



SMU

**SINGAPORE MANAGEMENT
UNIVERSITY**

ACCT 421

Analytics for Value Investing

AY2020-2021, Term 2

Group Project Report

Class section: G1

Prepared by:

Hou Hongzhou (01356839)

Liu Xiaorong (01358961)

Lu Kaiqi (01351798)

Song Zitong (01359190)

Prepared for:

Prof. Andrew Lee

Table of Content

Executive Summary	3
1. Introduction.....	4
2. Research methodology.....	4
2.1 Profitability	4
2.2 Leverage	4
2.3 Operating efficiency	5
2.4 Book-to-Market (BM) ratio	5
2.5 Market Capitalization	5
3. Research Design	6
3.1 Monthly stock return data	6
3.2 Accounting signals	6
3.3 Ratio computation	6
3.4 F_SCORE.....	6
3.5 Statistical significance of signals.....	7
4. Portfolio formation and assessment	7
4.1 Strategy 1: Long Only Portfolio	8
4.2 Strategy 2: Long/Short Portfolio.....	8
4.3 Strategy 3: Optimized Long/Short Portfolio	9
5. Evaluation of performance and interpretation	10
5.1 Performance metrics	10
5.2 Interpretation of Strategy 3	11
6. Limitations of strategies and potential future improvements	12
6.1 Lack of portfolio weight optimization.....	12
6.2 Insignificant difference in mean annual return between strategy and S&P 500.....	12
7. Conclusion	12
8. Appendix.....	13
8.1 Computation of financial ratios.....	13
8.2 Iterative parameter tuning	14
8.3 Visualization of Strategy cumulative annual return with an investment of \$1, 000 ...	15
8.4 Strategy 4: Optimized Long-Only Strategy.....	16
9. Reference	18

Executive Summary

This report summarizes the use of historical financial data as an investment strategy to generate alpha returns. The team adopted nine financial signals to measure three main areas of the firm's financial condition. These signals help to evaluate the firm's potential and ability, which would in turn reflect in their stock prices. In addition, the team considered market signals such as Book-To-Market ratio and Market Capitalization in the model. For the accounting signals, each signal is tied to a binary indicator variable where one is assigned when the result is good and zero otherwise. The sum of the nine binary signals would form the F_SCORE. The team created portfolio based on the three-selection criterion: F_SCORE, Market Capitalization and Book-to market ratio.

With reference to Piotroski (2000) and G&C (2013)'s selection criteria, the team considered three strategies: long only portfolio; long/short portfolio; and optimized long/short portfolio with different selection criteria on F_SCORE, market capitalization and Book-to-market ratio.

Among the strategies, Strategy 3: Optimized Long/Short Portfolio produced the highest 20-year CAGR of 22.47%, compared to S&P500's 9.45%. Sharpe Ratio of Strategy 3 (0.87) also outperformed S&P500's 0.34. The maximum drawdown of 45% compared to S&P500's 86% also showed that Strategy 3 provided a more stable return.

Our strategies demonstrate that Market Cap, BM ratio and F_SCORE are accurate criteria to select over-valued and financially weak stocks for shorting; and they are very likely to have monotonic relationships with the portfolio return.

In addition, the team recognized that due to the ease of implementation, there are certain limitations to the strategies, such as the use of equal-weighted portfolios. The t-test of difference in annual returns of strategies against benchmark also showed insignificant difference.

1. Introduction

The purpose of this paper is to develop an investment strategy using historical financial statement information that can generate alpha returns. The methodology will be mostly applicable to value stocks whether there will be more trustworthy financial data available and more undervaluation which signals the higher returns.

The investment strategy will take inspiration mainly from Joseph D. Piotroski (Piotroski)'s research *Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers* and well as Gray & Carlisle (G&C)'s research *Quantitative Value*.

2. Research methodology

With reference to Piotroski (2001), one could generate alpha by identifying a firm's intrinsic value or systematic errors in market expectation. If there are available analysts' earnings forecasts and well-developed accounting model, it would be more likely to successfully identify undervalued stocks. In reality, investors usually pay less attention to high book-to-market firms. As a result, analyst forecasts and stock recommendations are unavailable for these firms. In addition, voluntary disclosure by these firms may be viewed as uncredible due to their poor performance. Therefore, instead of relying on forecasts, financial reports appear to be a better source of information for investors.

Lev and Thiagarajan (1993) suggested that useful financial signals are correlated with contemporaneous returns after controlling for current earnings innovations, firm size and macroeconomic conditions. The market's inability to fully process all financial signals creates opportunity to yield abnormal returns. Piotroski (2001) suggested to use the best and most relevant source of information about a firm's financial condition to differentiate strong and weak firms. Following this idea, 9 fundamental signals were selected to measure three main areas of the firm's financial condition.

These signals were classified into three categories, namely profitability; leverage, liquidity, source of funds; and operating efficiency.

In addition, besides accounting signals, market signals such as Book-to-Market ratio and Market Capitalization are also used to indicate the firms' market performance.

2.1 Profitability

Current profitability and cash flow realizations suggest about the firm's ability to generate funds internally, and it is measured by its return on asset (ROA), cash flow from operations, change in yearly ROA and accrual. If a firm is able to generate positive cash flow despite its poor performance in the past, it is somehow evident that there are potential and capacity in this firm.

2.2 Leverage

Changes in a firm's capital structure and its ability to meet future debt obligations can be measure by its change in leverage, liquidity and issuance of common equity. An increase in

leverage implies additional burden on the firm to repay the increasing long-term debt, and also suggests that it does not have the capability to generate sufficient internal funds. Similarly, issuing common equity stocks also implies that the firm is unable to generate sufficient internal funds and could only rely on issuing of stocks to raise funds. On the other hand, an increase in liquidity signals that the firm is improving on its ability to fulfill its current debt obligations.

2.3 Operating efficiency

Additionally, it is important to evaluate how a firm utilizes its assets to generate profit, namely its operating efficiency. Change in the firm's current gross margin ratio helps to identify potential improvements in the firm's management on the cost and selling price of its products. Any increase in profit margin signals that the firm has the ability to either reduce cost or raise the selling price to generate increasing profit for future expansions. Similarly, asset turnover ratio signals the firm's productivity from the asset base. An improvement in asset turnover ratio could be resulted from more efficient operations or increase in sales, which both contributes to its operating profitability.

2.4 Book-to-Market (BM) ratio

The BM ratio compares the company's book value to its market value. It helps the investors to measure the true value of the stocks. In general, high BM ratio is preferred by value investors as it signals that the stocks are undervalued as they are trading at a lower price of what the companies are actually worth.

2.5 Market Capitalization

Market capitalization is calculated by multiplying the company's outstanding shares by the current share price. Instead of company's assets or sales, investors usually use market capitalization to measure the size of the stocks. The stock prices of companies with small market capitalization are usually more volatile. This allows the investors to make short-term returns while bearing more risks. On the other hand, companies with large market capitalization are likely to generate consistent returns in the long run.

3. Research Design

3.1 Monthly stock return data

The team extracted the monthly stock return data from Compustat for the period 1 January 2000 to 31 December 2020. The portfolio formation year variable is then created, where stock data released between 1 July of current year to 30 June of the following year will be included in portfolio formation on 30 June of current year. Monthly returns are then compounded to get annual returns at end-June of each portfolio and winsorization is used to limit the outliers using 2.5% benchmark cut-offs.

3.2 Accounting signals

The team collected the relevant annual financial statement signals for all tradable universe of U.S. companies excluding banks and financial institutions from Compustat for the period 1 January 2000 to 31 December 2020. Similarly, the portfolio formation year variable is then created, where stock data released between 1 April of current year to 31 March of the following year will be included in portfolio formation on 30 June of following year.

Variables from Compustat

IB: Income before extraordinary items
AT: Assets – Total
OANCF: Operating activities net cash flow
GP: Gross profit(loss)
REVT: Revenue – Total
DLTT: Long-term debt – Total
ACT: Current assets – Total
LCT: Current liabilities – Total
MKVALT: Market Value – Total – Fiscal
LT: Liabilities – Total
CSHI: Common shares issued

Table 1: Variables from Compustat

3.3 Ratio computation

Based on the signals that the team have identified in Research Methodology, the team computed the financial ratios and assigned binary variables according to the ratios. The detailed computations are illustrated in Appendix 8.1. Accordingly, the 3 financial performances are measured by:

- Profitability: F_ROA, F_CFO, F_ΔROA, F_ACCURAL
- Leverage: F_ΔLEVER, F_ΔLIQUID, EQ_OFFER
- Operating efficiency: F_ΔTURN, F_ΔMARGIN

3.4 F_SCORE

F_SCORE is the one of the primary signals the team will use in portfolio formation. It is the sum of F_ROA, F_CFO, F_ΔROA, F_ACCRUAL, F_ΔMARGIN, F_ΔTURN, F_ΔLEVER, F_ΔLIQUID and EQ_OFFER. The F_SCORE has a range from 0 to 9, with 0 meaning the weakest financial fundamentals and 9 meaning strongest financial fundamentals.

3.5 Statistical significance of signals

The team has selected three important signals for portfolio formation criteria: F_SCORE, BM ratio and Market Capitalization. In order to determine whether these ratios are statistically significant in predicting individual stock returns, the team further performed a year-by-year OLS regression of yearly stock return against (F_SCORE, BM ratio and Market Capitalization) and generated the following statistics:

	t-stat	p-value
Constant	-0.0339	0.9733
F_SCORE	1.8212	0.0836
BM ratio	1.1678	0.2566
Market Capitalization	-0.8308	0.4159

Table 2: T-statistics and p-values of year-by-year coefficients

The use of a year-by-year OLS regression instead of a pooled OLS regression is to reduce the impact of autocorrelation and heteroskedasticity on the t-test of coefficient estimates.

As seen from the regression statistics, the three ratios have p-value greater than 0.05, the significant level. This suggests that the evidence is not sufficient for us to reject the null hypothesis that the coefficient estimates of these variables are 0. As a result, this could mean that the three variables are not statistically significant in predicting the yearly stock return.

However, the team has still decided to use the variables for the portfolio selection due to the significance demonstrated in Piotroski's research. The team also aim to test if using these variables as criteria can help to construct an outperforming strategy.

4. Portfolio formation and assessment

After merging the stock return data with the accounting signal data, the team performed back testing with rolling validation where the portfolio will be rebalanced at each June end. The purpose of rolling validation is to preserve the autocorrelation in the cross-section and time series of the financial data as well as to avoid any hindsight bias.

The team has considered two types of methods: Long-Only and Long/Short strategies with a total of 3 strategies as explained below. The first two portfolios attempt to mimic Piotroski's and G&C's selection criteria and test if the strategies can still outperform the market in a different time period as the research papers. The last strategy is an optimized long/short portfolio using iterative parameter tuning.

4.1 Strategy 1: Long Only Portfolio

For each portfolio formation year, the team selects the stocks based on the below criteria:

Signals	Selection Criteria
Market capitalization	MKVALT \geq 40th percentile
Book-to-market ratio	BM ratio \geq 90th percentile
F-score	F-score \geq 8

Table 3: Signal selection criteria for long only portfolio

* Percentile is based on the sample for the particular portfolio formation year, not the whole dataset.

According to G&C (2013), stocks with a small market capitalization were able to generate a higher return. However, small market capitalization stocks tend to have wide bid/ask spreads and very little liquidity at the bid or ask. Hence, to avoid misleading performance data that could not be achieved in the real world, the team decided to drop the stocks that are in the lower 40th percentile of market capitalization.

Since the BM ratio measures the cheapness of the stock and whether the stocks are over or under-valued. Stocks with a higher BM ratio would imply that they are under-valued and cheaper. G&C (2013) dropped the bottom 90% on EBIT/EV, since EBIT/ EV and BM ratio are similar signals that measure the cheapness of the stock. The team decided to adopt the benchmark and drop the stocks that are in the lower 90th percentile of BM ratio.

The team also adopted the benchmark for F_SCORE used in Piotroski (2000). Firms with F_SCORE equals to 8 or 9, which signify strong financial fundamentals, are selected.

After selecting the stocks based on the selection criteria for each portfolio formation year, the team then compute the mean return for each portfolio formation year and calculate the performance metrics as shown in Section 5.1 and Appendix 8.3.

4.2 Strategy 2: Long/Short Portfolio

For each portfolio formation year, the team selected the stocks based on the below criteria:

Signals	Selection Criteria	
	Long	Short
Market capitalization	MKVALT \geq 40 th percentile	MKVALT \geq 40 th percentile
Book-to-market ratio	BM ratio \geq 90 th percentile	BM Ratio \leq 10 th percentile
F-score	F-SCORE \geq 8	F-SCORE \leq 1

Table 4: Signal selection criteria for long/short portfolio

For Long/Short Portfolio, the selection criteria for longing a stock remain the same as Strategy 1.

For the stocks to be shorted, the team selected those with BM ratio in the lowest decile which is less than or equal to 10th percentile to include the over-valued stocks. Within the selected stocks, the team then chose the financially weak companies with F_SCORE ≤ 1 . The returns of the shorted stocks are the negative of the original calculated returns.

The performance metrics are shown in Section 5.1 and Appendix 8.3.

4.3 Strategy 3: Optimized Long/Short Portfolio

For each portfolio formation year, the team selects the stocks based on the below criteria:

Signals	Selection Criteria	
	Long	Short
Market capitalization	MKVALT \geq 35th percentile	MKVALT \geq 35th percentile
Book-to-market ratio	BM ratio \geq 95th percentile	BM ratio \leq 5th percentile
F-score	F-SCORE ≥ 9	F-SCORE ≤ 2

Table 5: Signal selection criteria for optimized long/short portfolio

With reference to the performance metrics in Table 6, the performances of Strategy 1 and 2 appear to be mediocre with CAGRs slightly higher than S&P 500. To improve the performance of the strategy, the team used iterative parameter tuning which is explained in detail in Appendix 8.2.

The team generated a list of possible values for the cut-off points for the 5 parameters: Market_Cap_Cutoff, BM_Long_Cutoff, F_Long_Cutoff, BM_Short_Cutoff, F_Short_Cutoff which correspond to the selection criteria shown in Table 5. All possible combinations of the parameters are then used in the portfolio formation process to find out the most optimal cut-off points for the selection criteria that generates the best CAGR.

The performance metrics of Strategy 3 are shown in Section 5.1 and Appendix 8.3 and more detailed interpretation of Strategy 3 is explained in Section 5.2.

5. Evaluation of performance and interpretation

5.1 Performance metrics

To evaluate the performance of the investment strategies, the team have computed a few key performance metrics: CAGR (Compounded Annual Growth Rate), Sharp ratio, Sortino ratio, Max drawdown, t-stats and compared them against S&P 500.

	Strategy 1	Strategy 2	Strategy 3	S&P 500
CAGR/Compounded annual return	9.80%	11.37%	22.47%	9.45%
Sharpe ratio	0.42	0.47	0.87	0.34
Sortino ratio	1.63	1.60	1.59	1.53
Maximum drawdown	46.0%	45.0%	45.0%	86.0%
t-stats of difference in mean annual returns verse S&P 500 (p-value)	-0.2889 (0.7745)	-0.2294 (0.8201)	0.573 (0.5708)	NA

Table 6: Performance metrics

CAGR is a good indicator for evaluating how different strategies have performed over time, it measures the investment's annual growth rate over time, with the effect of compounding taken into account. From the table above, it can be observed that Strategy 3 has significantly outperformed the other strategies as well as S&P 500, earning a staggering 20-year CAGR of 22.47%. However, CAGR does not reflect volatility as it assumes growth to be constant throughout the investment's time horizon when in fact the value of the underlying investment can vary significantly. Hence, other performance metrics such as Sharp ratio and Sortino ratio are calculated to factor in the volatility portion of the investment.

Looking at the Sharpe ratio which measures the performance of the strategy in terms of excess return per unit of risk, strategy 3 has the highest Sharpe ratio, implying that Strategy 3 would be able to generate a higher return on a risk-adjusted basis.

Some investors are more concern for the downside risk as upside volatility is a plus for the investment. Hence, the Sortino ratio which is a variation of the Sharpe ratio that differentiates harmful volatility from total overall volatility, is a good indicator of negative volatility as it only considers the standard deviation of the downside risk. Even though the Sortino ratio of Strategy 3 is slightly lower than the other strategies, it is still higher than that of S&P 500. A higher Sortino ratio would mean that the strategy is earning more return per unit of the downside risk that it takes on.

The maximum drawdown on a strategy is the maximum cumulative percentage loss suffered under the strategy. It shows the volatility of the strategy in the past and provides an accurate way of predicting future price movements. A low maximum drawdown value indicates slight fluctuations in the value of the strategy. Since strategy 3 has a lower maximum drawdown value of 45% as compared to S&P 500 which has a value of 86.0%, this implies that strategy 3 is able to generate a more stable return.

Looking at the t-stats of difference in mean annual returns verse S&P 500, the team noted that the p-values for all the strategies are above the significance level of 0.05. Hence, there is insufficient evidence to reject the null hypothesis that the mean annual returns of these strategies are equal to that of S&P 500.

Comparing across the three strategies, Strategy 1 which is the Long-Only strategy produced the lowest performance for all the metrics. Even with iterative parameter tuning (see Appendix 8.3 - Strategy 4), the optimized Long-Only strategy (CAGR: 12.76%) only perform slightly better than an unoptimized Long/Short strategy (CAGR: 11.37%). This further demonstrates that Market Cap, BM and F_SCORE are accurate criteria to select over-valued and financially weak stocks for shorting; and they are very likely to have monotonic relationships with the portfolio return.

5.2 Interpretation of Strategy 3

With reference to the selection criteria used in Table 5, the team observed that they differ slightly from those of Piotroski's and G&C's but essentially align with the theories and research results.

Market capitalization cut-off: The market capitalization cut-off has reduced from 40th to 35th percentile, which means that the portfolio now includes smaller-cap stocks. This aligns with Section 2.5 that small-cap companies tend to be volatile and generate higher return. The inclusion of small-cap stocks could have contributed to the higher return in the portfolio.

BM ratio cut-offs: The BM cut-off for longing a stock has increased from 90th to 95th percentile, which suggests that the portfolio now focuses on the more undervalued stocks. This is consistent with Piotroski and G&C as they discovered that high BM stocks tend to generate higher return.

The BM cut-off for shorting a stock has also decreased from 10th to 5th percentile which suggests that the portfolio now includes the more overvalued stocks. This also aligns with the research that low BM stocks tend to generate higher loss and hence shorting them would mean a gain for the portfolio.

The difference in cut-off might simply be because the team are using a more granular iteration during the parameter tuning e.g., in steps of 5% instead of 20% as in Piotroski.

F_SCORE cut-offs: The F_SCORE cut-off for longing a stock has increased from 8 to 9 which suggests that the portfolio now considers the financially strongest stocks. This is consistent with Piotroski and G&C as they discovered the monotonic relationship between portfolio return and F_SCORE, that financially stronger firms tend to generate higher return.

The F_SCORE for shorting a stock has also increased from 1 to 2 which suggests that the portfolio now includes financially stronger stocks. This contradicts with the research as it implies that financially stronger stocks generate more loss. Possible explanations could be because either there were negatively abnormal events during the period that caused those stocks' prices to drop; or F_SCORE has lost predictive power especially for financially weaker stocks since the publication of the research papers.

6. Limitations of strategies and potential future improvements

6.1 Lack of portfolio weight optimization

Currently due to the ease of implementation, all the strategies have been using equal-weighted portfolios which means the weight assigned to each stock in the portfolio are not optimized in terms of risk diversification.

The team could potentially improve the risk-adjusted return by incorporating Modern Portfolio Theory introduced by Henry Markowitz. This technique essentially assigns different weights to each stock in the portfolio based on several parameters such as historical price volatility and stock excess return; and identify a set of weights that generates the highest Sharpe ratio, which is a measure of risk-adjusted return.

The technique can be automated using a 3rd party Python package called PyPortfolioOpt (Karve, 2021).

6.2 Insignificant difference in mean annual return between strategy and S&P 500

Referring to the t-test of difference in mean annual return against S&P500 in Table 6, there is insufficient evidence to reject the null hypothesis that the mean annual returns of these strategies are equal to that of S&P 500. This is concerning because the results suggest that the strategies might be only as good as S&P500. It in turn implies that they are passively managed portfolios that only follow the S&P500 movement.

One potential reason could be due to the small sample size as only have 20 annual returns available for the t-test. A small sample size can cause the estimation of sample means to be less precise, and hence can result in extreme t-statistics and p-value (Frost, 2020).

Hence, to increase assurance about the strategies, the team could potentially include more years of data to reduce the uncertainty in the sample mean estimation.

7. Conclusion

In summary, all the three strategies developed using historical financial statements signals are able to generate profits. Among the three, Strategy 3 has the best performance not only in terms of CAGR. Strategy 3 is relatively less volatile and manages to deliver higher returns per unit of downside risks. The limitations of the strategies lie in the lack of portfolio weight optimization which can be incorporated using Modern Portfolio Theory to improve the risk-adjusted return. Moreover, the insignificant difference between the strategies' returns and S&P 500 may further deter the investors from choosing this strategy.

8. Appendix

8.1 Computation of financial ratios

Financial statement signals	Computation	Binary
F_ROA: ROA = net income before extraordinary items scaled by beginning-of-the year total assets	$ROA = IB/Beg. AT$	F_ROA = 1 if ROA is positive, 0 otherwise
F_CFO: CFO = cash flow from operations scaled by beginning of the year total assets	$CFO = OANCF/Beg. AT$	F_CFO = 1 if CFO is positive, 0 otherwise
F_ΔROA: ΔROA = current year's ROA less the prior year's ROA	$\Delta ROA = \text{current year's ROA} - \text{prior year's ROA}$	F_ΔROA = 1 if ΔROA > 0, 0 otherwise
F_ACCRUAL: ACCRUAL = current year's net income before extraordinary items less cash flow from operations, scaled by beginning of the year total assets	$Accrual = (IB - OANCF)/Beg. AT$	F_ACCRUAL = 1 if CFO > ROA, 0 otherwise
F_ΔMARGIN: ΔMARGIN = current gross margin ratio (gross margin scaled by total sales) less the prior year's gross margin ratio.	$MARGIN = GP/REVT$	F_ΔMARGIN = 1 if ΔMARGIN is positive, 0 otherwise
F_ΔTURN: ΔTURN = current year asset turnover ratio (total sales scaled by beginning-of-the-year total assets) less the prior year's asset turnover ratio.	$TURN = REVT/Beg. AT$	F_ΔTURN = 1 if ΔTURN is positive, 0 otherwise
F_ΔLEVER: ΔLEVER = historical change in the ratio of total long-term debt to average total assets	$Leverage\ ratio = DLTT/[(Beg. AT + End AT)/2]$	F_ΔLEVER = 1 if leverage ratio fell in the year preceding portfolio formation, 0 otherwise
F_ΔLIQUID: ΔLIQUID = historical change in current ratio (ratio of current assets to current liabilities) between the current and prior year	$Current\ ratio = ACT/LCT$	F_ΔLIQUID = 1 if current year's current ratio > prior year's current ratio, 0 otherwise
EQ_OFFER: Whether the firm issues common equity in the year preceding portfolio formation	$CSHI - Beg\ CSHI$	EQ_OFFER = 1 if the firm did not issue common equity in the year preceding portfolio formation, 0 otherwise
BM (Book to Market)	$BM = (AT - LT)/MKVALT$	

Table 7: Computation of financial ratios

8.2 Iterative parameter tuning

Refer to Python file “parameter_optimization.py”.

Iterative parameter tuning means fitting different combinations of self-generated parameters into the portfolio formation process. The 5 parameters are generated as followed:

Market_Cap_Cutoff	= values from 30 th to 95 th percentile in step of 5% = list(x for x in np.arange(start=0.3, end=1.0, step=0.05)) * The lower bound of 30 th percentile is set to reduce the number of combinations while still adequately respecting the theory described in Section 4.1 that 40 th percentile is good cut-off.
BM_Long_Cutoff	= values from 55 th to 95 th percentile in step of 5% = list(x for x in np.arange(start=0.55, end=1.0, step=0.05))
F_Long_Cutoff	= integer from 5 to 9 in steps of 1 = list(x for x in range(start=5, end=10, step=1))
BM_Short_Cutoff	= values from 5 th to 45 th percentile in step of 5% = list(x for x in np.arange(start=0.05, end=0.50, step=0.05))
F_Short_Cutoff	= integer from 0 to 4 in steps of 1 = list(x for x in range(start=0, end=5, step=1))

Table 8: Generation of parameters

Using a Python module itertools, the team was able to generate all 28,350 combinations of these parameters:

Combination	Market_Cap	BM_Long	F_Long	BM_Short	F_Short
1	0.3	0.55	5	0.05	0
2	0.3	0.55	5	0.05	1
...
28349	0.95	0.95	9	0.45	3
28350	0.95	0.95	9	0.45	4

Table 9: Illustration of combinations

The team iterated through all these combinations by fitting them into the portfolio formation process. As few combinations are so extreme that there were more than 5 years with no stocks selected, the team dropped these combinations. The combination that generated the highest CAGR is strategy 3 which is explained in Section 5.3.

8.3 Visualization of Strategy cumulative annual return with an investment of \$1, 000

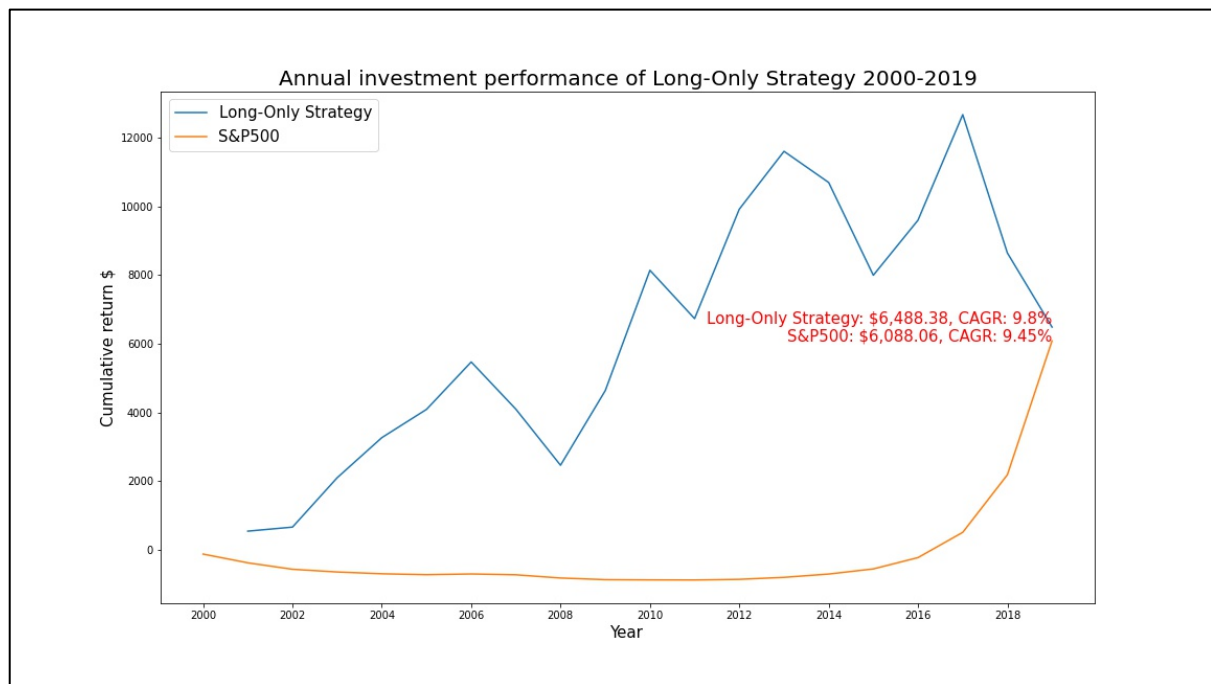


Figure 1: Performance of Strategy 1

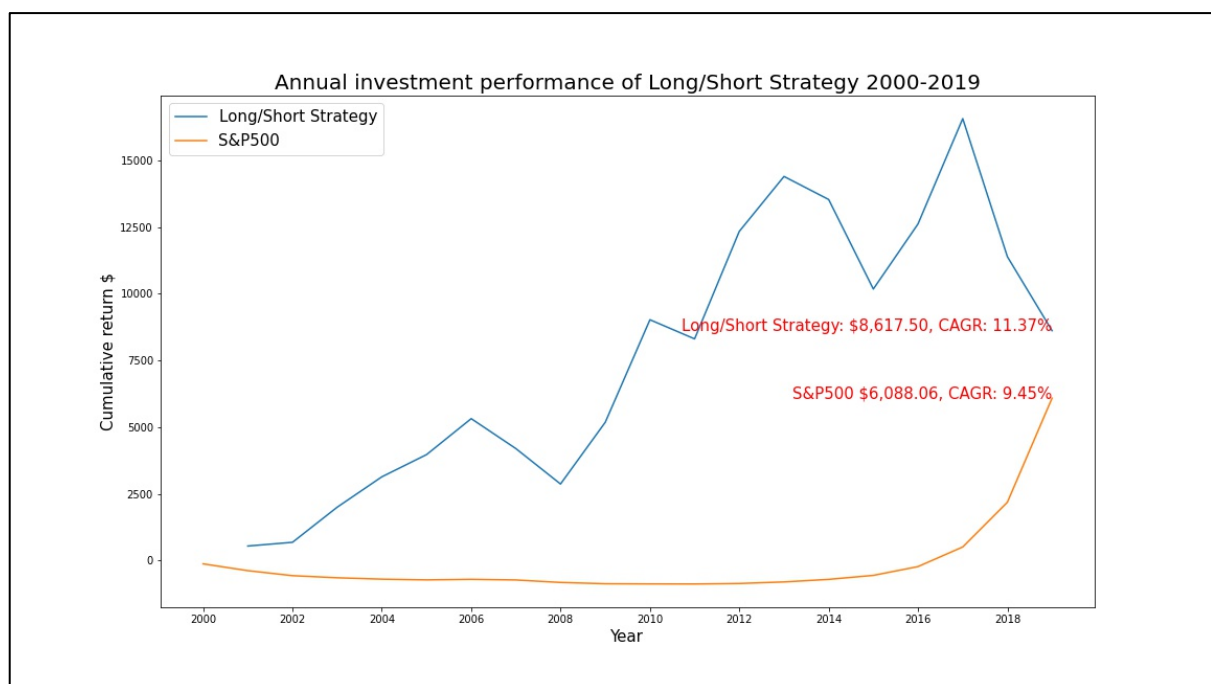


Figure 2: Performance of Strategy 2

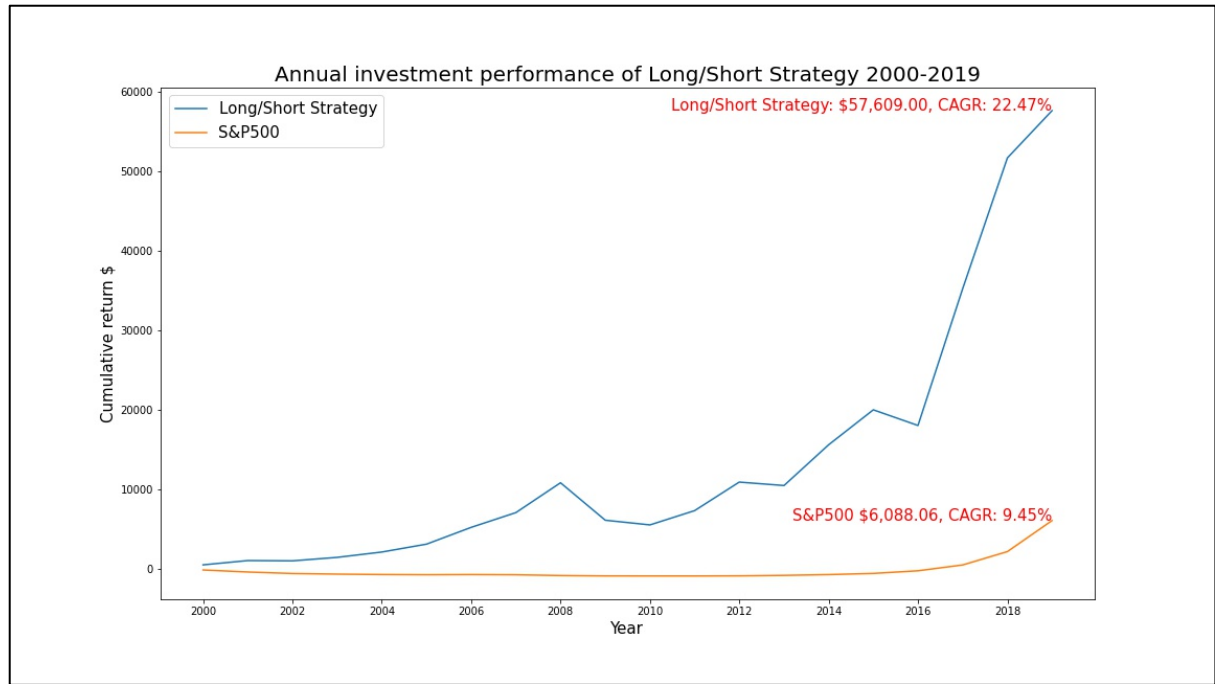


Figure 3: Performance of Strategy 3

8.4 Strategy 4: Optimized Long-Only Strategy

For each portfolio formation year, the team selects the stocks based on the below criteria:

Signals	Selection Criteria
Market capitalization	MKVALT \geq 40th percentile
Book-to-market ratio	BM ratio \geq 95th percentile
F-score	F-score \geq 8

Table 10: Signal selection criteria for optimized long only portfolio

	Strategy 1	Strategy 2	Strategy 3	Strategy 4	S&P 500
CAGR	9.80%	11.37%	22.47%	12.76%	9.45%
Sharpe ratio	0.42	0.47	0.87	0.47	0.34
Sortino ratio	1.63	1.60	1.59	1.69	1.53
Maximum drawdown	46.0%	45.0%	45.0%	57.0%	86.0%
t-stats of difference in annual returns verse S&P 500 (p-value)	-0.2889 (0.7745)	-0.2294 (0.8201)	0.573 (0.5708)	0.051 (0.9596)	NA

Table 11: Performance metrics

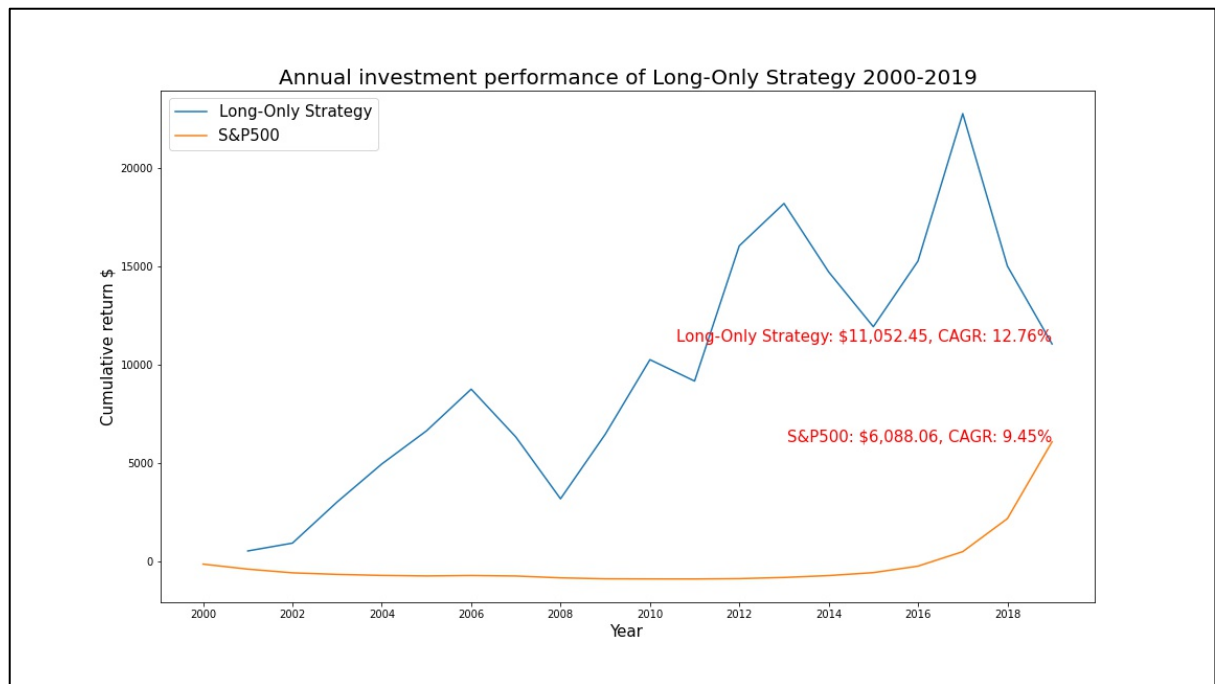


Figure 4: Performance of Strategy 4

9. Reference

- Frost, J. (2020, July 15). How t-tests work: T-values, t-distributions, and probabilities. Retrieved April 08, 2021, from <https://statisticsbyjim.com/hypothesis-testing/t-tests-t-values-t-distributions-probabilities/#:~:text=The%20sample%20size%20for%20a,the%20t%2Ddistribution%20become%20thicker.>
- Gray, W. R., & Carlisle, T. E. (2013). *Quantitative value a practitioner's guide to automating intelligent investment and eliminating behavioural errors*. Hoboken, N.J: Wiley.
- Karve, S. (2021, March 05). Automating portfolio optimization using python. Retrieved April 08, 2021, from <https://towardsdatascience.com/automating-portfolio-optimization-using-python-9f344b9380b9>
- Piotroski, J. (2001). *Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers*.