ECMT2130: Semester 2 2020 Assignment

SID: 500494994

SID: 500494994
UOS: ECMT2130

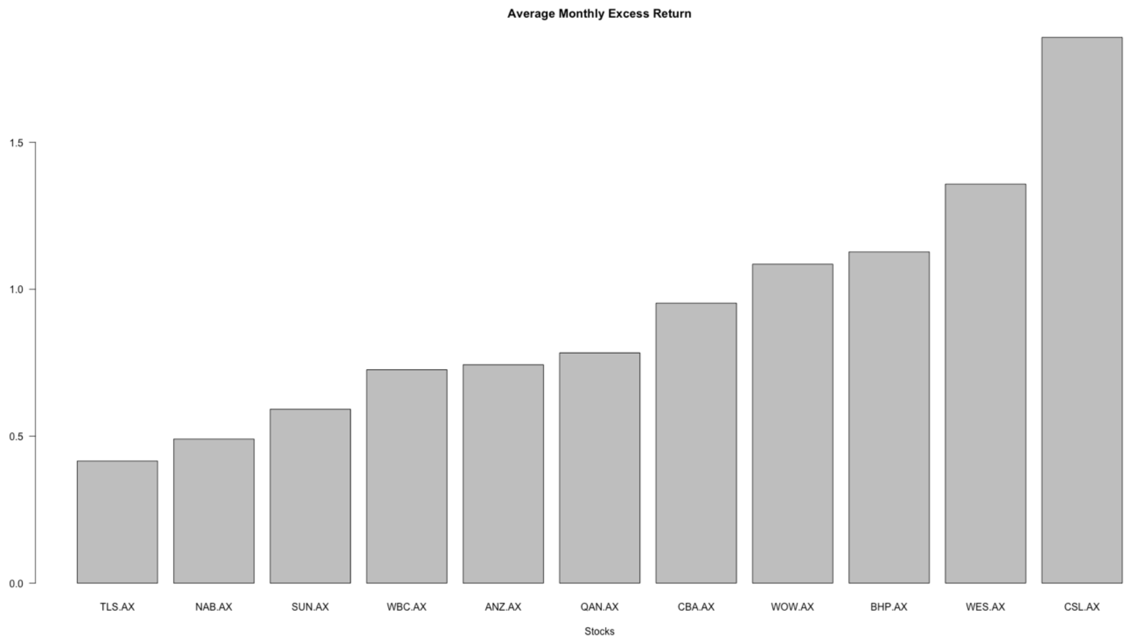
Suitability of the input data

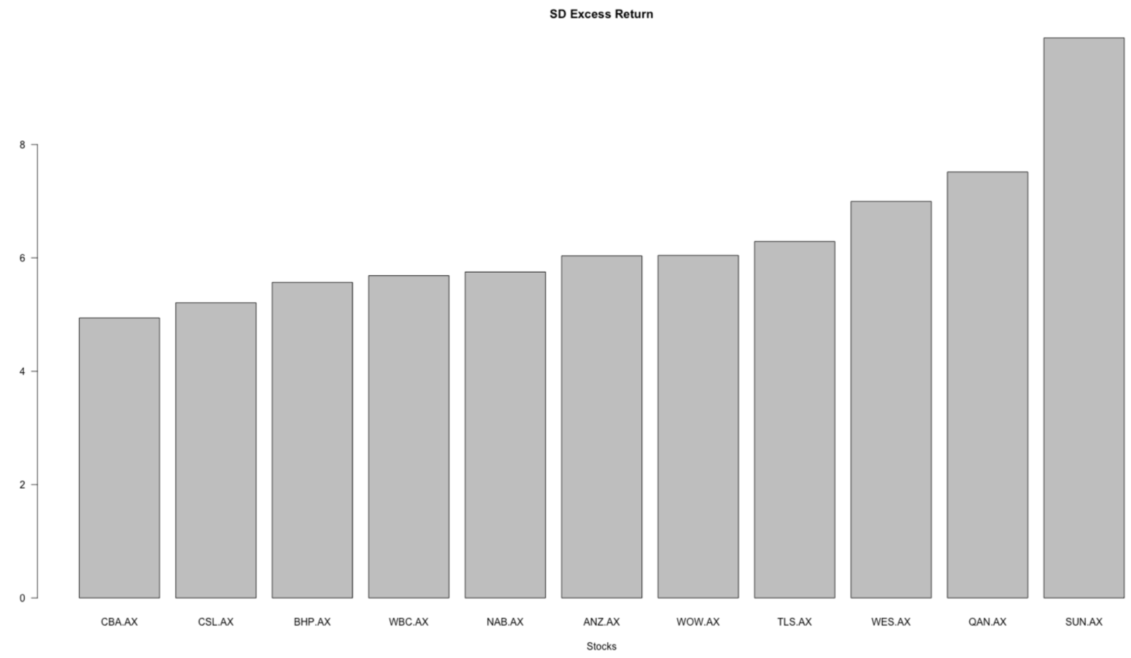
Stocks allowed to be invested

CBA, CSL, BHP, WBC, NAB, ANZ, WOW, TLS, WES, QAN, SUN

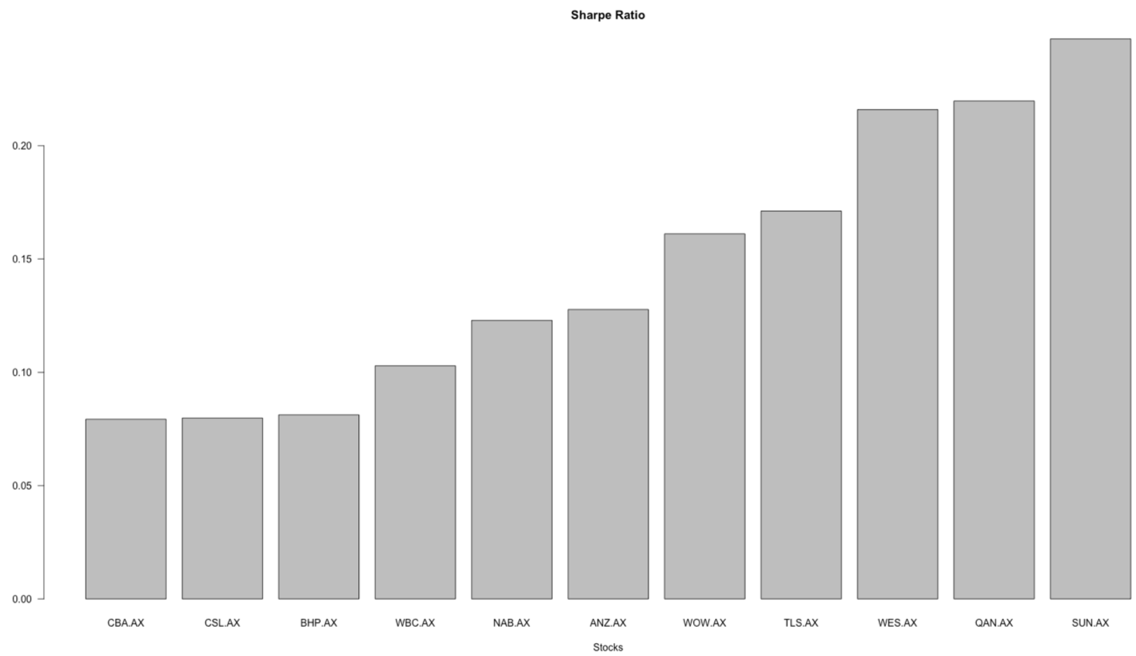
Flaws in Yahoo Finance and Reserve bank of Australia

Data sourced from Yahoo Finance have a 20-minute delay when referring to quotes for stocks during market hours, which could possibly affect the data as it is not completely up to date. In addition, Yahoo Finance has an optional premium package which includes deeper analytics. The Reserve Bank of Australia is a government run banking service, which is privately owned by the Commonwealth of Australia. Due to this, it is considered trustworthy and reliable.



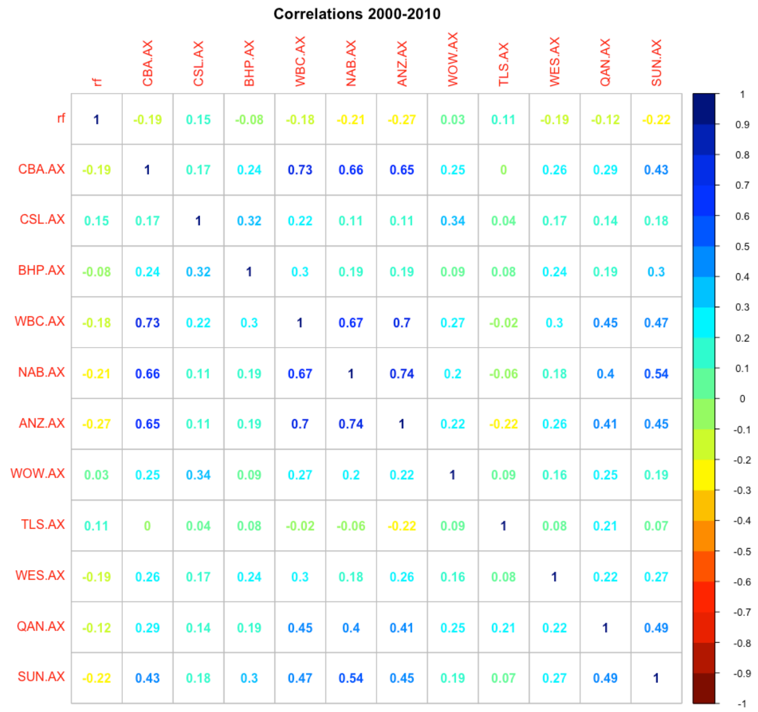


SID: 500494994
UOS: ECMT2130



Stocks with Higher expected return, lower risk-free rate and higher standard deviation provide a better Sharpe ratio.

Stock Correlation

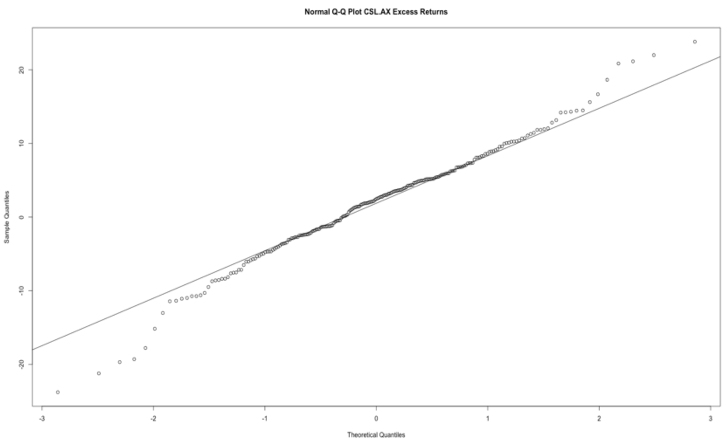


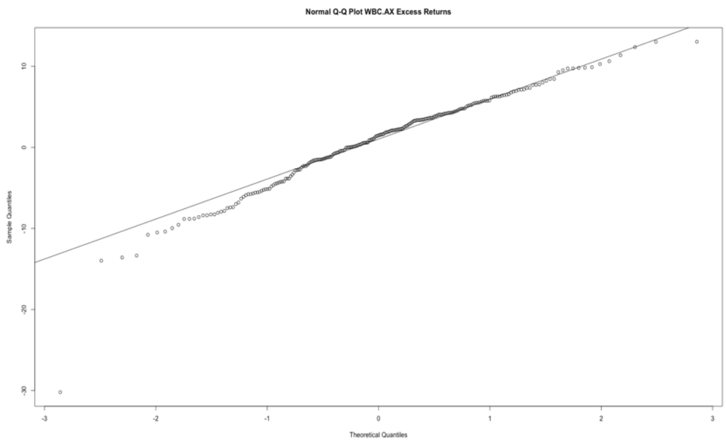
Distinct features in this graph shows the close correlation by sector, namely the financial sector. Companies such as Westpac, Commonwealth Bank, NAB, and ANZ all have a correlation around 0.8. Due to this, in order to invest in a diverse portfolio, investors should not just buy stocks with high correlation as if one has a negative return, the second has a high chance of following the same trend. Risk averse investors can invest in low correlated stocks such as Telstra and CBA to decrease risk as neither stock will affect the other.

SID: 500494994
UOS: ECMT2130

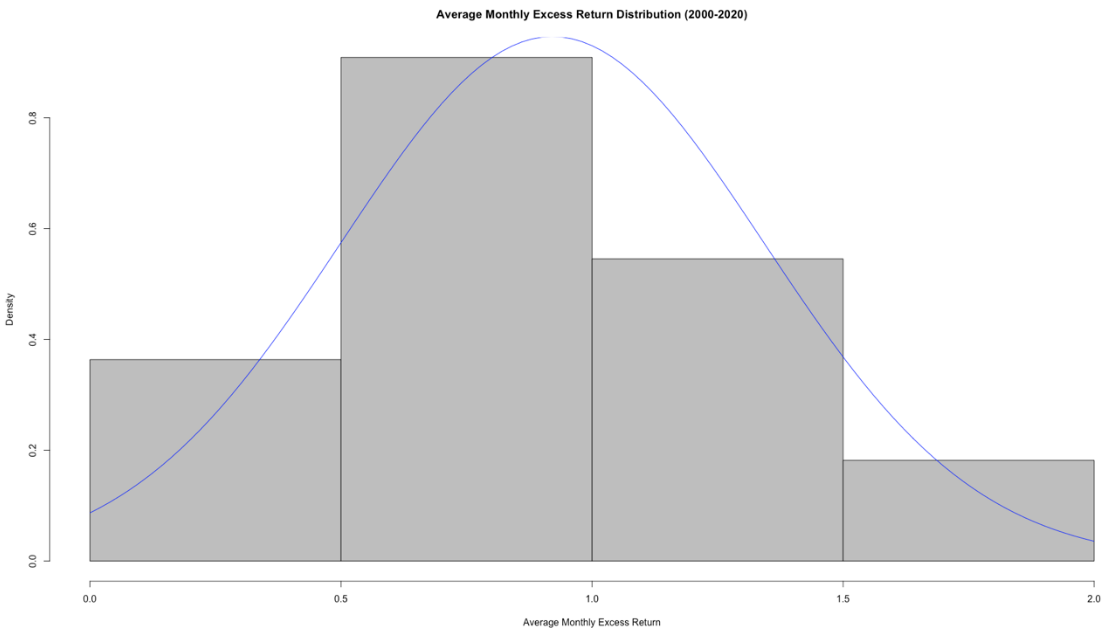
Appropriateness of Data

Dataset is from 1/11/2000 to 30/6/2020, with entries of ~7300 per stock. Investors should be aware of the possibility that historical data be less relevant due to the changes in nature of various factors including global socioeconomic factors. Contextual information is also important in order to understand these datasets. Example includes the Global Financial Crisis in 2007 which lowered number of share prices throughout all sectors. CSL is shown to be the most normal, whilst WBC the least.



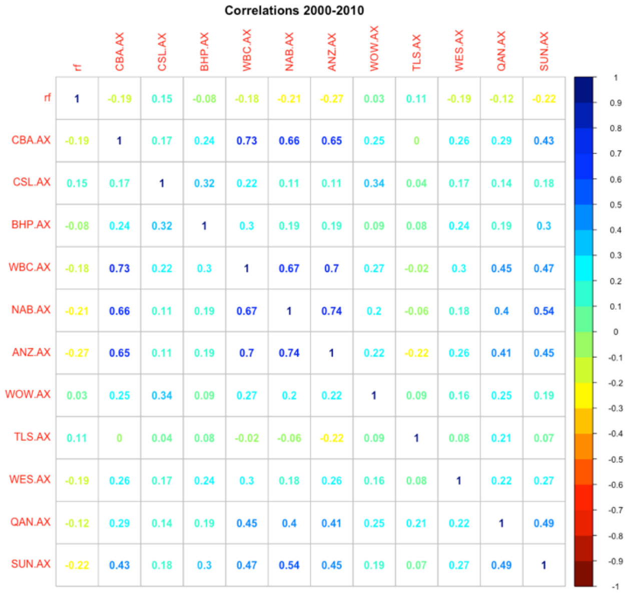


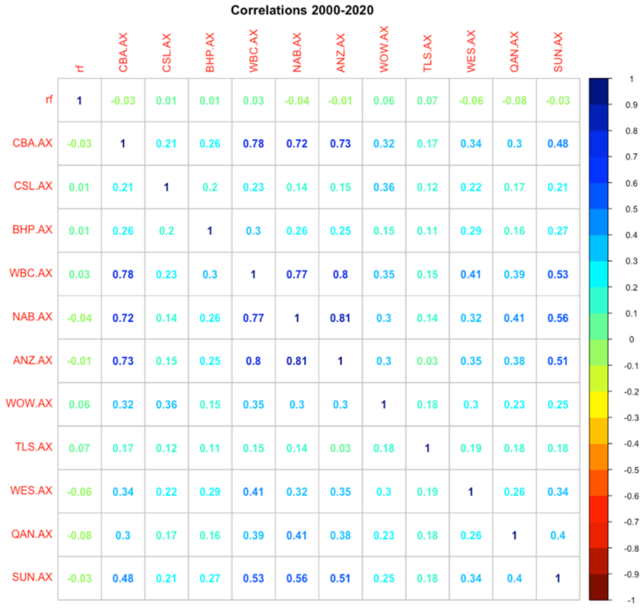
Stability over time of the distribution of excess returns

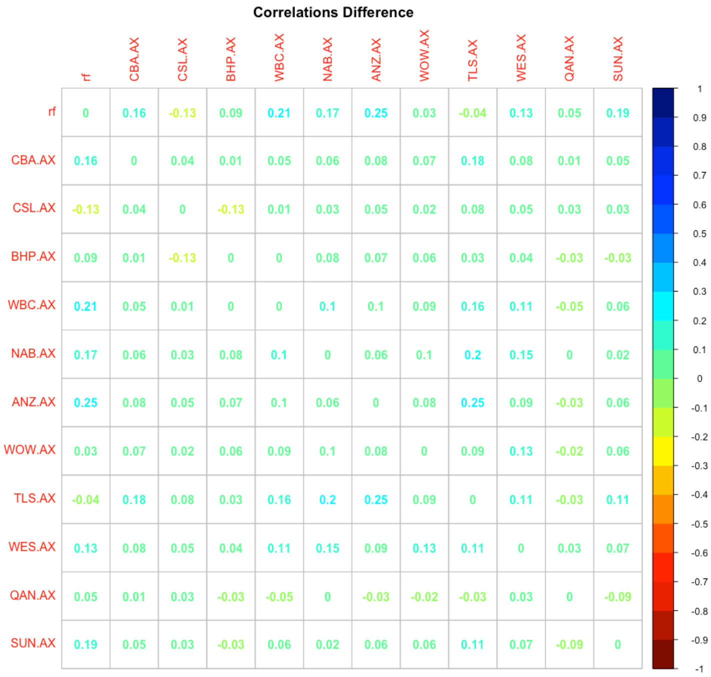


Normal distribution of excess returns lessens instability due to nature of data

SID: 500494994
UOS: ECMT2130



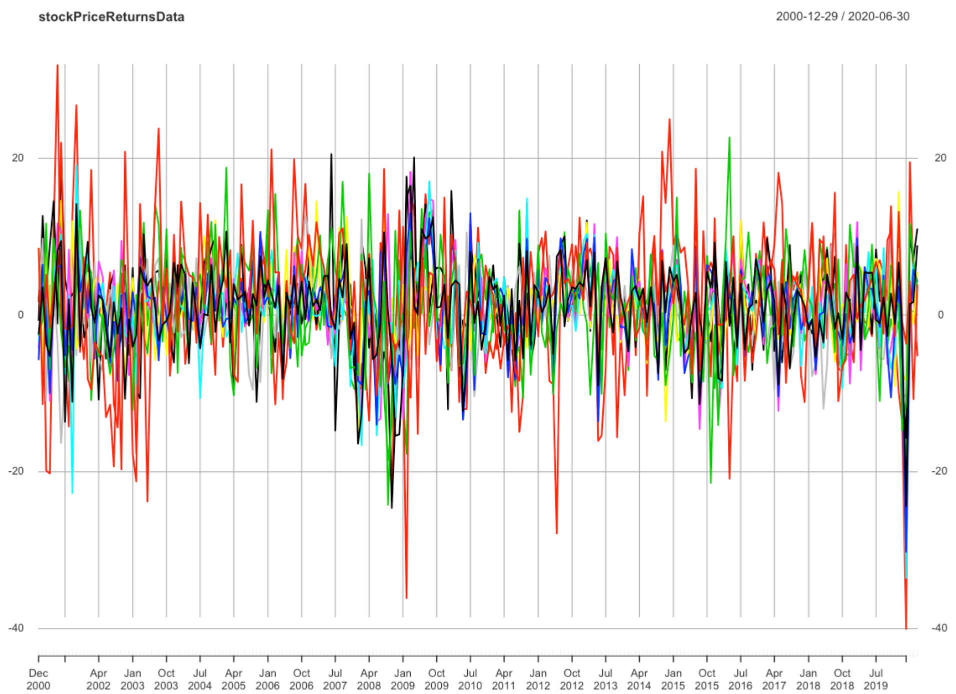




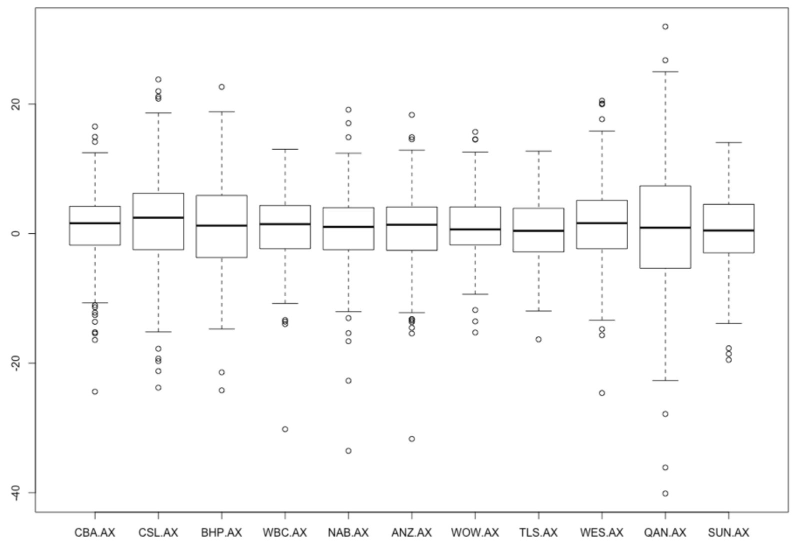
The differences between the two correlation time series can be shown above. Evidently there are no values which exceed 0.25, which occurs at TLS vs ANZ and ANZ vs RF. Apart from these examples, there is not much movement in correlation between stocks, thus making this portfolio relatively stable. As shown in the histogram above, the data is distributed normally, which indicate stable nature, as there are no bias or impacting amounts of outliers.

SID: 500494994
UOS: ECMT2130

Extreme outliers and data preparation process



This graph represents the return price on stock and features three main points of outliers, around 2001, 2009 and 2020. Each of these outlying points have been due to global factors which affects stock prices. In 2001, the 911 attacks occurred which led to a 684-market point loss. Furthermore, early 2009 was the peak of the GFC where the market consequently crashed around 777%. More recently the outbreak of the Corona Virus in March 2020 led to an 8000-point drop in the Dow Jones. This data has been prepared properly as this is objective data, with no biasness. The outliers have been affected mostly due to global stock crashes.



In addition, the boxplot above labels the said outliers and which stocks they come from.

SID: 500494994
UOS: ECMT2130

Adequacy of monthly return computation

Portfolio monthly returns is computed through taking one hundredth times the difference between portfolio values (between 2010-01-29 to 2019-11-29) divided by the portfolio values time series. By dividing these two returns, the number outputted is a ratio, thus return is turned into a percentage due to it being multiplied by 100. Since these returns are computed using 20 years of historical, investors should base trends on their investment strategies. For example, if you were a long-term investor, all 20 years should be looked at, however if you wanted a shorter investment, 5 or 10 years of data would be more valuable as it represents a more similar era in regard to the stock market and global factors.

Adequacy of missing value compensation

The program omits all rows where there are values missing. This means that there will be certain number of dates where there are no data at all. Although having missing rows is better than having rows with empty entries, these values could be found in online archives, similar to Yahoo Finance. The most unbiased datasets would have no missing data and no missing rows, however currently, omitting rows is a better course of action than preceding with datasets which rows have missing values. There are only 11 missing values in this dataset, which is near-perfect.

SID: 500494994
UOS: ECMT2130

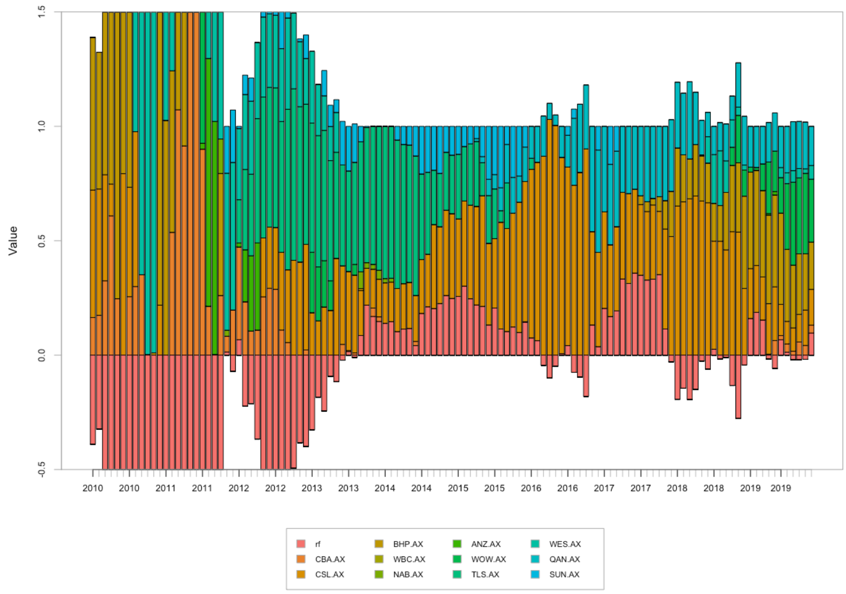
Trading Algorithm

Features of the trading algorithm and impact on portfolio weights

The algorithm imports historical data from 2010 – 2019 and defines the portfolio as risky asset names. By adding a constrain “Full Investment” the trading algorithm maximises weights at a sum of 1. Furthermore, the algorithm assumes that the investor can borrow funds, thus not limiting weightings -1 < w < 1. The objective of this algorithm is to minimise risk through standard deviations. Using ROI, the program scans for the optimal weights for the investor to invest in. These attributes can be summarised below.

```
PortfolioAnalytics Portfolio Specification
**************************************************

Call:
portfolio.spec(assets = riskyAssetNames)

Number of assets: 11
Asset Names
 [1] "CBA.AX" "CSL.AX" "BHP.AX" "WBC.AX" "NAB.AX" "ANZ.AX" "WOW.AX" "TLS.AX" "WES.AX" "QAN.AX"
More than 10 assets, only printing the first 10

Constraints
Enabled constraint types
                - full_investment
                - long_only
                - return

Objectives:
Enabled objective names
                - StdDev
```

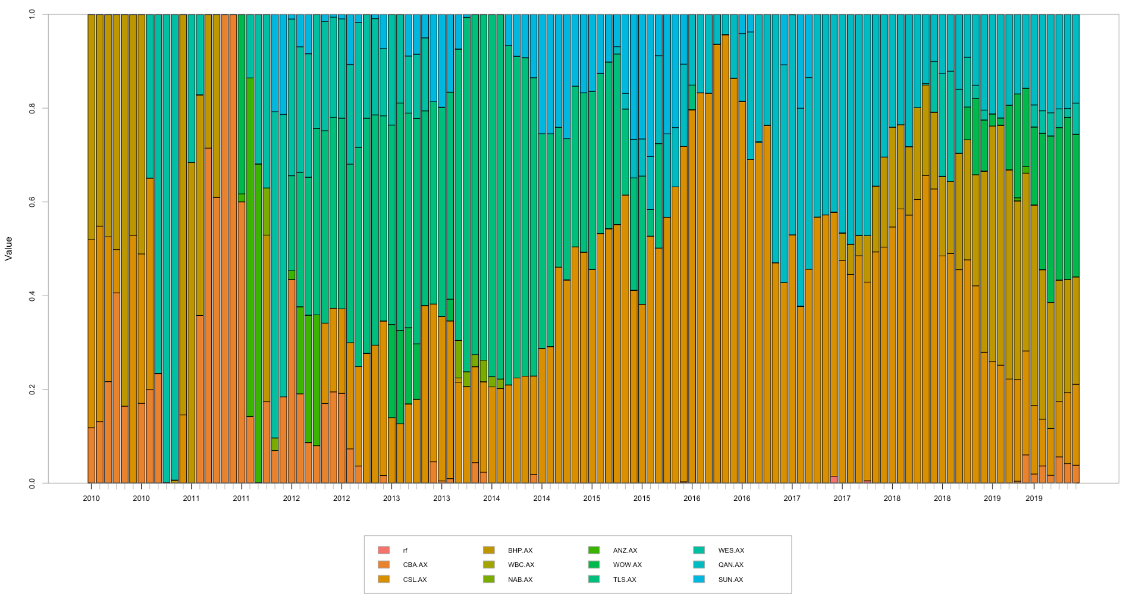
Using the program, the weights are as shown



The program sets leverage ceiling at -0.5. This is the risk-free weight limit. In addition, target return was set at 2% along the CAL, and average monthly portfolio return was 1.06%. Using the heldback data, portfolio return was 1.4%

SID: 500494994
UOS: ECMT2130

For a risk averse investor, set leverage ceiling at 0, and thus we can also set target return at 3% due to expected returns having less risk. Using this, weights are shown below.



Using this, the mean monthly portfolio returns are 0.93% whilst heldback data returned 1.44%

Features which can compromise informativeness

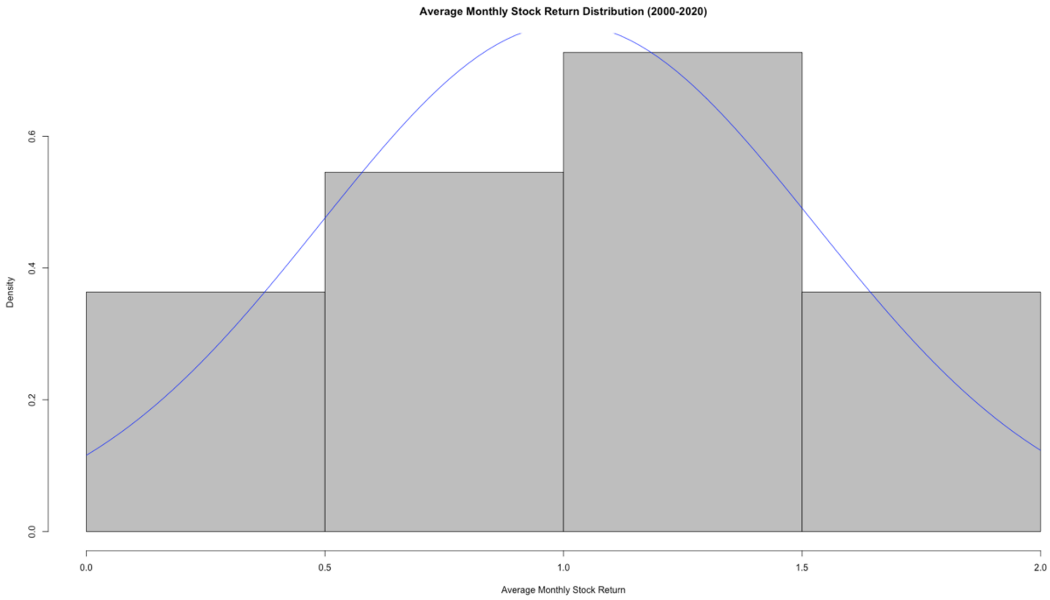
The trading algorithm allows leverage at a negative rate. This means that short selling is viable. The algorithm will not let you invest at a higher leverage of -0.5, which incurs higher risk. The development process must set up an algorithm where the leverage and target return are at an achievable rate and can be tested using the training data. However, as this is only the training data, the algorithm must also be tested using the held back data, in order for prediction be as accurate as possible. The training data must also be in a time period where data is as normally distributed as possible in order to limit number of outliers due to global affects. For example, if data were used from a recession, expected returns would be lower. Conversely, a booming economical period will have a higher rate of return, thus these factors can compromise informativeness. Evidently in the previous subheading, both portfolios had a higher return when using heldback data, as it included 10 more years of data, which makes the data more normal.

SID: 500494994
UOS: ECMT2130

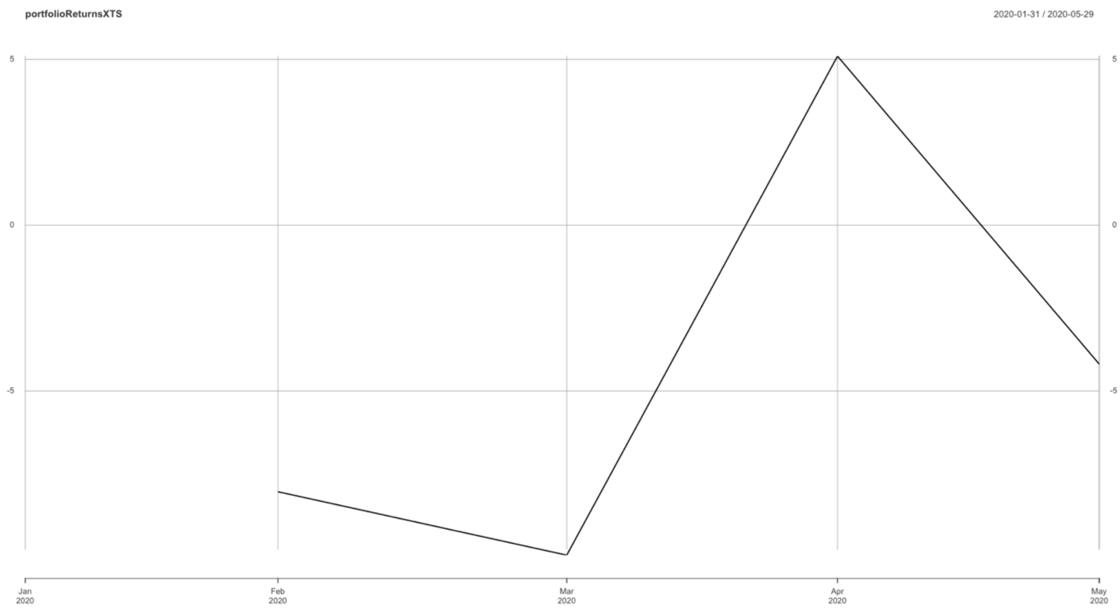
Performance Assessment

Likeliness that predicted profits will be informative

Using a Shapiro-Wilk normality test on 2000-2020 data, the p-value returns as 0.675. As this is above the 0.05 cutoff, we retain the normality hypothesis.

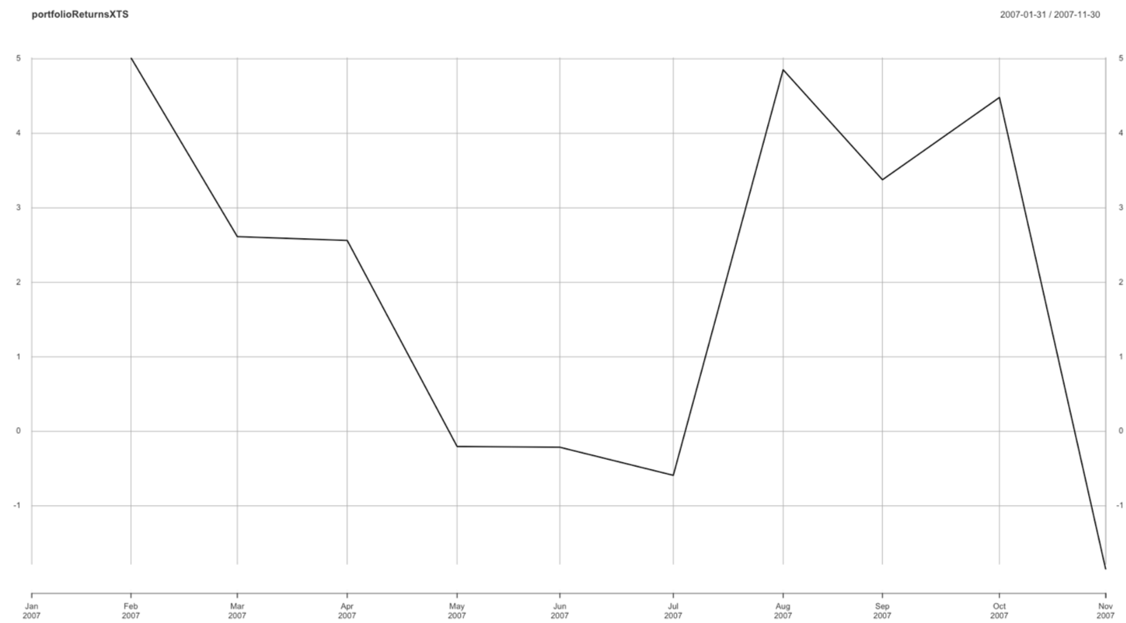


Consequently, as the data is normal, there is no biasness, meaning this data is viable for future predictions. However, the algorithm performs different depending on different kinds of market, especially recessions and booms. Using the coronavirus recession 2020/01/01 – 2020/06/30, mean monthly portfolio returns were at -4.27% (sd = 6.69)



SID: 500494994
UOS: ECMT2130

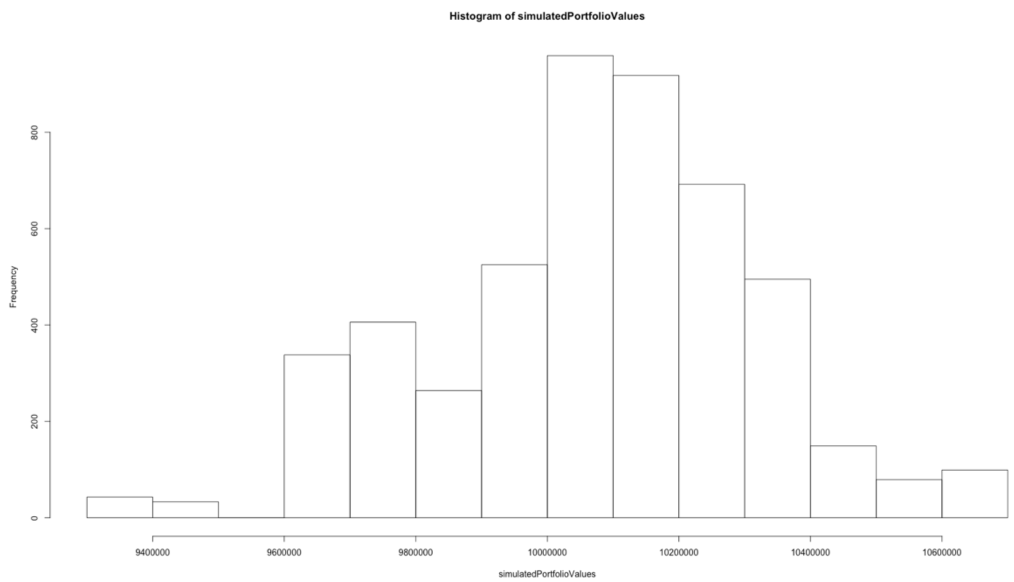
Conversely, in a boom (2007-2008) the average returns were at 2% (sd = 2.5)



Evidently, which dataset you use will greatly affect the profit estimation, thus investors should ensure data used includes both recessions and booms in order to predict for the future. As there are both of these in timespan 2000-2020, the dataset is likely to be informative about future performance.

VaR

Due to a substantial amount of data entries between 2010-2019, a historical simulation is used for a 1-month time horizon at the 99th percentile. The dataset is not extended from 2000, as the older the data, the higher chance of irrelevance. Drawing random values from 2010-2019, the VaR is $9657162. This means the worst scenario in the month (1% chance), the investor would lose $9657162. The VaR uses the last row of weights, in order to produce the most optimal outcome.



Suggestions and conclusion

The client should use this trading algorithm however must adapt it to your needs. Ceiling leverage, weight limits, and target return should be modified based on investors level or riskiness and goals. The data used proves to be normal with unbiases and has missing values omitted, thus being suitable for this algorithm. Long term investments should adopt the whole timespan (2000-2020) whilst short investments should adopt shorter timespans (2015-2020 / 2010-2020). If the investor wanted to ensure that this dataset is trustworthy, and to test the algorithm again, using another source to double check data is viable (e.g. ASX). From here, the investor should modify personal information in regard to initial portfolio value, transaction costs, and the possibility of investing in different stock.