the Monte Carlo model has better extensibility as it is extendable to any possible situation, and it is only restrained by the computational capacity of the underlying hardware [15].

Furthermore, the preexisting, such as the Black-Scholes Model, assumes a Geometric Brownian Motion. If financial market behaviour does not have the qualities of a log-normal distribution, then the output would be inaccurate and unreliable [14].

A feature of the Monte Carlo Model's extensibility is its ability to add characteristics to its model, such as the Geometric Brownian Motion assumed in the Black-Scholes Model. Geometric Brownian Motion is a stochastic process occurring in continuous time in which the logarithm of the randomly varying quantity (in this case, stock prices) follows a Brownian Motion with the going trend or growth rate (drift). The Brownian Motion refers to the mathematical model used to describe such random movements prevalent in particle theory [16]. When Geometric Brownian Motion is used in mathematical finance, it incorporates random noise with modelling stock prices - making it account for volatility within the stock market.

4 | Problem Analysis

4.1 Objective

Financial markets are the quintessential complex system, carrying information on individual and aggregate preferences, as revealed through every economic interaction and relation between the individuals, institutions, and businesses embedded within them. It is, therefore, unsurprising that they have been deemed by many as opaque to prediction. The process of investing can be overwhelming to the masses, so I envisaged a system that would help reduce the complexities of research and find the best way of investing.

My main intention is to determine the best collection of stocks and predict the profitability of large-cap stocks (at least 10 billion market cap) and evaluate the performance of fundamental and technical analysis throughout the process. Furthermore, I aim to find the best method of constructing a portfolio for an individual investor.

Experiment and evaluate which portfolio optimisation strategy (CAPM, Fama French 3 Factor and 5 Factor Models, Monte Carlo optimisation, and Black Litterman) provides the best results in terms of profitability, risk mitigation and future projection. Lastly, present a comparable solution in terms of performance to the market standards when measured using industry-specific parameters and evaluate whether the generated portfolio can beat the S&P 500 index and official funds made by institutional competitors within the market.

4.2 Assumptions

One of the critical assumptions is the absence of transaction costs in calculating a portfolio's worth. There are additional costs in the real world, such as brokerage fees, market costs - like the difference in the buying and selling price of stocks dependent on transaction time, taxes, etc. These costs depend on various factors such as country of residence and the platform an individual investor uses; all these factors are beyond the scope of this dissertation.

The Sharpe ratio is a ratio to reflect the risk to reward of a portfolio - it is prominent in Markowitz mean-variance analysis. The ratio measures the excess return over standard deviation [8]. It is the backbone of portfolio optimisation. Over time, however, newer metrics and discovery of imperfections in the existing Markowitz mean-variance model has drawn questioned its trustworthiness. Hence, other models should be tested alongside the mean-variance analysis.

4.3 Mathematical Formulation

The portfolio consists of m assets. The closing prices of all assets form the *price vector* v_t for each time period t. Similarly, $v_t^{(hi)}$ and $v_t^{(lo)}$ denote the highest and lowest prices of the period.

The *rate of return* is

$$\rho_t := \frac{p_t}{p_{t-1}} - 1 \tag{1}$$

If there is no transaction cost, the final portfolio value will be

$$p_f = p_0 \exp(\sum_{t=1}^{t_f+1} r_t) \tag{2}$$

where p_0 is the initial investment amount. The objective of the investor is to maximise p_f for a given time frame t_f with minimal risk instead of investing via government bonds or leaving money in a savings or ISA account.

4.4 Problems

The process of coming up with an innovative trading method is unlikely to be found in research papers, but the paper can provide greater insight to more efficient and safer methods of trading.

My approach of analysing fundamental data may lead to a limited amount of data compared to acting solely on historical price data; publicly traded companies may have only to provide this data annually, and some companies have key metrics omitted from the dataset. Having missing data that dates back to 2005 may produce a few risks when trying to conjure the best possible portfolio and price prediction. The occasional omission of fundamental metrics is problematic if it is a regular occurrence. The risk is considered, and the risk will fundamentally affect short-term investing as the data would probably be outdated or missing, but long term investing and the risks are greatly reduced, but this risk must be kept in mind.

5 | Methodology and Implementation

The initial step of my methodology is data collection and retrieving the data for my selected market (NYSE) from Yahoo Finance and compiling them into a ".csv" file. Once data is collected for fundamental and technical analysis, I proceed to the second step of exploratory data analysis. I tailor each form of analysis, fundamental - comparative analysis within specified sectors and for technical analysis - Q-Learning to evaluate the best technical analysis. For each form of data analysis, data is investigated by plotting and calculating summary statistics to check correlations and attributes. The third step, feature extraction & pre-processing, is where I clean the data, generate the technical indicators to be used as features, and define target variables for training and testing my models. Once I have such a cleaned data-set, constituted exclusively of those labels and features, I proceed to the fourth step, in which I split the data into training and testing periods. Fifthly, group values of a similar range and map the grades assigned to the group to their performance regarding profitability and risk. After that stage, I will implement back-testing in portfolio optimisation to evaluate the profitability of different forms of analysis and composite analysis. The results will undergo a series of statistical test and will be compared against funds and the S&P 500 index. Lastly, I will implement forecasting to gauge how accurately (using mean squared error) the portfolio's future projections are concerning its actual performance in real life. For graphical summarisation, refer to Figure 8.4.

5.1 Data Compilation

For price data, I derive the structured historical stock data of interest, from Yahoo Finance. In particular, daily data is gathered from various noticeable stocks. The historical stock data is provided by the yfinance Python module and comprises information for the stock using the date as an index. Refer to Figure 8.2 for an example of price data being retrieved using the 'yfinance' python module.

5.2 Technical Indicator Generation

5.2.1 Price data - feature extraction and pre-processing

I begin by using the historical price data that include open, close, high, low, and Adjusted Close (Adj Close) to compute simple technical indicators. I use the data to compute simple and exponential moving averages. These moving averages will be used to create various distinct technical indicators operating on the same price data.

Furthermore, I compute more complex technical indicators, measuring momentum, volume, and other such characteristics for the specified stock or indices such as Bollinger Bands are computed for the same periods, rolling maxima and minima, and others classically.

The Relative Strength Index (RSI) is the first of such advanced indicators to be computed and is used to judge whether a stock is being overbought or oversold. Precisely, the RSI measures whether a stock is over or under-priced and measures the degree to which given stock or asset is inappropriately priced and can be thought of as a comparison between recent price changes and the ratio of average gain to loss over the period n. In particular, the RSI is calculated as:

$$RSI_n = 100 - \frac{100}{1 + rs_n}$$
$$rs_n = \frac{average\ gain_n}{average\ loss_n}$$

Where n is the ”look-back” period under consideration, for which the average gain or loss will be computed. Given the nature of the RSI formula, its values can be interpreted as percentages, with a value of 70 or above indicating that the asset is comparatively overbought; hence, sell as the price should decrease soon. Conversely, a value of 30 or below is an indicator that the price is about to increase indicating to buy.

Next is the computation of the Moving Average Convergence/Divergence (MACD) trend-momentum indicator. This indicator will be subject to varying parameters (window size) to manipulate the effect of the moving average on the indicator. The signal line (a denoised version of the MACD) acts as an indicator that the market is over-pricing or under-pricing an asset according to whether its trend has a steep upward slope or a downward one. This offers an alternative to the RSI, which, rather than following a smoothed or denoised trend of the plot, instead examines the changes in daily price as compared to recent gains and losses.

Stochastic oscillator takes the 3-period moving average of %K, to generate a bounded value. The bounded value can, once again, be thought of as a percentage. Here, values over 80% are considered indicators of an overbought stock, whereas values below 20% are the opposite. The Stochastic Oscillator is meant to be used within smaller trading windows, which means it is less dynamic than the other momentum measures.

$$\%K = (\frac{C - L_n}{H_n - L_n}) * 100$$

William's %R is an alternative to the Stochastic Oscillator and thus, behaves very similarly, except that its bounding values are changed from 0 and 100 to -100 and 0. The negative values maintain the meanings of their corresponding positive counterparts. However, the indicator is calculated as follows:

$$W\%R = \frac{Highest\ High_n - Close}{Highest\ High_i - Lowest\ Low_i}$$

Volume Weighted Average Price (VWAP) calculates the average price of an asset in a given time range. VWAP reflects whether or not the entry price or exit price is profitable relative to the average market price at the time. VWAP is calculated by dividing the cumulative average price the asset has traded at throughout the day by the cumulative volume to get an unbounded value.

The indicator is calculated as follows:

$$VWAP = \frac{\sum^t i = 0 v_i p_i}{\sum^t i = 0 v_i}$$

where v_i is the volume, and p_i is the price in the same time period

Bollinger Bands (BB) is a volatility indicator with three bands: an upper, middle, and lower band. When the asset's price touches the lower band, a buy signal is produced, and when the price reaches the upper band, a sell signal is generated. The lower and upper bands correspond to the positive and negative standard deviations from the middle band (the moving average mean of the stock price). The variable factor is the moving average, and it depends on the timeframe of the stock trades.

5.2.2 Training and Testing Split Methodology

I take the data to prepare it for the analysis, which now occurs at different periods between 01/01/2007 to 01/04/2021 and split it into a training and testing period. Having extensive training data ensures the models implemented have enough training examples relative to the testing data. From there, data is divided into features and targets. This will allow me to consecutively train the model on one set of data and then evaluate its performance using another.

5.3 Fundamental Analysis

5.3.1 Methodology

The system rates stocks out of 100 based on their valuation, profitability, growth, and price-performance metrics - their scores relative to other companies' metrics within the same sector. The stocks are rated using a grading system.

Grading System

The grading system acts on the laurels of normal distribution of values for a chosen metric for a specified sector. For each chosen metric, the stock's metric is compared to the same metric of every stock within the sector, and the stock's metric grade is based upon its percentile in the distribution of values. Each metric is classified as a fundamental feature, profitability, leverage, valuation, or operating ratios.

The metrics chosen are: Return on Equity, Return of Assets, Price to Book, Earning per Share (EPS), Debt to Asset, Debt to Equity, Profit after Tax Margin, EBITDA Margin, Return of Capital Expenditure, Price to Earnings (P/E), Total money spent on investments, and Return on Investments.

Interest Coverage Ratio, Fixed Assets Turnover Ratio, Working Capital Turnover Ratio, and Total Assets Turnover Ratio were omitted due to the abundance of missing or inaccessible data needed for their calculations.

The grading system takes the standard deviation of the set of values and divides that number by 3 to acquire the 'Change' value, which outlines grade boundaries. The grades range from A+ to F for each metric. Each grade is associated with a weight and the stock's total weighted value out of 100, translating to a grade (between A+ - F). Visual representation of the grade boundaries - Figure 8.1

5.3.2 Data

Initially, data was collected through the python module, yfinance [17], which gathers its fundamental data from Yahoo finance. I ran into an issue where the fundamental data provided by the yfinance[17] module is minimal as it only dates back to 2015. The final solution acquired its data from a python module called finpie[18]. The finpie[18] python module required some alterations on my side, but it could gather data from macrotrends that has all vital fundamental data required by the program and dates back to 2005.

6 | Results

6.1 Metrics used

The Sharpe ratio (SR) is used to take risk into account. The ratio adjusts for its risk, dividing the expected return on the asset from the market and the risk-free rate of return on the asset by the extra volatility that one endures for performing that strategy. Formula:

$$S = \frac{\mathbb{E}[\rho_t - \rho_F]}{\sqrt{var(\rho_t - \rho_F)}} \tag{1}$$

where ρ_t are periodic returns, and ρ_F is the rate of return of a risk-free asset.

In these experiments, the risk-free asset is based on default-free (government) zero-coupon bonds - valued at 0.36 (2016-2021) and 1.66 (2007-2010).

Although the Sharpe Ratio considers volatility of the portfolio components, there is frailty in its calculation as it equally treats upward and downward movements. In reality, upward volatility contributes to positive returns, but downward volatility contributes to losses. To highlight the magnitude of downwards deviations, Maximum Drawdown (MDD) is also considered. MDD is the most significant loss from a peak to a trough, and its mathematical formula is:

$$MDD = \max_{t>\tau} \frac{p_t - p_\tau}{p_t} \tag{2}$$

Markowitz Mean-Variance analysis is not enough, so the inclusion of a newer risk metric is necessary. Value-at-risk (VaR) is an ideal and prevalent risk metric because it summarises the overall market risk faced by an institution or an individual investor. The VaR summarises the worst loss over a set window with a given level of confidence, in this paper 5%

confidence level. Formula:

$$VaR_\alpha = \min_{\tau \in R} Pr[f(S, y) \leq \tau] \geq \alpha \tag{3}$$

Statistical tests are also implemented. ANOVA stands for analysis of variances. It tests for differences between group means. ANOVA tests for significance by comparing the variance between individuals within groups to variance among groups. Formula:

$$\frac{variance\, among\, groups = group + individual + error}{variance\, within\, groups = individual + error} \tag{4}$$

$$variance_{total} = variance_{within} + variance_{among}$$

T-test, similarly to ANOVA, allows me to compare results in terms of that variation. At a given confidence level, it allows me to see if there is a statistical significance between the two set of results. To see if there's a statistical significance, T-test does a test of means to see if both sets of results truly different or they have different means by chance. Formula:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

6.2 Technical Analysis

6.2.1 In-sample Profitable Rules

For each indicator, the best performing rule is found based on the largest mean profit factor (returns) together with a high Sharpe ratio. Sharpe ratio is an information ratio that provides context on the strategy significance.

The criteria for a high Sharpe ratio is that it should not be less than 80% of the value for the rule with the highest Sharpe ratio. Figure 6.2 below shows mean profit factor (returns) since the start of trading, the Sharpe ratios and max drawdown.

	MACD_12_26	MACD_50_200	William_R	BB_10	BB_21	BB_50	RSI	Stoch_Osc	VWAP	ATR_2	ATR_10	ATR_30	ATR_50
Mean Profit Factor	1.398227	1.449167	1.457457	1.303946	1.311114	1.27133	NaN	1.50792	1.112248	1.527209	1.474286	1.472738	1.407959
Mean Sharpe Ratio	2.307742	2.626183	2.60538	4.538849	3.679397	4.170849	2.945783	2.6487	6.415181	2.773598	2.650075	2.5432	2.577249
Mean Max Drawdown	20.45222	11.083096	10.456071	15.288868	13.894227	9.988619	NaN	10.895522	9.98041	14.414019	15.138879	16.825047	18.923551

Figure 6.1: Technical indicator risk and performance analysis across all sectors (2015–2020, bullish environment)

From the results above, the best performing technical indicators in each

	MACD_12_26	MACD_50_200	William_R	BB_10	BB_21	BB_50	RSI	Stoch_Osc	VWAP	ATR_2	ATR_10	ATR_30	ATR_50
Mean Profit Factor	1.238529	0.899494	0.974266	1.135596	1.060602	0.994894	0.869982	1.046097	1.018708	0.927451	0.895602	0.811133	0.778142
Mean Sharpe Ratio	1.829642	2.09079	1.87398	4.570709	3.163679	2.562306	1.904446	1.34731	5.509929	1.380726	1.326381	1.321832	1.418796
Mean Max Drawdown	26.051532	26.714804	29.20056	13.717544	18.525664	21.161947	27.353097	36.215044	9.138053	39.050442	41.438938	43.743363	42.681416

Figure 6.2: Technical indicator risk and performance analysis across all sectors (2007–2010, bearish environment)

type of economic climate are Bollinger Bands and Volume Weighted Average Price (VWAP). In a bull period, 50-Day Moving Average Bollinger Bands (Group 6) and VWAP (Group 8), respectively in Figure 8.5, outperformed other indicators. Both indicators have the best performance when risk, volatility and profitability are taken into account in various climates compared to the other technical indicators. However, in a bearish climate, the best performing Bollinger Band has a shorter window of 10 - Group 4 in Figure 8.6 - and Group 8 is VWAP which is still one of the best performing indicators.

When ANOVA statistical test is performed on the technical indicators, there is no statistical difference during the bullish period (2016-2021) at a 5% confidence level but in a bearish climate, there is a statistical difference (p-value is 0.99982 which is greater than 0.975) displaying VWAP and the short window bollinger bands superiority compared to other indicators.

6.3 Fundamental Analysis

6.3.1 In-sample Profitable Rules

Traditional fundamental analysis analyses the intrinsic value of a stock or portfolio, and therefore, trading decisions can still be made by looking at the company profile. Many fundamentalists conduct sector analysis, comparing the best companies within a sector.

For each sector, the stocks are compared with other companies in the same sector - the best performing stock is found through a grading system. All of the stocks ranked A- to A+ is placed into a portfolio, and its performance is based on the most significant mean returns together with a high Sharpe ratio.

The grading system shows a fairly consistent trend where highly graded stocks (regarding their fundamentals) have a low max drawdown, a relatively high profit factor, and a relatively more robust Sharpe ratio than the lower graded stocks. On occasion, lower graded stocks achieve a higher profit factor than the higher graded stocks, but there is a trade-off by being more volatile and risky by scoring a higher max drawdown percentage change and lower Sharpe ratio.

The behaviour is consistent in both bull and bear climates. However, when

ANOVA is implemented, there is no statistical difference (at a 5% confidence level) during the bull period (Figure 8.7) as Grade Boundaries A+ to A- (Group 1), B+ to B- (Group 2), and C+ to C- (Group 3) are similar in regards to variation. However, in a bearish climate (Figure 8.8) there is a statistical difference (p-value is 0.99667 which is greater than 0.975) where Group 1 is better than the other groups as it has the smallest standard deviation and standard error (Figure 8.9) - Group 4 (D+ to D-) is seemingly close, but its raw metrics show it is more volatile and riskier.

	A+	A	A-	B+	B	B-	C+	C	C-	D+	D	D-
Mean Sharpe Ratio	2.027853	2.140103	1.987664	1.541906	2.109931	1.565583	2.080657	1.650222	2.088542	1.589264	1.108505	0.965036
Mean Profit Factor	1.867000	1.952000	1.962000	2.291000	1.934000	2.462000	1.948000	2.294000	2.084000	2.340000	1.114000	3.764000
Mean Max Drawdown	6.000000	12.400000	7.200000	9.600000	10.200000	8.800000	7.500000	9.900000	15.000000	15.100000	68.200000	16.000000

	A+ to A-	B+ to B-	C+ to C-	D+ to D-
Mean Sharpe Ratio	2.128952	1.6849	1.890732	1.728989
Mean Profit Factor	1.940000	2.2540	2.119000	1.648000
Mean Max Drawdown	9.000000	9.5000	10.200000	51.100000

Figure 6.3: Fundamentally graded stocks in equal weighted portfolios between 2015-2020 (2015–2020)

	A+	A	A-	B+	B	B-	C+	C	C-	D+	D	D-	A+ to A-	B+ to B-	C+ to C-	D+ to D-
Mean Sharpe Ratio	1.866053	1.887486	2.256292	1.447158	2.187401	1.813221	1.911771	1.780081	1.274542	1.737949	1.480123	1.456529	2.276611	2.00172	1.650073	1.720837
Mean Profit Factor	0.877	1.533	1.242	1.215	1.075	1.408	1.155	1.097	0.857	1.467	1.185	1.942	1.232	1.183	1.031	1.419
Mean Max Drawdown	53.8	31.2	40.6	49.2	49.2	42.8	46.7	55.4	65.5	44.8	54.3	41	41	47.7	56.3	45.5

Figure 6.4: Fundamentally graded stocks in equal weighted portfolios between 2007-2010 (a bear climate due to the financial crash)

6.3.2 Fundamental and Technical Analysis

Implemented Q-Learning and backtesting on a series of fundamentally good stocks to evaluate whether technical analysis helps improve fundamental analysis that currently uses position trading (buy and hold) strategy.

	MACD_12_26	MACD_50_200	William_R	BB_10	BB_21	BB_50	RSI	Stoch_Osc	VWAP	ATR_2	ATR_10	ATR_30	ATR_50
Mean Profit Factor	1.249	1.274271	1.299979	1.146667	1.173729	1.11925	1.345875	1.300208	0.989958	1.267	1.244062	1.240563	1.190333
Mean Sharpe Ratio	2.546646	2.827208	2.791146	5.035771	4.1235	4.78875	3.142167	2.994229	10.653229	3.212875	2.925875	3.082604	2.973979
Mean Max Drawdown	20.108333	13.477083	12.095833	15.0	14.072917	11.11875	7.558333	13.802083	10.3	12.897917	16.920833	18.33125	21.76875

Figure 6.5: A+ to A- graded stocks traded using technical analysis between 2015-2020 (Fear & Greed Index - Greed)

	MACD_12_26	MACD_50_200	William_R	BB_10	BB_21	BB_50	RSI	Stoch_Osc	VWAP	ATR_2	ATR_10	ATR_30	ATR_50
Mean Profit Factor	1.1856	1.263133	1.238733	1.122	1.1558	1.139667	1.326	1.278733	0.972267	1.082	0.9986	1.082133	1.091733
Mean Sharpe Ratio	2.582267	2.936467	2.5258	5.9388	5.159267	5.681	3.2224	2.9818	10.118067	3.083133	3.153467	3.243133	3.357067
Mean Max Drawdown	22.813333	16.393333	18.82	7.526667	9.746667	8.706667	8.1	15.5	10.946667	18.613333	21.78	20.8	24.233333

Figure 6.6: A+ graded stocks traded using technical analysis between 2015-2020 (Fear & Greed Index - Greed)

To evaluate whether fundamental analysis helps improves technical analysis performance, I conducted a t-test that shows there is minimal improvement when the fundamental school of thought is applied to technical analysis. The only significant difference (at a 5% confidence level) is during a bull period; fundamental analysis helps increase profits gained without impeding the pre-existing risk-mitigating qualities of technical analysis. There is no noticeable difference between technical analysis and technical analysis assisted by fundamental analysis in the bear market. Refer to Appendix, where section 8.4 has the tabled results.

Both show that VWAP and Bollinger Bands remain the best technical indicators for technical analysis, which helps minimise risk, reducing profit gained and increasing maximum drawdown (MDD). There is no technical indicator that attributes a high profit factor and a considerate or high Sharpe ratio and a minimal max drawdown. Technical analysis should solely be used to potentially mitigate risk but should not solely be used to decide what stocks to place in their portfolios.

6.4 Comparison between portfolio optimised portfolios with Indexes and Funds

In a generally bull market (2016-2021), all of the portfolios mimic the mutual funds and S&P 500 index movement patterns during the five year span - noticeable decreases in portfolio returns occurred in 2018 and 2020, where the economic climate was a downturn. The Black Litterman and Monte Carlo Model tread closely with the funds and S&P 500 index rate of movement, and they achieved a similar cumulative return as the S&P 500, which acted as a benchmark. Both Black Litterman and Monte Carlo were within +15% to the S&P 500. Whereas the Capital Asset Portfolio Management (CAPM) and the Fama French Models statistically superior results than mutual funds as they displayed a significantly higher rate of movement through a higher mean return yearly and higher Sharpe ratio for the same duration of time. However, mean-variance analysis alone is not enough, so Value at Risk must be considered. Value at Risk shows that funds and the S&P 500 index are better performers as they are more stable than self-made portfolios. Funds and the S&P 500 index have a smaller stake (value) at risk to fluctuations in price than the self-made portfolios.

However, market conditions disproportionately affect the overall performance of self-made portfolios. During a generally bearish market (2007-2010), in the financial crash, the self-made portfolio's Sharpe ratios are noticeably worse than the risk-free assets, scoring negative Sharpe ratios and are comparatively worse than the funds and the S&P 500 index on all risk metrics. Funds and the S&P 500 index have less value at risk, more minor downward deviations in portfolio value (max drawdown) and better Sharpe ratios.

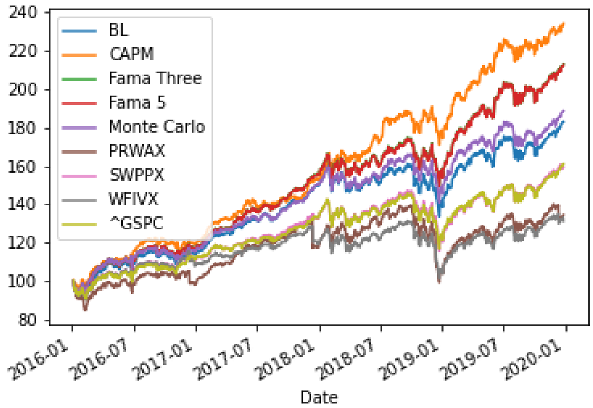


Figure 6.7: Self-made portfolio, funds and indexes' cumulative return comparisons (2015-2020, a bullish climate, Fear & Greed Index - Greed)

5 Year Performance							
Ticker	Mean Return	Max Drawdown	Sharpe Ratio	Cumulative Return	Profit Factor	Value at Risk (VaR) - Monthly	Morning Star Rating
PRWAX	24.16%	30.50%	1.34	66.17%	1.63	5.67%	5 Stars
SWPPX	16.25%	33.80%	1.01	83.27%	1.79	5.35%	4 Stars
WFIVX	15.86%	20.37%	0.97	42.55%	1.19	5.49%	3 Stars
S&P 500	14.74%	20%	0.88	85.18%	2.88	5.53%	-

5 Year Performance, 0.36 Risk Free Rate, Frequency = 1260						
Portfolio Strategy	Mean Return	Max Drawdown	Sharpe Ratio	Cumulative Return	Profit Factor	Value at Risk (VaR) - Monthly
Monte Carlo	14.13%	39.20%	1.02	93.30%	1.18	6.99%
Fama French 3 Factor	24.76%	35.50%	2.07	201.38%	1.29	7.63%
5 Factor	25.18%	34.66%	2.12	206.49%	1.3	7.40%
Black Litterman	14.52%	43.06%	1.01	96.62%	1.17	8.40%
CAPM	31.10%	23.12%	2.99	331.60%	1.36	8.09%

Figure 6.8: Risk comparisons between the S&P 500 index, funds and self made portfolios (2015-2020, a bullish climate, Fear & Greed Index - Greed)

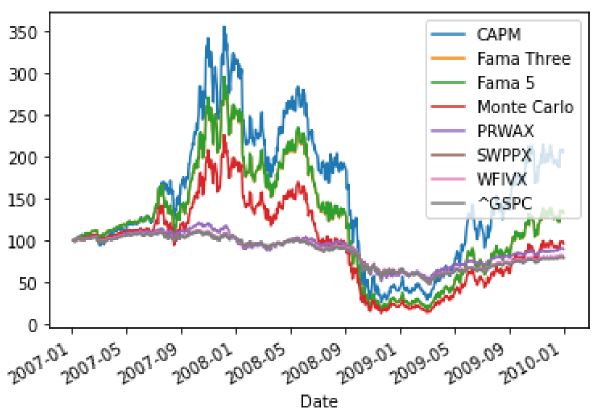


Figure 6.9: Cumulative return comparisons between the S&P 500 index, funds and self made portfolios (2007-2010, a bearish climate, Fear & Greed Index - Fear)

Disclaimer: In the image above, Fama Three and Fama Five performance were closely matched hence orange's lack of clarity.

3 Year Performance (2007 - 2010), 1.66 Risk Free Rate, Frequency = 755						
Portfolio Strategy	Mean Return	Max Drawdown	Sharpe Ratio	Cumulative Return	Profit Factor	Value at Risk (VaR) - Monthly
Monte Carlo	13.55%	60.20%	-2.19	About 128%	1.05	15.33%
Fama French 3 Factor	24.51%	67.10%	-1.52	About 170%	1.1	16%
5 Factor	24.60%	67.10%	-1.52	about 170%	1.1	16.02%
Black Litterman	N/A	N/A	N/A	N/A	N/A	N/A
CAPM	39.80%	48.90%	-0.68	about 252.5%	1.16	17.59%

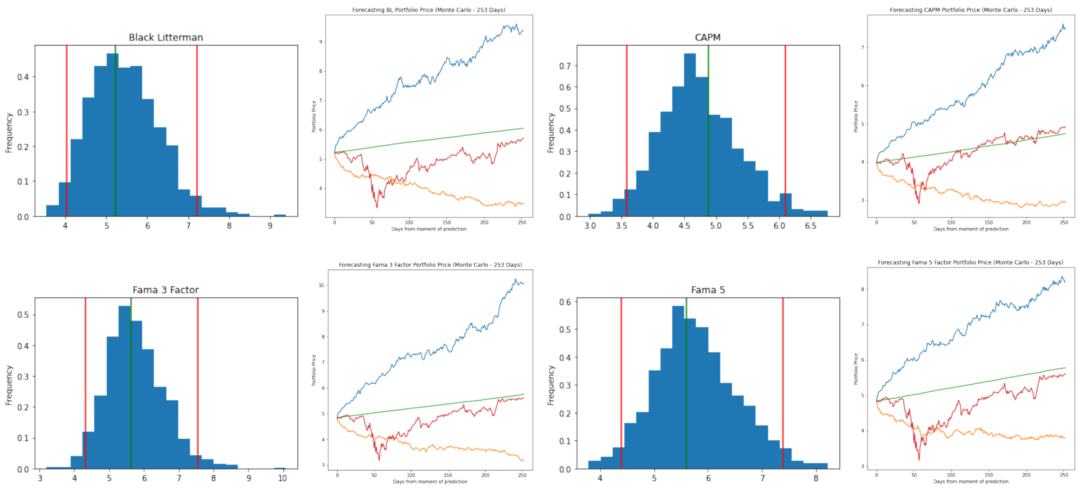
3 Year Performance (2007 - 2010)							
Ticker	Mean Return	Max Drawdown	Sharpe Ratio	Cumulative Return	Profit Factor	Value at Risk (VaR) - Monthly	Morning Star Rating
PRWAX	-%	53.03%	0.02	118.06%	-	14.94%	5 Stars
SWPPX	-%	55.06%	-0.29	162.30%	-	12.68%	4 Stars
WFIVX	-%	55.43%	-0.29	109.33%	-	13.13%	3 Stars
S&P 500	-%	56.78%	-0.38	165.15%	-	13.02%	-

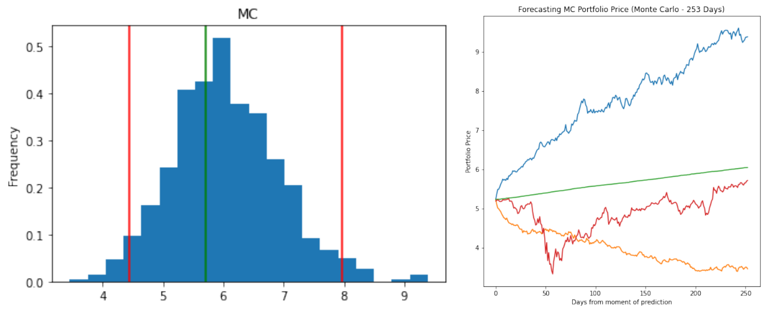
Figure 6.10: Risk comparisons between the S&P 500 index, funds and self made portfolios (2007-2010, a bearish climate, Fear & Greed Index - Fear)

6.4.1 Out-of-sample comparison with real life performance

As well as portfolio optimisation, the Monte Carlo model is useful for the projection of portfolio performance for financial forecasting. The chosen models underwent 1000 simulations to produce a variety of forecasts of the stock's trajectory over a fixed period, an example is Figure 8.3.

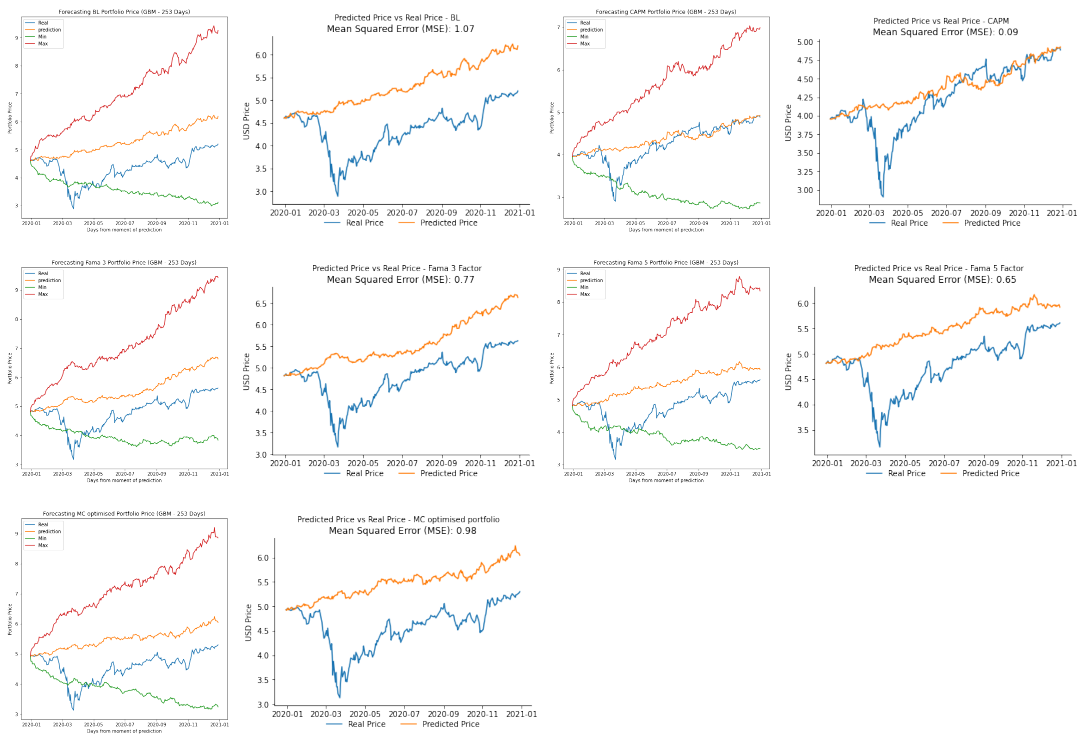
Running the Monte Carlo simulation to project the performance of the portfolio over the course of a year (2020) provided dubious results where a majority of results lay within the 5% confidence level, and the performance tends closely to the Monte Carlo predictions. However, it failed to mitigate the spurious pattern in the market as it was incapable of predicting the sudden downturn in March 2020, where the portfolio price decreased further than the lowest simulated price. The observation is prevalent in all portfolios.





One year projections

The Monte Carlo Model has displayed a reasonable amount of accuracy regarding projecting the portfolio's performance. However, Monte Carlo model's inability to predict the randomness with the market led to the variation of the model - Geometric Brownian Motion (GBM) model. Geometric Brownian Motion is a stochastic process - it is an effective method of gauging volatility and uncertainty with a financial asset. Geometric Brownian motion is useful as it is always positive, similar to stock prices and the standard deviation and volatility depends on the price of the stock.



One year projections

Disclaimer: Black Litterman was omitted from the decade long projections as the model could not be generated due to the staggering absence of views due to the inaccessibility of information beyond 2010. Furthermore, the financial crash period led to a negative risk aversion in the black Litterman model, thus producing a sub-optimal portfolio strategy.

Geometric Brownian Motion is a stochastic process, so it captures the volatility in market movement significantly better than the standard Monte Carlo model. For both projection models, 10000 simulations were conducted.

In the short and mid-term, Geometric Brownian Motion is preferable; however, the volatility of the market minimises the standard Monte Carlo model in the long term. Refer to section 8.2 to see the various Geometric Brownian Motion forecasts for varying time periods.

The trend perseveres in both bull and bear climates. Geometric Brownian motion captures the market's volatility better than the standard Monte Carlo model; refer to sections 8.1 and 8.2 to see the various forecasts at varying time lengths for both the Monte Carlo model and Geometric Brownian Motion.

The model forecasts are slightly better when the look-back period is greater than the forecast's period - makes logical sense as the financial market is a complex adaptive system, so if the simulation models have more data that explicitly capture low-level behaviour and the subtleties of the many entities and the trend in market movement.

6.5 Conclusion

Technical Indicators useful for active trading as the short term technical indicators performs better than the long term. Fundamental analysis is beneficial for position trading like traditional investors as the fundamental has very poor performance in short term analysis.

Technical analysis only requires historical stock data such as price movement and volume. Compared to the fundamental analysis, technical analysis save more research cost and time for the investors.

Technical Indicators useful for active trading as the short term technical indicators performs better than the long term. However, implementing active trading for a prolonged period of time is less profitable than the buy and hold strategy. In the long run, passive is better than active investing. Therefore, fundamental analysis should be used for position trading like traditional investors as the technical analysis has very poor performance in long term analysis. Fundamental analysis for long term investing as the fundamental analysis insight on short term price movement is minimal compared to the technique analysis.

To put this in simple words for traders and investors if they want to maintain their current respective strategies:

1. In a bullish environment, short term traders should use technical analysis on a portfolio full of fundamentally good stocks to increase

their profits without affecting other risk aspects.

2. In a bearish environment, short term traders should use volume weight average price or a bollinger bands with a short window for their chosen technical indicators.

3. Long term investors should opt for a well known mutual funds or should compose their portfolio with fundamentally strong stocks, and time when to add or remove a stock from their portfolio using technical analysis then implement the buy and hold strategy.

Sharpe ratio is a good metric to evaluate a portfolio as it covers performance attribution to tests of market efficiency to risk management, so in theory, the best portfolio optimisation strategy is the Capital Asset Pricing Model due to its high Sharpe ratio and profitability in comparison to the other self-made models. In regards to portfolio theory and its real-world application, I would opt for the Fama French models as their construction is based on the CAPM and solved its defects. Moreover, if the Sharpe ratio is discounted by 50%, the 5-factor Fama French model achieves a Sharpe ratio of 1.06 in a bull market which is comparable to highly rated fund Sharpe ratios in a five-year time span. Beyond having a comparable Sharpe ratio to actual funds, the cumulative returns are better than the mutual funds.

The issue with the Black Litterman model is that the model is not sub-optimal when the risk aversion (delta) is negative, and this is especially likely in a recession, for example, the 2008 financial crash.

During a recession, steer clear of the self-made portfolios as it is preferable to invest in a zero-coupon government bond or invest in a well respected fund, such as the 5 Star MorningStar rated PRWAX. Both options offer more stable and less volatile performance, which is preferable in a climate where investors are fearful and risk-averse.

For portfolio analysis and forecasting, Geometric Brownian Motion should be implemented to gauge the volatility of the assets within the portfolio, and the methodology has the best mean squared error in comparison to the standard Monte Carlo model. However, despite its advantages, Geometric Brownian Motion acts under the random walk assumption, which is beneficial for large scale modelling - beneficial for mid term passive investors but useless for short term investors or day traders. For long term passive investors, Monte Carlo Model is advisable to mitigate any potential unforeseen circumstances as it is extendable to any possible situation whereas Geometric Brownian Motion is only feasible in a market that follows a log-normal distribution which is not a guarantee in the future.

To put this in simple words for traders and investors if they want to improve their current strategies:

1. Short term investor should transition to a long term strategy implementing a buy-and-hold.
2. Investors should primarily use fundamental analysis to assets the valuation and future performance of a stock (or company) rather than technical analysis.
3. Long term investors should opt for a well known mutual funds as conducting fundamental analysis is time exhaustive or they should compose their portfolio with fundamentally strong stocks, and use portfolio strategy to help mitigate risks and improve risk to reward ratio. Fama French 3 or 5 Factor portfolio strategy is the preferred choice due to its commendable Sharpe ratio in comparison to other funds and indexes.
4. To help plan for the future in regards to an investor's net worth and return on investment, investors should forecast portfolio's performance. For mid term portfolio projection, implement Geometric Brownian Motion and for long term to gauge long term portfolio performance use a standard Monte Carlo model.

7 | Evaluation

7.1 Problems remained unsolved

A potential drawback in my solution is the absence of transactions, as its inclusion in methodology could significantly reduce the profitability and effectiveness of trading strategies. Despite some portfolios being exemplary compared to mutual funds and the S&P 500 index, the inclusion of transaction costs might deem the optimised portfolios as sub-optimal when actualised in the real world. Furthermore, the absence of transaction costs concerns active investing more than passive investing and active investing is often attributed with chartists (technical analysis).

Another issue in regards to real-world practicality is the application of the risk-free rate. To gauge a risk-free rated asset is difficult, especially with variable interest rates - it is difficult to determine a fixed risk-free rate over a span of time. Other researchers consider using a 5-year treasury bond, but it is not truly a risk-free asset since the bond coupons will be reinvested at interest rates that cannot be predicted at the moment of prediction. The risk-free rate for a five-year time horizon has to be the expected return on a default-free government five-year zero-coupon bond - this will have a more fixed valuation of the long term risk-free rate.

Another drawback is the inability to gather the market's or analysts' views on a particular stock in the past due to missing data. The issue affects the Black Litterman portfolio as the premise behind the portfolio strategy is Bayesian logic, where the individual (agent) has prior information or views on the stock before initiating any form of optimising analysing. The missing data of the user prior views due to the absence of the market's views on a stock means that the prior views are nothing - the individual is going in blind. A user going in blind is not feasible in theory; however, with the recent surge of individual investors, it is becoming even more likely as amateurs with little to no knowledge of the financial market. For consistency, I did not include the market's and analysts' views on stocks in the present day as I did not want disparities between the Black Litterman models from 5 or 10 years ago and a portfolio generated with 2021 views included.

The drawbacks primarily affected fundamental analysis rather than technical analysis as price data is prevalent and spans for decades, whereas, with fundamental analysis, there is either missing or incorrect data the further you back in history [19].

Whether the perfect information with fundamental analysis would change the findings significantly is debatable, and hopefully, with the advancement of technology and algorithmic trading, that answer would be answered. Current developments such as feedforward search algorithms detect erroneous data by identifying anomalies or Multiple Imputation of by Chained Equations (MICE) to populate and fix missing data [20], but these methods are still under review.

7.2 Future Work

In the future, the property of the prediction in the prediction modules is to be investigated. Moreover, the risk-free rate will not be assumed to be a constant one and will be a varying quantity with time. Finally, to have more information to make a decision, more underlying financial metrics will be considered in my model rather than only the past price data.

The rise of retail investors has increased the popularity of sentimental analysis as many investors are operating under the influence of 'Wisdom of the Crowds', and this spurred an increase in research in the field where Thethi et al. [21] and Kurdonis et al. [22] discovered changes in the public sentiment could affect the stock market. If a variable is capable of causing a change in the stock market, it can be used as a metric to forecast future financial market movement. To ensure extensibility and provide a more holistic perspective to financial forecasting, I would need to delve and include sentimental analysis to evaluate its aspects with predicting a company's profitability and future price movement.

Furthermore, the recent developments in regards to pairs trading research [23], there is a possibility of implementing mean reversion (pairs trading) to help reduce risk in the portfolio optimisation process. There is a rise in research delving into quantitative strategies with fundamental and technical analysis - there is a coined term 'quantamental' for particular investors that use quantitative strategies with fundamental analysis. Pairs trading could potentially reduce the volatility in a portfolio's performance as the quantitative strategy cointegrates and identifies stocks or assets that operate in tandem, so if one experiences a significant decrease in valuation, the associated stock will counteract the drop in price with a similar increase in price, in theory. Hence, the net change in the portfolio is minimal. The potential implementation of pairs trading could help reduce volatility.

Part of future work includes incorporating a lot more stock tickers from various stock exchanges, including NSE, LSE, ASX, etc. The continuation of this paper's objectives in different countries and cultures ensures the extensibility of the paper. Furthermore, it helps me evaluate whether the results from trading strategies and portfolio optimisation and forecasting techniques are universal or help analyse significant differences on a country-by-country basis.

7.2.1 Limitations to the evaluation

Although the prior view is absent, the precision and accuracy of analysts' forecasts is a polarising topic among researchers. There are many doubts regarding the use of analysts' recommendations. Becker and Gürtler [24] argue that using analysts outright target prices and recommendations are that stocks and assets analyses are not available to the extent required for all participants in the market, especially individual investors. Their argument proves true with this dissertation as there are barriers to access to investor's view in the past. One of the most critical findings stems from Womack [25], who discovered that stock prices move in the same way as the analysts' forecasts, showing that analysts' forecasts are self-fulfilling in the short-term and not necessarily to improve public information. Analyst's recommendations being self-fulfilling in the short term, questions their long term viability of the stock, questions the integrity of individual investors and questions what stock qualities attract analysts.

The financial market is a rapidly changing industry, the rise of financial technology companies and brokerages offering commission-free trading platforms have helped reduce overall transaction costs, reducing the cost of trading and conducting business. However, market conditions can negatively affect the process of conducting business as a robust institutional environment can quickly become a weak institutional environment due to cost surges and shutdown of businesses which can exacerbate transaction costs for all participants in the market. The transaction costs are heavily likely to experience fluctuations. The inclusion of transaction costs in future work could make the future work outdated due to the ever-changing nature of the market.

There are discrepancies between the institutions and individual investors. Institutions occasionally cause a rift between them and the individual investors by setting barriers of entry. For example, in recent events, commission-free brokerages locked out customers from trading particular stocks such as Gamestop and AMC during the short squeeze phenomenon of January 2021. These unmitigated changes and restrictions in the institutional environment and behaviour cannot be considered for setting transaction costs and cannot be mitigated by an analysis.

The advancement in technology proves to be very promising to the realm of algorithmic trading by reducing transaction costs with the use of blockchain and significantly faster data processing and manipulation with quantum computing. Despite these fascinating technologies, individual investors are likely to be priced out due to their scarcity and potentially high costs. Institutions will have an overwhelming advantage, causing a disproportionate advantage to funds and the S&P 500 index compared to self-made portfolios.

Other accuracy and risk measures would need to be considered with the future work planned, such as judging the portfolio against market volatility using the VIX as the benchmark and the use of entropic value at risk to judge risk performance. Furthermore, I would need to test over more extended periods of time as mean squared error offers limited information in short timeframes as practically all mean squared error scores are relatively close to zero. Testing over long periods provides more information, thus better information.

7.3 Discussion

The paper focuses on two trading strategies, technical analysis strategies. The primary focus is to compute the best entry and exit points based on price movement and fundamental analysis, where fundamental metrics are constantly being processed and compared to other companies with the same sector and same market. The results showcased that there has been limited change to market behaviour, therefore, fundamental analysis is still relevant. The only difference in the financial market is that investing has become available to wider audience and has become mainstream introducing a new cohort of retail investors. Fundamental analysis outperformed technical analysis in the long run (observed through backtesting) and financial forecasting has provided positive results where portfolio projections closely matched real life performance.

The paper has raised a series of fascinating possibilities, such as developing rigorous models with advanced technology, such as quantum computing, helping the process of data preparation, data manipulation and data formulation in preparation for statistical modelling that is as holistic as possible. Improvements in algorithmic trading are paramount as one of the main issues is missing data and expansive data processing that cannot be dealt with on unspecialised retail hardware.

The results used various vital metrics to judge performance, but how will the study vary once the evaluation methods are changed. How do people measure success and failure? My method centres on accuracy and the disparity between a portfolio's returns compared to other funds and indexes - thus providing a individual investor with the best selection of stocks to boost their net worth. However, a measure of success and failure for other individuals will be the accuracy of the price movement i.e. the use of the brier score metrics to judge how accurate predictions are to the actual world values or for the socially conscious, Environment, Social and Governance (ESG) value and the score is a metric of judging success.

Other Researchers tend to do their research, overseeing all cycles as one that will be a considerable time period i.e. 20 years, whereas this paper explores how each cycle (recession, downturn, expansion and peak) affects the portfolio's performance. Furthermore, I provide further context by associating the portfolio to the mentality of the economy at the time using CNN's Fear and Greed Index. The intention is to provide an insight into how well a portfolio operates in different economic climates and give a more holistic view for a reader when it comes to gauging performance.

Difficulties with the paper are data processing and data collection, where there were many latency issues. However, the future is bright due to blockchain and quantum computing that will allow fetching a lot more data and process data more to make more comprehensive portfolios that will be adaptive enough for all scenarios as it will be able to backtest a lot further into history and be exposed to a wider range of scenarios and market climates. For a well-made model, the model needs to be back-tested a lot further back than the current paper as this covers 15 years (2005 to 2020).

A feature that needs to undergo consideration is shorting stocks to improve the self-made portfolios' performance. Shorting stocks is commonly implemented by hedge funds rather than retail investors; however, incorporating this method of investing will significantly improve the holistic view of investing that this paper offers.

Sharpe ratio is a crucial metric that acts as an estimator where its statistical properties will often depend on the portfolio's investment style being evaluated.

Providing possible interpretations for the results, I know that all self-made portfolios are risk-averse by scoring minimal value at risk scores and strong Sharpe ratio scores; however, these findings are open to interpretations. The issue with the Sharpe ratio is that the performance of more volatile investment strategies, such as the ones I have created in this dissertation, is more difficult to gauge than that of less volatile strategies. Therefore, objectively speaking, results garnered in this dissertation imply that, for example, Sharpe ratios are likely to be more correctly estimated for funds than for the self-made portfolios. Furthermore, the data mining process increases the likelihood of inaccuracies with my generated Sharpe ratios [26]. As a counter-measure, like other researchers, the Sharpe ratios need to be discounted by approximately 50% [26].

Bibliography

[1] R. Sullivan, A. Timmermann, and H. White, “Data-snooping, technical trading rule performance, and the bootstrap,” *The Journal of Finance*, vol. 54, pp. 1647–1691, 1999.

[2] B. Graham and D. Dodd, *Security Analysis*. New York: McGraw-Hill, 1934.

[3] J. J. Murphy, “Technical analysis of the financial markets: A comprehensive guide to trading methods and applications,” vol. 2, 1999.

[4] P. Jorion, *Value at risk*. New York: McGraw-Hill, 2007.

[5] J. Seager and R. Ramakrishnan, “Designing a robust backtesting framework,” 2018.

[6] L. Cao, J. Wang, L. Lin, and C. Zhang, “Agent services-based infrastructure for online assessment of trading strategies,” in *Proceedings. IEEE/WIC/ACM International Conference on Intelligent Agent Technology, 2004. (IAT 2004).*, 2004, pp. 345–348. DOI: `10.1109/IAT.2004.1342967`.

[7] S. Maier-Paape and A. Platen, *Backtest of Trading Systems on Candle Charts*, `https://arxiv.org/abs/1412.5558`, [Online; accessed 04-Mar-2021], 2014.

[8] H. Markowitz, “Portfolio selection,” *The Journal of Finance*, vol. 7, p. 77, 1952.

[9] W. F. Sharpe, “Capital asset prices: A theory of market equilibrium under conditions of risk,” *The Journal of Finance*, 1964.

[10] F. Black and R. Litterman, “Global portfolio optimization,” *Financial Analysts Journal*, vol. 48, pp. 28–43, 1992.

[11] E. F. Fama and K. R. French, “The cross-section of expected stock returns,” 1992.

[12] E. Fama and K. French, “Size and book-to-market factors in earnings and returns,” *The Journal of Finance*, pp. 131–155, 1995.

[13] F. Gökgöz, *The applicability of the three factor model in asset pricing in the Istanbul stock exchange*, https://dergipark.org.tr/en/download/article-file/35993, [Online; accessed 13-Mar-2021], 2008.

[14] P. P. Boyle, "Options: A monte carlo approach," *Journal of Financial Cconomics*, vol. 4, pp. 323–338, 1977.

[15] S. A. Maidanov, "Monte carlo european options pricing implementation using various industry library solutions," *New York City: INTEL*, 2010.

[16] P. Morters and Y. Peres, *Brownian Motion*. Cambridge University Press, 2010, vol. 30.

[17] R. Aroussi, *Yahoo! Finance market data downloader*, https://pypi.org/project/yfinance/, [Online; accessed 1-Apr-2021], 2021.

[18] P. la Cour, *Finpie*, https://pypi.org/project/finpie/, [Online; accessed 1-Apr-2021], 2021.

[19] A. Pagano, D. Perrotta, and S. Arsenis, *Imputation and outlier detection in banking datasets*, https://www.sis-statistica.it/old_upload/contenuti/2013/09/RS12-Imputation-and-outlier-detection-in-banking.pdf, [Online; accessed 27-Mar-2021], 2012.

[20] B. Meghanadh, L. M A, B. Joshi, R. Sathiamoorthy, and M. Kumar, "Imputation of missing values in the fundamental data: Using mice framework," *Journal of Quantitative Economics*, vol. 17, Oct. 2018. DOI: 10.1007/s40953-018-0142-7.

[21] J. K. Thethi, A. Pandit, H. Patel, and V. Shirsath, "Stock market prediction and portfolio management using ml techniques," *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NTASU*, vol. 9, 2021.

[22] J. Kordonis, S. Symeonidis, and A. Arampatzis, "Stock price forecasting via sentiment analysis on twitter," in *20th Pan-Hellenic Conference on Informatics*, 2021.

[23] D. Kazakov, N. Chotai, and C. Erten, "Pair trading with an ontology of sec financial reports," 2020.

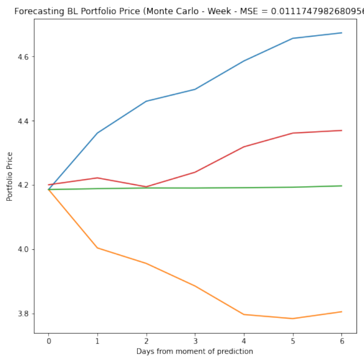
[24] F. Becker and M. Gürtler, "Quantitative forecast model for the application of the black-litterman approach," *SSRN Electronic Journal*, 2010.

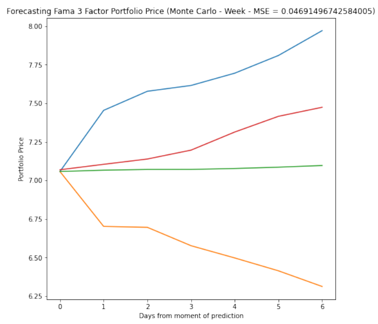
[25] K. Womack, "Do brokerage analysts' recommendations have investment value?" *The Journal of Finance*, vol. 51, pp. 137–167, 1996.

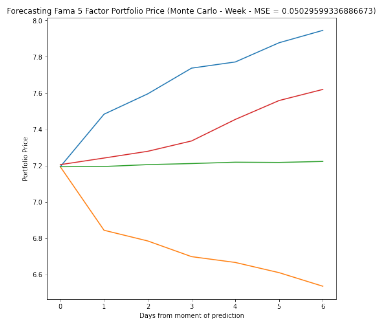
[26] C. Harvey and Y. Liu, "Backtesting," *The Journal of Portfolio Management*, vol. 42, pp. 13–28, 2015.

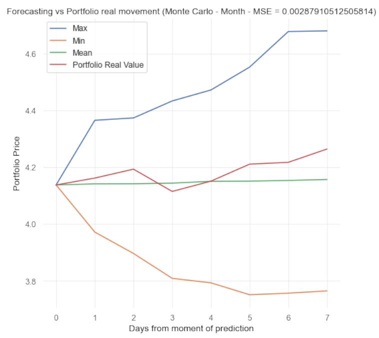
8 | Appendix

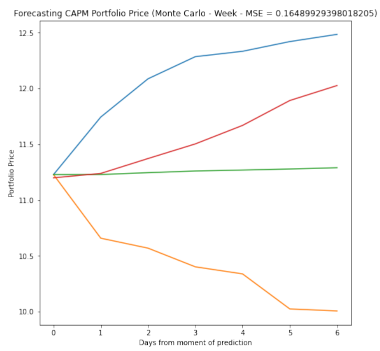
8.1 Monte Carlo forecasts



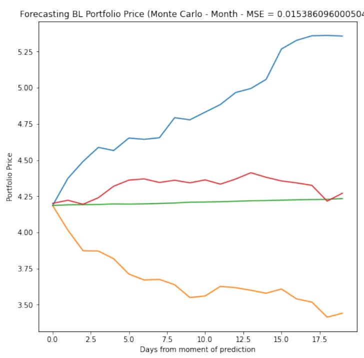


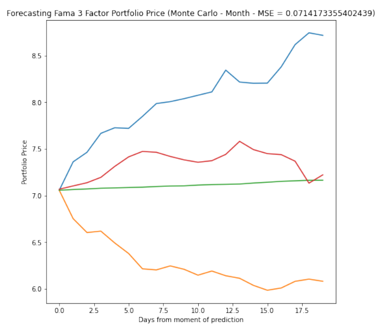


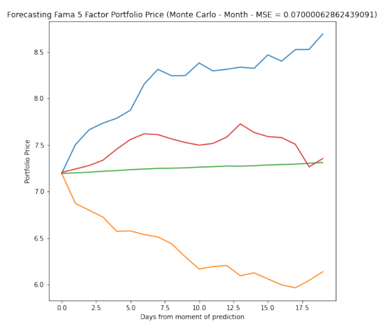


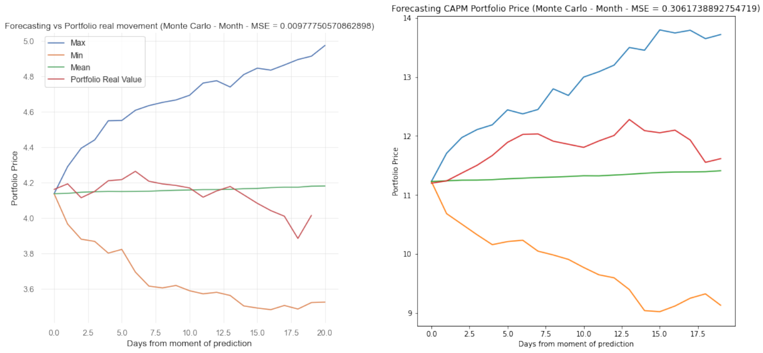


Week projections (4-11 Jan 2021)

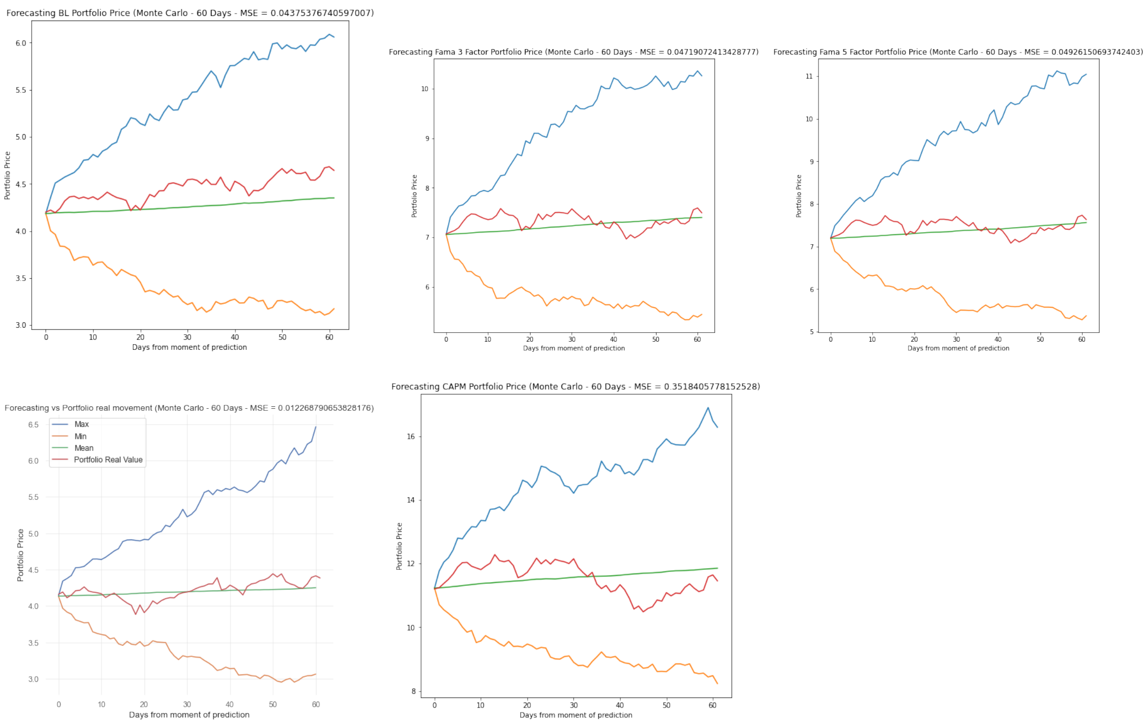




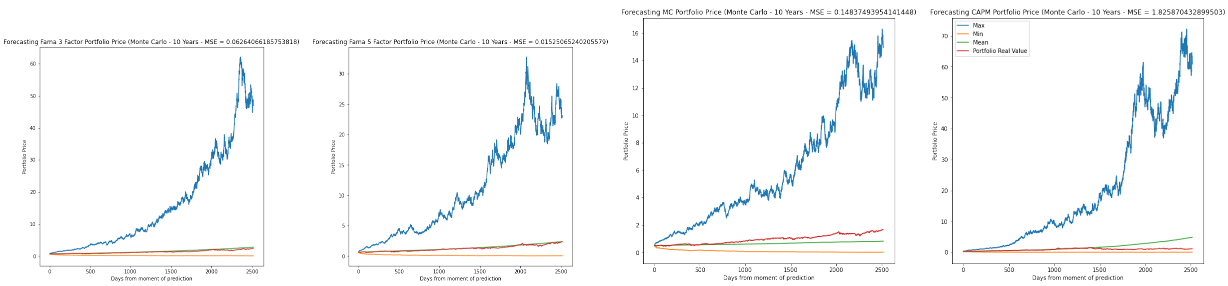




Month projections (Jan 2021)

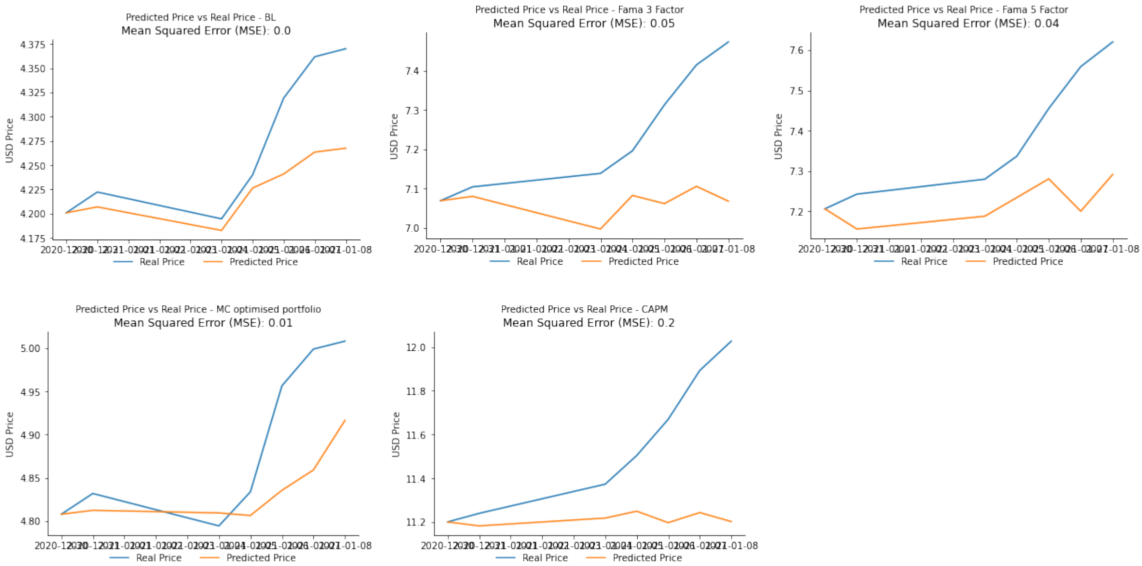


3 Month projections (Jan 2021 - Mar 2021)



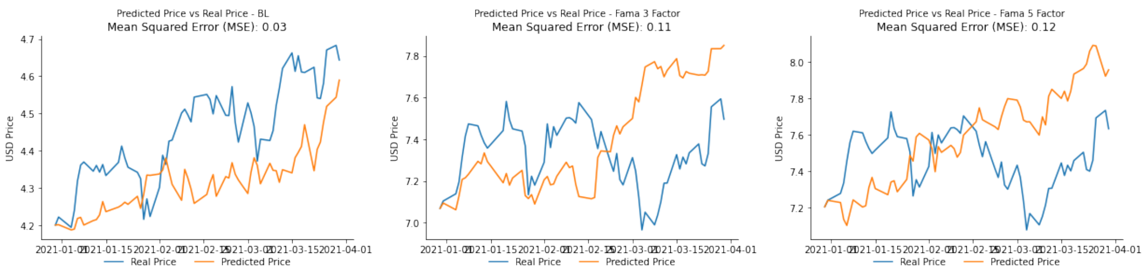
Decade projections (2010 - 2020)

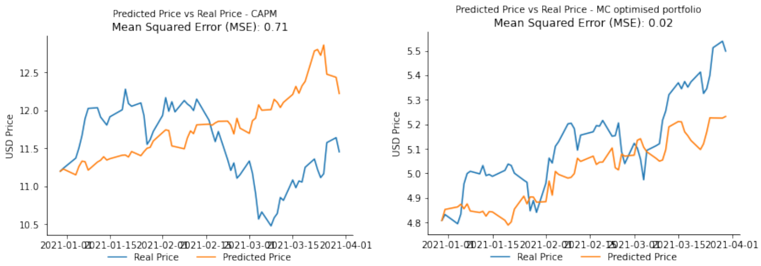
8.2 Geometric Brownian Motion Forecasts



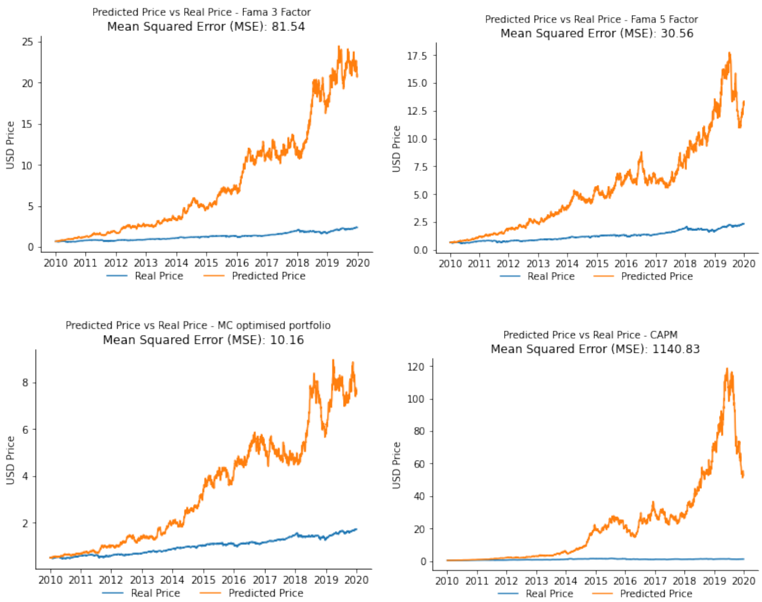
Week projections (4-11 Jan 2021)

Month projections (Jan 2021)





3 Month projections (Jan - Mar 2021)



Decade projection projections (2010 - 2020)

8.3 Code Snippets

```
# Black Litterman

S = risk_models.CovarianceShrinkage(data[assets]).ledoit_wolf()
delta = black_litterman.market_implied_risk_aversion(market_data)
```

```
In [13]:
```

```
mcaps = {}
for t in assets:
    stock = yf.Ticker(t)
    mcaps[t] = stock.info["marketCap"]
```

```
market_prior = black_litterman.market_implied_prior_returns(mcaps, delta, S)

bl = BlackLittermanModel(S, pi="market", market_caps=mcaps, risk_aversion=delta, absolute_v

ret_bl = bl.bl_returns()
S_bl = bl.bl_cov()

ef = EfficientFrontier(ret_bl, S_bl)
ef.add_objective(objective_functions.L2_reg, gamma=0.1)
ef.max_sharpe()
#ef.min_volatility()
#ef.portfolio_performance(verbose=visualize)
BL_Weights = ef.clean_weights()

BL_Weights = pd.Series(BL_Weights)
```

Example of code to generate Black Litterman portfolios

```python
# Time for CAPM
x = mkt_diff.values
x = sm.add_constant(x)

Beta_1  =  []
for i in assets:
    y_i =  port_diff[i].values
    mod =  sm.OLS(y_i, x)
    res =  mod.fit()
    Beta_1.append(res.params[1])

Beta_1 = np.matrix(Beta_1)

# Using scikit Learn
y = port_diff.values
x = mkt_diff.values
lst = []
for i in mkt_diff.values:
    lst.append([i])

x = lst
reg = LinearRegression().fit(x, y)
Beta_2 = np.matrix(reg.coef_).T

# Combining both betas to show them
B1 = pd.DataFrame(Beta_1.T, index = assets, columns = ['Beta_1'])
B2 = pd.DataFrame(Beta_2.T, index = assets, columns = ['Beta_2'])

B1_B2 = pd.concat([B1, B2], axis = 1)


################################################## ######################
# Calculating the vector of means and the matrix of covariances from
# the CAPM model
########### ################################################ ###########

Sigma_rm = np.matrix(np.var(mkt_diff, ddof = 1))

x = mkt_diff.values
x = sm.add_constant(x)

mu_F_1 = []
Sigma_e_1 = []
for i in assets:
    y_i = port_diff[i].values
    mod = sm.OLS(y_i, x)
    res = mod.fit()
    error = np.matrix(y_i) - np.matrix(res.predict(x))
    mu_F_1.append(np.mean(res.predict(x), axis = 0))
    Sigma_e_1.append(np.asscalar(np.var(error.T, ddof = 1, axis = 0)))

mu_F_1 = np.matrix(mu_F_1)
Sigma_e_1 = np.diag(Sigma_e_1)
Sigma_F_1 = Beta_2.T * Sigma_rm *  Beta_2 + Sigma_e_1

x = mkt_diff.values
lst = []
for i in mkt_diff.values:
    lst.append([i])
```

```python
x = lst
error_2 = y - reg.predict(x)

mu_F_2 = np.matrix(np.mean(reg.predict(x), axis = 0))
Sigma_e_2 = np.diag(np.var(error_2, ddof = 1, axis = 0))
Sigma_F_2 = Beta_2.T * Sigma_rm *  Beta_2 + Sigma_e_2

################################################## #####################
# Calculating optimal portfolios with historical estimates
# and CAPM estimates
############# ################################################## #########

# Defining the optimization function
def Max_Sharpe(mu, Sigma):
    w = cv.Variable((mu.shape[1], 1))
    k = cv.Variable(1)
    rf = cv.Parameter(nonneg=True)
    rf.value=0
    #defining the problem, objective function and
    constraints = [(mu - rf)*w == 1,
                   w >= 1e-10,
                   k >= 1e-10,
                   cv.sum(w) == k]
    prob = cv.Problem(cv.Minimize(cv.quad_form(w, Sigma)), constraints)
    #solving the problem
    prob.solve(solver=cv.ECOS)

    return np.matrix(w.value/k.value)
```

```python
mu = np.matrix(np.mean(port_diff.values, axis = 0))
Sigma = np.cov(port_diff.values.T)

w1 = Max_Sharpe(mu, Sigma)
w2 = Max_Sharpe(mu_F_2, Sigma_F_2)


Historical_Weights = pd.DataFrame(w1, index = assets, columns = ['Historical'])
CAPM_Weights = pd.DataFrame(w2, index = assets, columns = ['CAPM'])
```

Example of code to generate Capital Asset Pricing Model portfolios

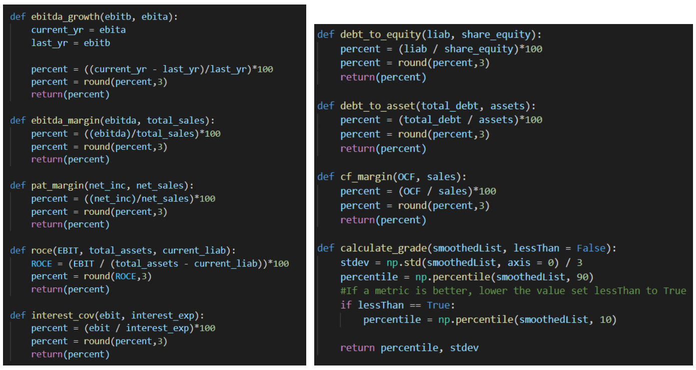
```python
# Fama Three
three_factor = pd.read_csv('FF_three.csv', index_col=0, header=0)/100
three_factor = three_factor.loc[start: end][:-2]
x = three_factor[['Mkt-RF', 'SMB', 'HML']].values
y = port_diff.values
reg = LinearRegression().fit(x, y)
B = np.matrix(reg.coef_).T


Sigma_F = np.cov(x.T)
error = y - reg.predict(x)
Sigma_e = np.diag(np.var(error, ddof = 1, axis = 0))

mu_fama = np.matrix(np.mean(reg.predict(x), axis = 0))
Sigma_fama = B.T * Sigma_F *  B + Sigma_e_2


w_fama_three = Max_Sharpe(mu_fama, Sigma_fama)

F3_Weights = pd.DataFrame(w_fama_three, index = assets, columns = ['Fama Three'])
```

```python
five_factor = pd.read_csv('FF_five.csv', index_col=0, header=0)/100
five_factor = five_factor.loc[start: end][:-2]
x = five_factor[['Mkt-RF', 'SMB', 'HML', 'RMW','CMA']].values
y = port_diff.values
reg = LinearRegression().fit(x, y)
B = np.matrix(reg.coef_).T

Sigma_F = np.cov(x.T)
error = y - reg.predict(x)
Sigma_e = np.diag(np.var(error, ddof = 1, axis = 0))
mu_fama = np.matrix(np.mean(reg.predict(x), axis = 0))
Sigma_fama = B.T * Sigma_F *  B + Sigma_e_2
w_fama_five = Max_Sharpe(mu_fama, Sigma_fama)


F5_Weights = pd.DataFrame(w_fama_five, index = assets, columns = ['Fama Five'])
```

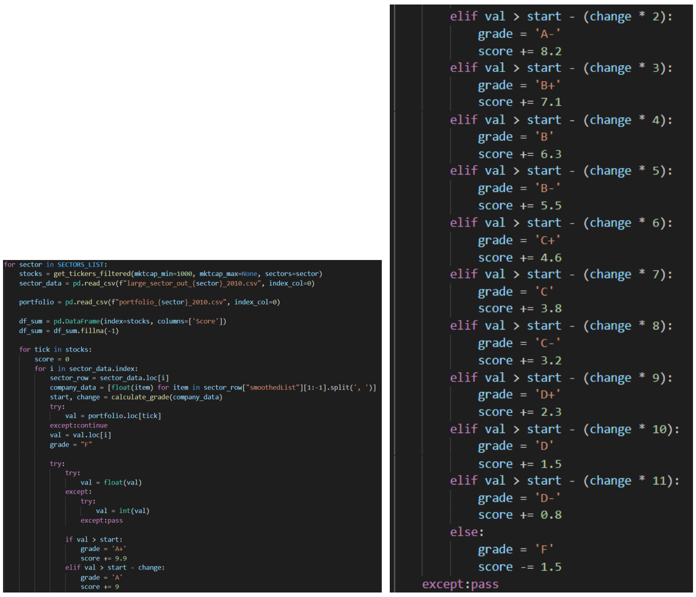
Example of code to generate Fama French 3 and 5 Factor portfolios

```python
def getAverage(df, lessThan = False):
    data_return = []
    for col in df:
        if col == None:
            continue
        metrics = df[col]
        try:metrics = metrics[metrics != None]
        except:pass
        try:metrics = metrics[metrics != '-']
        except:pass
        try:metrics = metrics[~np.isnan(metrics)].tolist()
        except:pass
        elements = np.array(metrics)
        try:
            mean = np.mean(elements, axis=0)
            sd = np.std(elements, axis=0)

            for i in range(2):
                smoothedList = [x for x in metrics if (x > mean - 2 * sd)]
                smoothedList = [x for x in smoothedList if (x < mean + 2 * sd)]

                mean = np.mean(smoothedList, axis=0)
                sd = np.std(smoothedList, axis=0)

            stdev = np.std(smoothedList, axis = 0) / 3
            start = np.percentile(smoothedList, 90)
            smoothedAvg =  np.percentile(smoothedList, 50)

            if lessThan == True:
                start = np.percentile(smoothedList, 10)

            data_return.append([col, smoothedAvg, start, stdev, smoothedList])
        except:pass

    return data_return
```

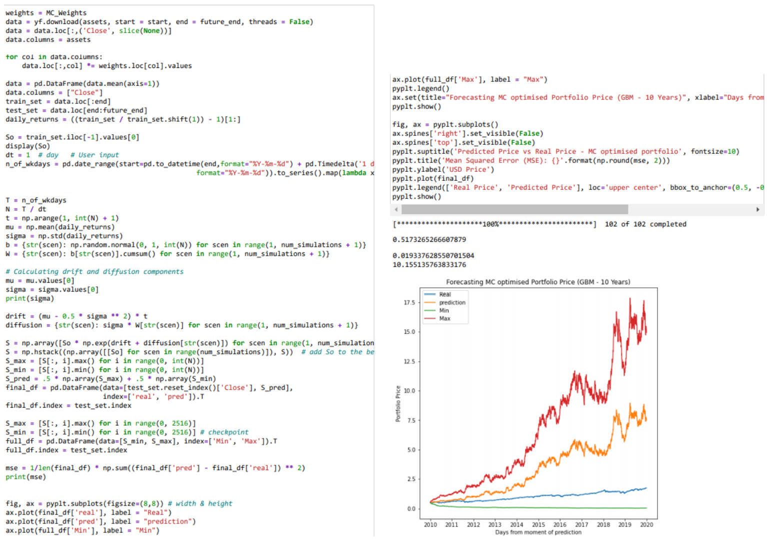
Example of code to generate grading system’s grade boundaries



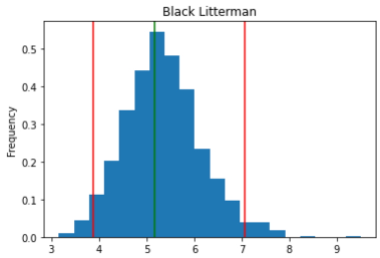
Example of code to generate fundamental metrics



Example of code applying weights to certain grades



Example of code producing a geometric brownian motion forecast of a portfolio



Example of code showing how the forecast fell within the boundaries of 5% Confidence levels

```python
weights = BL_Weights
data = yf.download(list(weights.index), start = start, end = end, threads = False)
data = data.loc[:,('Close', slice(None))]
data.columns = list(weights.index)

for col in data.columns:
    data.loc[:,col] *= weights.loc[col]

data = pd.DataFrame(data.mean(axis=1))
data.columns = ["Close"]
port_diff = pd.DataFrame(data.pct_change().dropna())
port_diff.columns = ["Close"]
avg_daily_return = port_diff.mean()["Close"]
std_dev_daily_return = port_diff.std()["Close"]
# Set number of trading days and get last closing price of TSLA from DataFrame
stock_last_price = data['Close'][-1]
# Initialize empty DataFrame to hold simulated prices for each simulation
simulated_price_df = pd.DataFrame()

# Run the simulation of projecting stock prices for the next trading year, `1000` times
for n in range(num_simulations):
    simulated_ticker_prices = [stock_last_price]

    for i in range(num_trading_days):
        simulated_price = simulated_ticker_prices[-1] * (1 + np.random.normal(avg_daily_ret
        simulated_ticker_prices.append(simulated_price)

    simulated_price_df[f"Simulation {n+1}"] = pd.Series(simulated_ticker_prices)


simulated_ending_prices = simulated_price_df.iloc[-1, :]
simulated_ending_prices.value_counts(bins=20) / len(simulated_ending_prices)
confidence_interval = simulated_ending_prices.quantile(q=[0.025, 0.975])

data = yf.download(list(weights.index), start = end, end = future_end, threads = False)
data = data.loc[:,('Close', slice(None))]
data.columns = list(weights.index)

for col in data.columns:
    data.loc[:,col] *= weights.loc[col]

data = data.mean(axis=1)
port_diff = pd.DataFrame(data.pct_change().dropna())
actual_price = data.iloc[-1]

simulated_ending_prices.plot(kind='hist', density=True, bins=20)
pyplot.title("Black Litterman")
pyplot.axvline(confidence_interval.iloc[0], color='r')
pyplot.axvline(confidence_interval.iloc[1], color='r')
pyplot.axvline(actual_price, color='g')
```

How Monte Carlo simulations were created

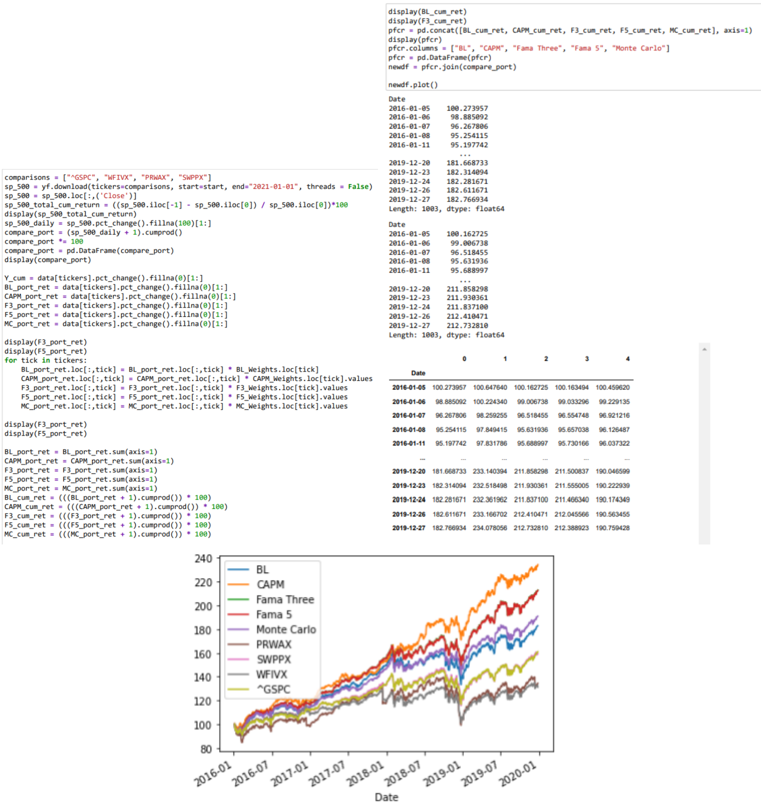
```python
try:
    eps = income_df["eps__earnings_per_share"]
    if eps == None:
        raise Exception
except:
    try:
        eps = income_df["basic_eps"]
        if eps == None:
            raise Exception
    except:eps = '-'

try:
    ROA = ratios_df["roa__return_on_assets"]
    if ROA == None:
        raise Exception
except:ROA = '-'

try:
    ROI = ratios_df["roi__return_on_investment"]
    if ROI == None:
        raise Exception
except:ROI = '-'


try:
    roe = ratios_df["roe__return_on_equity"]
    if roe == None:
        raise Exception
except:roe = '-'

try:
    investments = balance_sheet["longterm_investments"]
    if investments == None:
        raise Exception
except:
    investments = '-'
```

```python
try:
    OCF = income_df["cash_flow_from_operating_activities"]
    if OCF == None:
        raise Exception
except:OCF = '-'


try:
    EBIT = income_df["ebitda"]
    if EBIT == None:
        raise Exception
except:EBIT = '-'


try:
    gross_profit = income_df["gross_profit"]
    if gross_profit == None:
        raise Exception
except:gross_profit = '-'


try:
    revenue_cost = income_df['cost_of_goods_sold']
    if revenue_cost == None:
        raise Exception
except:revenue_cost='-'


try:
    total_revenue = income_df["revenue"]
    if total_revenue == None:
        raise Exception
except:
    total_revenue = '-'
```

```python
for sector in SECTORS_LIST:
    stocks = get_tickers_filtered(mktcap_min=1000, mktcap_max=None, sectors=sector)
    print(sector)
    columns = ['ROE', 'ROA', 'Price-to-Book', 'eps', 'Debt_Asset', 'Debt_Equity','PAT_Margin', 'EBITDA_Margin', 'ROCE', 'Investments', 'ROI']
    df = pd.DataFrame(index=stocks, columns=columns)
    df = df.fillna(-1)

    for ticker in stocks:
        try:
            balance_sheet = pd.read_csv(f"all_stocks\{ticker}_balancesheet_2005_2020.csv", index_col="date").loc[2015]
            if len(balance_sheet) > 1:
                balance_sheet = balance_sheet.iloc[0].combine_first(balance_sheet.iloc[1])
            if balance_sheet == None:
                raise Exception
        except:pass

        try:
            ratios_df = pd.read_csv(f"all_stocks\{ticker}_ratios_2005_2020.csv", index_col="date").loc[2015]
            if len(ratios_df) > 1:
                ratios_df = ratios_df.iloc[0].combine_first(ratios_df.iloc[1])
            if ratios_df == None:
                raise Exception
        except:pass

        try:
            income_df = pd.read_csv(f"all_stocks\{ticker}_income_2005_2020.csv", index_col="date").loc[2015]
            if len(income_df) > 1:
                income_df = income_df.iloc[0].combine_first(income_df.iloc[1])
            if income_df == None:
                    raise Exception
        except:pass
```

```python
try:
    cashflow_df = pd.read_csv(f"all_stocks\{ticker}_cashflow_2005_2020.csv", index_col="date").loc[2015]
    if len(cashflow_df) > 1:
        cashflow_df = cashflow_df.iloc[0].combine_first(cashflow_df.iloc[1])
    if cashflow_df == None:
        raise Exception
except:pass

#########################################################

try:
    total_assets = balance_sheet["total_assets"]
    if total_assets == None:
        raise Exception
except:total_assets = '-'


try:
    total_current_liab = balance_sheet['total_current_liabilities']
    if total_current_liab == None:
        raise Exception
except:total_current_liab = '-'

try:
    price_to_book = ratios_df['book_value_per_share']
    if price_to_book == None:
        raise Exception
except:price_to_book = '-'
```

Example of code collecting fundamental data about companies



Example of code retrieving funds, the S&P 500 index and self made portfolio performance

8.4 T-test - Technical against Technical with Fundamental

Conditions	T-value	p-value
Bull Climate	1. Profit factor = 5.31926 2. Sharpe Ratio = -1.22101 3. Max Drawdown (MDD) = -0.87099	1. Profit factor = 0.000021 2. Sharpe Ratio = 0.234453 3. Max Drawdown (MDD) = 0.392761
Bear Climate	1. Profit factor = -0.02574 2. Sharpe Ratio = 0.02887 3. Max Drawdown (MDD) = 0.01229.	1. Profit factor = s 0.979679 2. Sharpe Ratio = 0.488602 3. Max Drawdown (MDD) = 0.495147

8.5 Figures

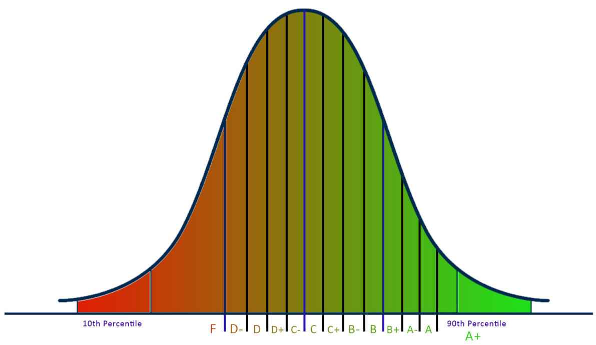


Figure 8.1: Visual representation of grading system

Date	Open	High	Low	Close	Adj Close	Volume
1999-01-22	1.750000	1.953125	1.552083	1.640625	1.509998	67867200.0
1999-01-25	1.770833	1.833333	1.640625	1.812500	1.668188	12762000.0
1999-01-26	1.833333	1.869792	1.645833	1.671875	1.538759	8580000.0
1999-01-27	1.677083	1.718750	1.583333	1.666667	1.533965	6109200.0
1999-01-28	1.666667	1.677083	1.651042	1.661458	1.529172	5688000.0
...	...	...	...	...	...	...
2020-05-13	316.700012	323.140015	303.790009	311.200012	311.200012	15646300.0
2020-05-14	313.670013	321.440002	307.500000	321.220001	321.220001	15057800.0
2020-05-15	315.589996	340.019989	314.959991	339.630005	339.630005	24691500.0
2020-05-18	350.420013	356.660004	347.220001	350.010010	350.010010	19410100.0
2020-05-19	351.609985	363.500000	350.510010	352.220001	352.220001	17882700.0

Figure 8.2: Nvidia (NVDA) price data retrieved using yfinance

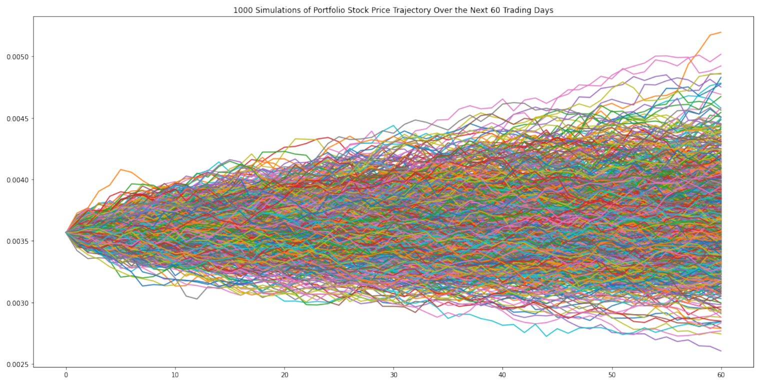


Figure 8.3: Example of all of the possible formulations formed for a forecast

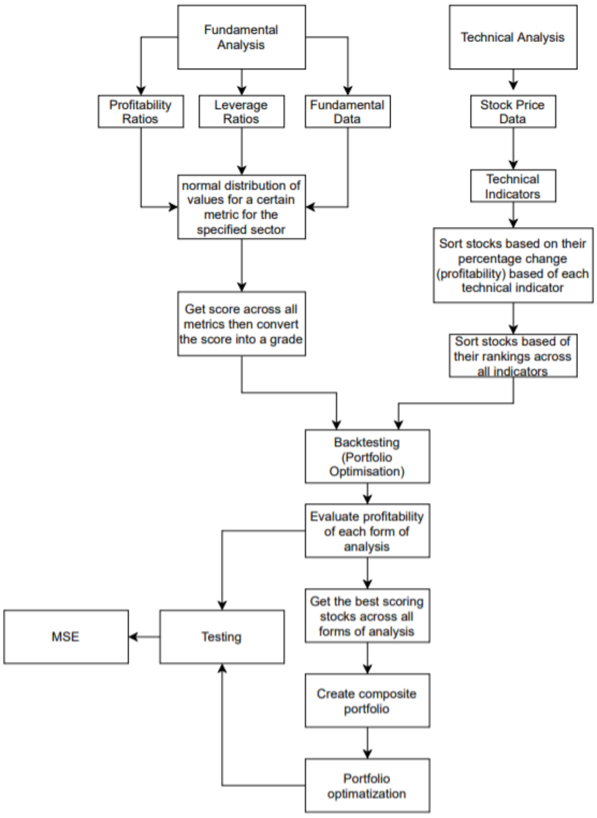


Figure 8.4: Methodology Flowchart

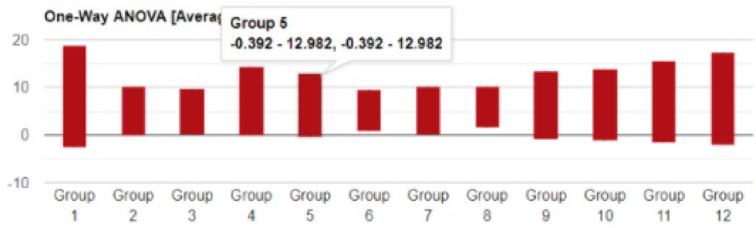


Figure 8.5: Bullish Climate (2015-2020), F-Statistic = 0.11508, p-value = 0.99982
Disclaimer: RSI is omitted

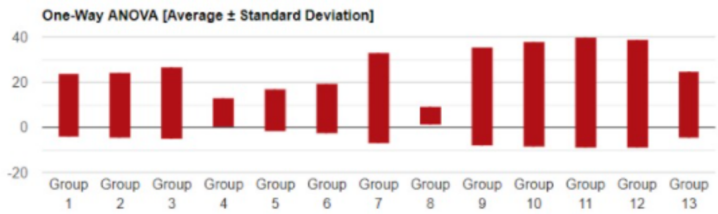


Figure 8.6: Bearish Climate (2007-2010), F-Statistic = 0.5946, p-value = 0.81463

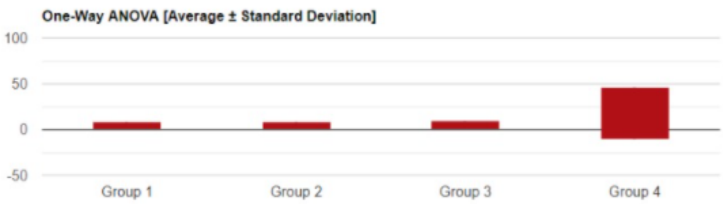


Figure 8.7: Bullish Climate (2015-2020), F-Statistic = 0.64038, p-value = 0.61011

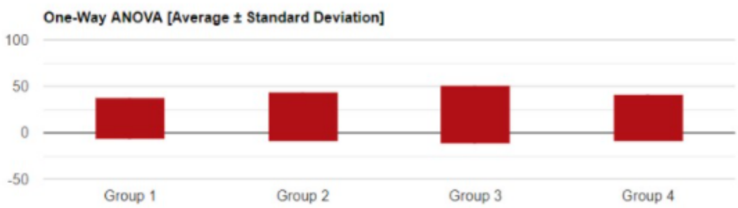


Figure 8.8: Bearish Climate (2007-2010), F-Statistic = 0.0172, p-value = 0.99667

Data Summary				
Groups	N	Mean	Std. Dev.	Std. Error
Group 1	3	14.8362	22.6645	13.0854
Group 2	3	16.9616	26.6234	15.371
Group 3	3	19.6604	31.7324	18.3207
Group 4	3	16.2133	25.3635	14.6436

Figure 8.9: Bearish Climate metrics (2007-2010)

9 | Glossary

	Definition
Active Investing	Investment strategy where investors engages in ongoing buying and selling - often a in short term horizons.
Adjusted (Adj) Close	Closing price is adjusted using appropriate split and dividend multipliers, adhering to Center for Research in Security Prices (CRSP) standards.
Backtesting	The process of testing a strategy, model etc. using historical data.
B/M factor	Refers to the outperformance of value stocks over growth stocks.
ESG Value	ESG stands for Environmental, Social and Governance, and refers to the three key factors when measuring the sustainability and ethical impact of an investment in a business or company. The ESG Value ranks between 0 and 31.
Forecasting	Best guesses about the future. The forecasts economists make are often badly wrong - sometimes they are likened to a dart throwing chimpanzee. Some of the inaccuracies in forecasts reflect badly designed models; often, the problem is that the future actually is unpredictable.
Law of large numbers	The more times you run a simulation or experiment, your overall result will converge closer to the expected value. Large deviations from the expected value become so improbable that the likelihood of reaching those values become infinitesimally small.
MorningStar Rating	MorningStar is an American firm that conducts investment research and management. Ratings conducts ranking given to publicly traded mutual funds which evaluates the people involved, the risk aspects, the portfolio's past performance, fund's process, and the institution managing the fund.
Moving Average	Calculation to analyse data points within a set window (fixed sized subsets of the full data set) creating a series of averages.
Mutual Funds	An investment fund of pooled money from many investors managed by an asset management firm.
Passive Investing	Investment strategy where investors invest by holding assets for longer terms - often using a buy and hold strategy.

	Definition
Retail Investor	Individual investors that take full control of their economic situation and economic future by trading by themselves and doing independent market research.
Sentimental Analysis	Subset of Natural Language Processing field where subjective views are extracted and used to determine an asset or company's future projections.
Shorting	Selling an asset, such as a share, that you do not currently own, in the expectation that its price will fall by a fixed period. If the price does fall, you can buy the asset at the lower price and make a profit. The risk is that if the price rises, it will result in a loss.
Size Factor (SMB)	Measurement to gauge the verified phenomenon that mid- and small-cap stocks outperforming large-cap stocks.
Stochastic Process	A process that exhibits random behaviour. For example, Brownian motion is a stochastic process and it is often used to describe changes in share prices in an efficient market.
Sustainable Growth	A term much used by environmentalists, meaning economic growth that can continue in the long term as it contributes minimal damage to the environment as its supply or use does not use non-renewable resources. Mainstream and television economists use the term to describe a rate of growth that an economy can sustain indefinitely without causing a rise in inflation.
Technical Analysis	Trading discipline of valuing assets' future price movement and its profitability based on the examination of its past price movements.
Quantamental	Combination of Quantitative Analysis for screening and ranking purposes and a Fundamental Analysis in determining the intrinsic value.