WallStreetQuant Capstone Project

Statistical Arbitrage in Cryptocurrencies

Kevin Tan YueJun

Table of Contents

1. INTRODUCTION 3
2. KEY PARAMETERS AND DATA 3
2.1 Cryptocurrency Data 3
2.2 Parameters 3
3. PERFORMANCE EVALUATION 3
4. QUANTITATIVE STRATEGY 1 4
4.1 Strategy: Cross-Sectional Equal Volatility Rolling Mean Strategy 4
4.2 Approach 4
4.3 Performance Metrics 5
4.4 Strategy Improvement 6
4.5 Turnover and Net Returns of Strategy 7
4.6 Strategy Conclusion 8
5. QUANTITATIVE STRATEGY 2 9
5.1 Strategy: Moving Average Cross-Over Equal Volatility Strategy 9
5.2 Approach 9
5.3 Performance Metrics 11
5.4 Strategy Improvements 12
5.5 Turnover and Net Returns of Strategy 15
5.6 Strategy Conclusion 15
6. LIMITATIONS 16
7. FUTURE IMPROVEMENTS 16
8. CONCLUSION 16

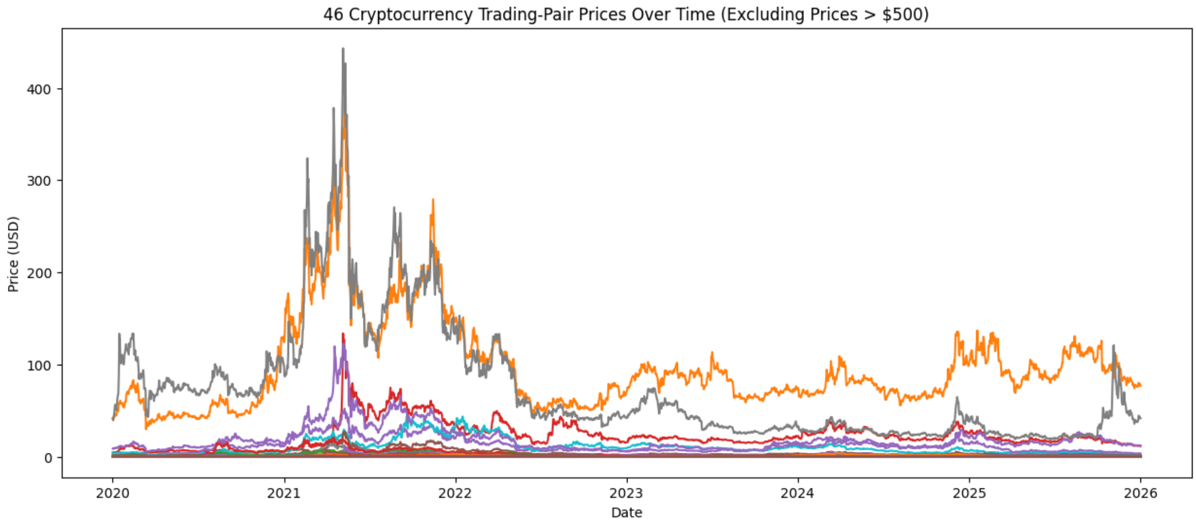
1. Introduction

Statistical Arbitrage is a class of strategies aimed at discovering price-volume patterns that predict returns. It is one of the most popular and successful quantitative hedge-fund strategies, and given that the cryptocurrency markets are still relatively new, it should provide fertile grounds for finding market inefficiencies using statistical arbitrage techniques. The two main patterns that are exploited in statistical arbitrage are namely 1) momentum and 2) reversal.

In the aim of this capstone project is to present two different types of strategies that employ a series of quantitative methods to a basket of cryptocurrencies. By doing so, we seek to curate a portfolio of securities that generates excess returns (Alpha) and outperforms the benchmark.

2. Key Parameters and Data

2.1 Cryptocurrency Data



The cryptocurrency data that we will be using for our strategy construction, backtesting, portfolio allocation and performance evaluation can be retrieved from Binance over a period of 5 years (2020 – 2025). We will set our *Training Data* to be across the period from *2020 – 2024*, and our Test Data to be data from the year 2025.

Binance has a total of **491+** cryptocurrency trading pairs, and after filtering for data within the period of analysis, we were down to approximately **51** trading pairs, which would suffice for our strategy formulation.

2.2 Parameters

Cryptocurrencies can have commissions of **~ 7bps**. While total slippage is unknown and will depend on the trader's position sizing and volume, we will be assuming another **13 bps**. So total all-in execution costs will be **20 bps** (i.e 0.2%) for market-orders. Limit orders will just have the 7 bps of commissions.

- Market Orders: 0.2%
- Limit Orders: 0.07%

For simplicity, our strategies involve placing orders using the closing price of each day and as such, can be considered as a **market order**. With that, we will now explore the two strategies that have been devised as part of the capstone project.

3. Performance Evaluation

To assess the performance of our strategy, we will be evaluating and comparing it against our benchmark, which is a buy-and-hold strategy of Bitcoin ($BTCUSDT$) over the period of analysis. The following table shows the metrics that we will be using to assess our strategy's performance relative to the benchmark.

Metric	Detail / Explanation
Annualised Return	The average return per year, scaled from periodic returns. This metric allows for performance comparison across strategies with different time horizons
Annualised Volatility	The standard deviation of returns, scaled to a yearly basis. It measures risk and variability of returns.
Sharpe Ratio	Excess return (over risk-free rate) per unit of volatility. This metric evaluates risk-adjusted performance of our portfolio. The higher the Sharpe, the better the strategy.
Hit Rate	Percentage of periods (or trades) with positive returns. It indicates consistency and win frequency, independent of payoff size.
Max Drawdown	Largest peak-to-trough loss over the period. This metric captures worst-case capital loss and tail risk.
Alpha	Return in excess of what is explained by market exposure (beta). It measures true skill or strategy edge.
Beta	Sensitivity of the strategy's returns to market returns. It quantifies systematic risk and market dependence.
Information Ratio	Active return divided by tracking error (volatility vs benchmark). It assesses consistency of outperformance relative to a benchmark.

We can derive the *Annualised Return, Annualised Volatility, Sharpe Ratio, Hit Rate and Max Drawdown* of our strategy using the returns that we will be computing for our portfolio on a daily basis. As for the remaining three metrics (*Alpha, Beta, Information Ratio*), we will be running an Ordinary Least Square (OLS) Regression of our portfolio returns against our benchmark. It can be modelled as follow:

$$Portfolio_t \sim \alpha + \beta * Benchmark_t + \varepsilon_t$$

Alpha (α) measures the excess return not explained by the market:
- $\alpha > 0$: Strategy outperformed what the market model predicts
- $\alpha < 0$: Strategy underperformed what relative to its market exposure

Beta (β) measures systematic risk:
- $\beta > 1$: Strategy moves more than the market (high volatility, higher risk)
- $\beta < 1$: Strategy moves less than the market (defensive behavior)
- $\beta \cong 1$: Strategy moves roughly in sync with the market

With the OLS regression model constructed, we can retrieve details on the following key variables:
- $\alpha + \varepsilon_t$ which is the 'tradeable strategy'. The higher the α the better our strategy
- β which will tell us how correlated our asset is to the benchmark. There is no hard and fast rule as to what constitutes a 'good' or 'bad' value for β, but just note that β constitutes high volatility
- Information ratio, which is defined as the **risk-adjusted mean** of our α component

4. Quantitative Strategy 1

4.1 Strategy: Cross-Sectional Equal Volatility Rolling Mean Strategy

Our first strategy is called the **Cross-Sectional Equal Volatility Rolling Mean Strategy**, a momentum-based strategy that seeks to long crypto trading pairs that exhibit high returns from a rolling window basis. Similar to equities, it is possible to find momentum in the shorter timeframe (from weeks to months) before reversion kicks on the longer timeframe (e.g reversion to the mean). The intuition behind this strategy is that assets that performed well over a given period of time will continue to display momentum and outperform in the near term, which assets that delivered poorer returns will continue to underperform in the short run. Henceforth, if we can identify the basket of cryptocurrencies that outperformed its peers in the short run, we can continue to long those cryptocurrencies for a given time period (i.e holding period), before reversion kicks in.

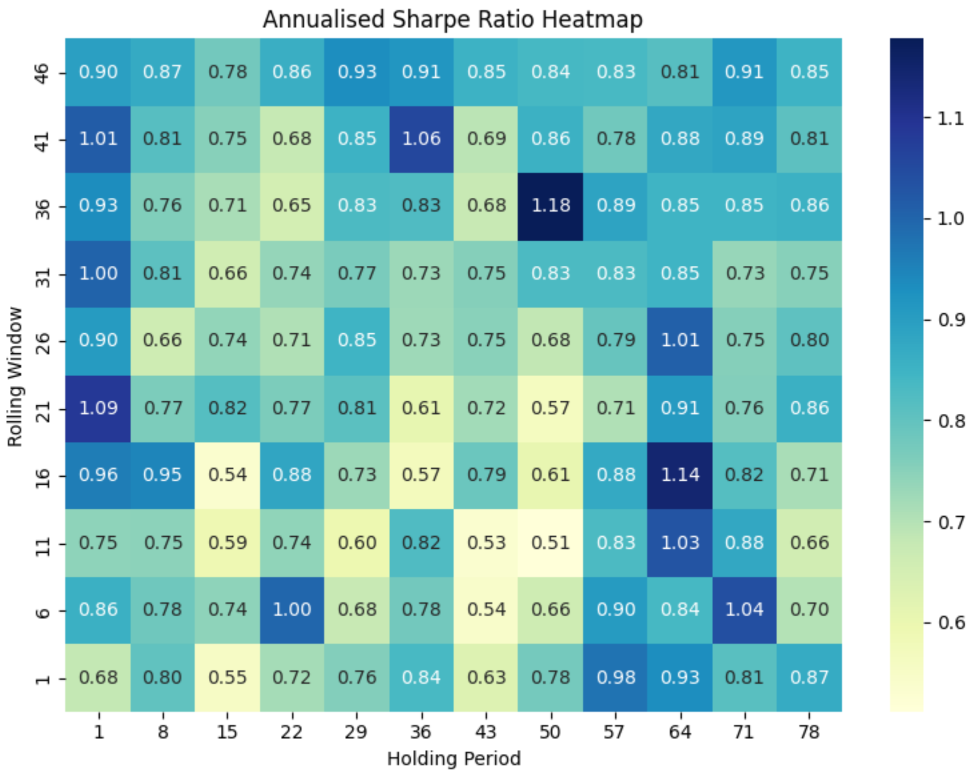
4.2 Approach

First off, we need to identify the ideal rolling window (rw) and holding period (hp) for our strategy, and for that, we will be performing a grid search using our training data to identify the pairs (rw_i, hp_i). For the selection of our cryptocurrencies, we will be using a cross-sectional ranking approach to rank the trading pairs based on their rolling-window returns. Afterwards, we will select the top 10 pairs and **equal weight** (initial approach) them into our portfolio.

For our grid-search, we will be iterating over a period of 50 days for our rolling window (rw) and holding period (hp). For each pair (rw_t, hp_t), where t refers to the number of days, we will be calculating the portfolio returns of our LONG-only portfolio and derive its Sharpe Ratio, assuming the risk-free return r_f is 0.

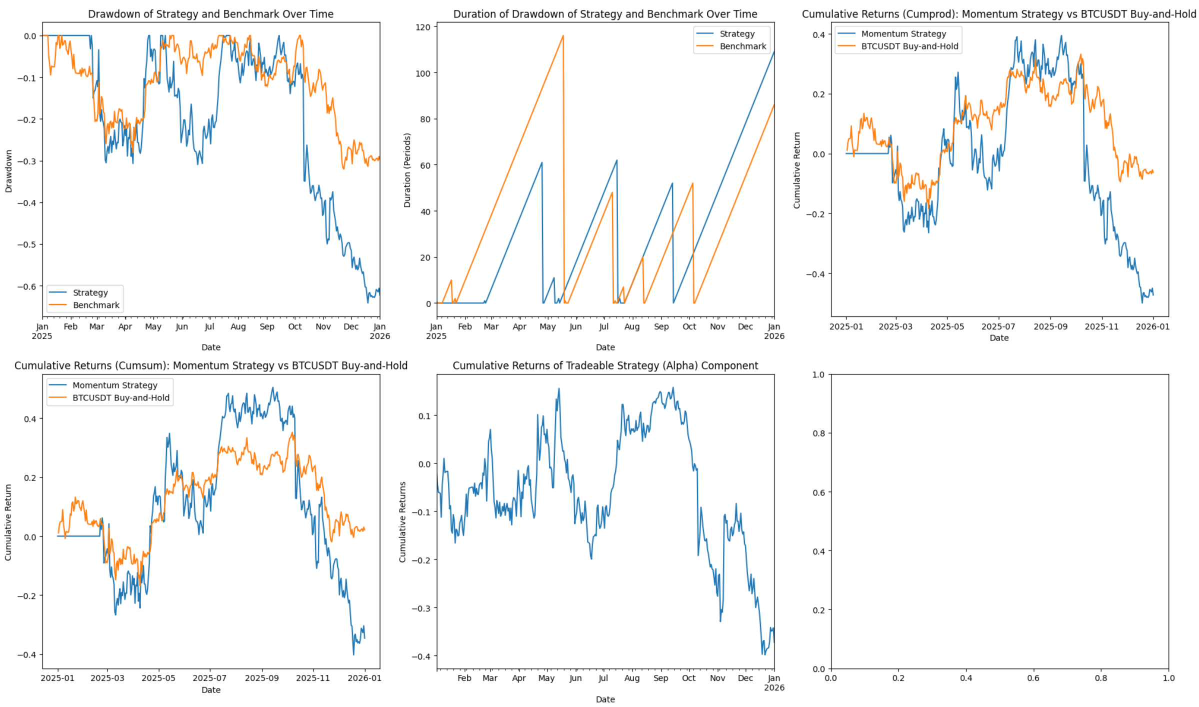
$$Sharpe\ Ratio = \frac{E(r_p)}{\sigma_p}$$

Iterating over each pair (rw_t, hp_t) and plotting the heatmap of its corresponding Sharpe Ratio, we will get the following diagram below.



Based on the heatmap from our training data, we derived our optimal (rw_t, hp_t) pair to be $rw_t = 36$ and $hp_t = 50$. Using this optimal pair, we will test it on our out-of-sample data (full year 2025)

4.3 Performance Metrics



When we ran our **cross-sectional rolling mean** momentum strategy on our test data, we severely underperformed the benchmark across all metrics.

Performance Metric	Strategy	Benchmark
Annualised Return	-0.212186	0.014711
Annualised Volatility	0.629334	0.346389
Sharpe Ratio	-0.378782	0.042162
Hit Rate	0.434426	0.500000
Max Drawdown (%)	64.169652	32.022515
Max Drawdown Duration (periods)	109.000000	116.000000

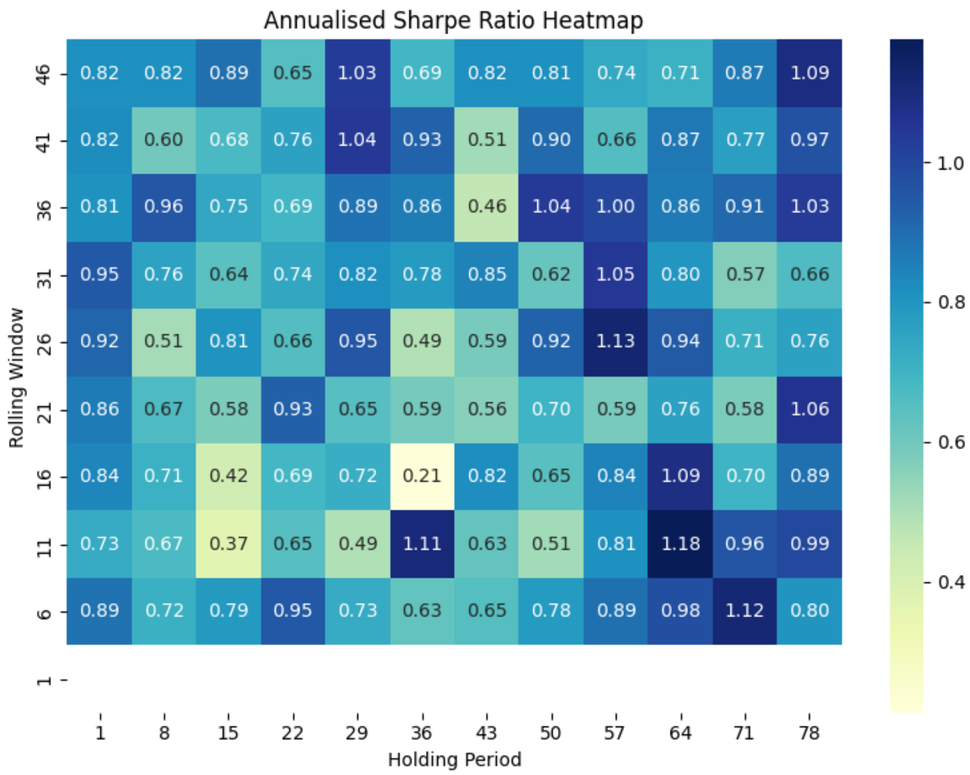
Our strategy generated negative returns with a higher volatility and a longer maximum drawdown (period and depth).

Performance Metric	Strategy
Alpha	-0.00104734
Beta	1.2536
Information Ratio (IR)	-0.57879
Alpha t-stat	-0.696571

As shown above, our remaining three metrics also underperformed with this strategy version.

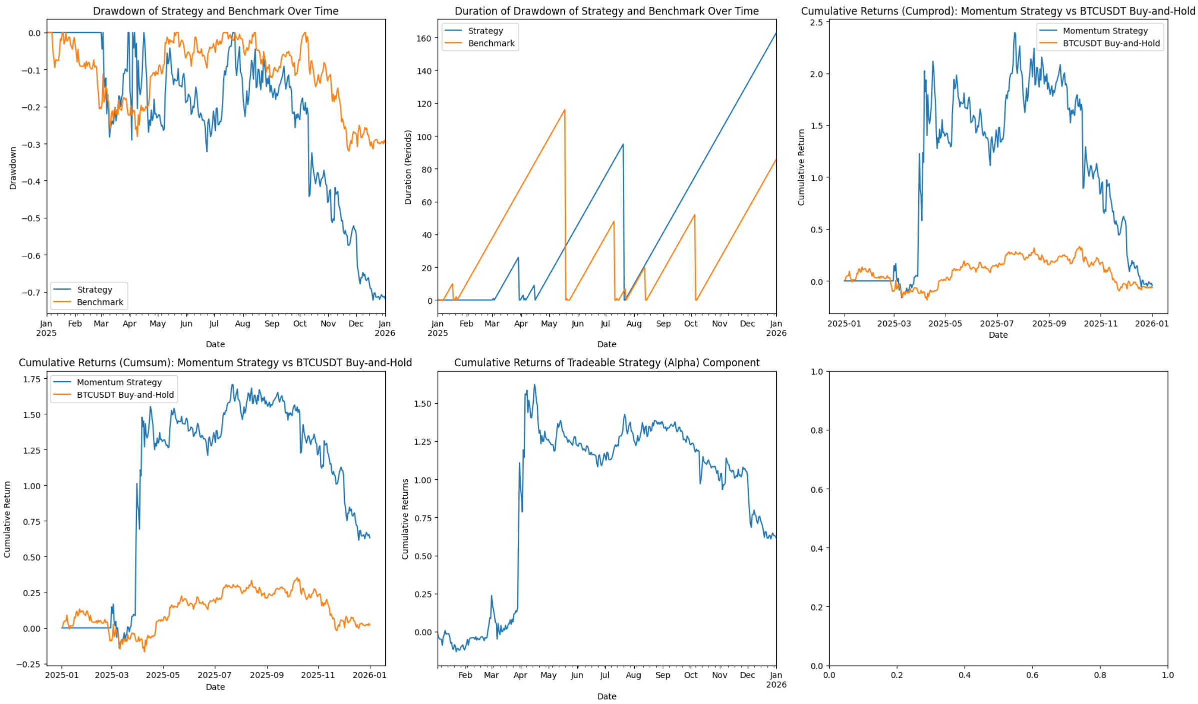
4.4 Strategy Improvement

This is where we introduce our strategy improvement. From the above metrics, we can see that our strategy has a significantly higher annualised volatility compared to the benchmark. That alone will not be ideal as it penalises our risk-adjusted returns. One improvement that could be done is to employ an equal volatility portfolio, such that the volatility weight of each cryptocurrency contribution is the same. Running the **cross-sectional equal-volatility rolling mean** momentum strategy on our training set yielded the updated Sharpe Ratio heatmap.



Although we can see that the optimal (rw_t, hp_t) pair is $rw_t = 11$ and $hp_t = 64$, we ran it on our test data and yielded a worst performance than the previous iteration. Next we looked to our second most optimal (rw_t, hp_t) pair, which is $rw_t = 26$ and $hp_t = 57$. This pair is different from the SR optimal pair, in that our rolling windows are significantly different. Having a rolling window with 11-days lookback may capture too much noise and lead to volatile swings in the short term. Therefore, our next optimal value of 26-days lookback would be better in that it smooths out the noise in the short term.

Running this **cross-sectional equal-volatility rolling mean** momentum strategy on our test data, we suddenly see a reversal and outperformance against the benchmark across most metrics.



Performance Metric	Strategy	Benchmark
Annualised Return	0.543730	0.014711
Annualised Volatility	1.018845	0.346389
Sharpe Ratio	0.426538	0.042162
Hit Rate	0.415301	0.500000
Max Drawdown (%)	72.228194	32.022515
Max Drawdown Duration (periods)	163.000000	116.000000

Our strategy generated positive excess returns with a higher volatility and a longer maximum drawdown (period and depth).

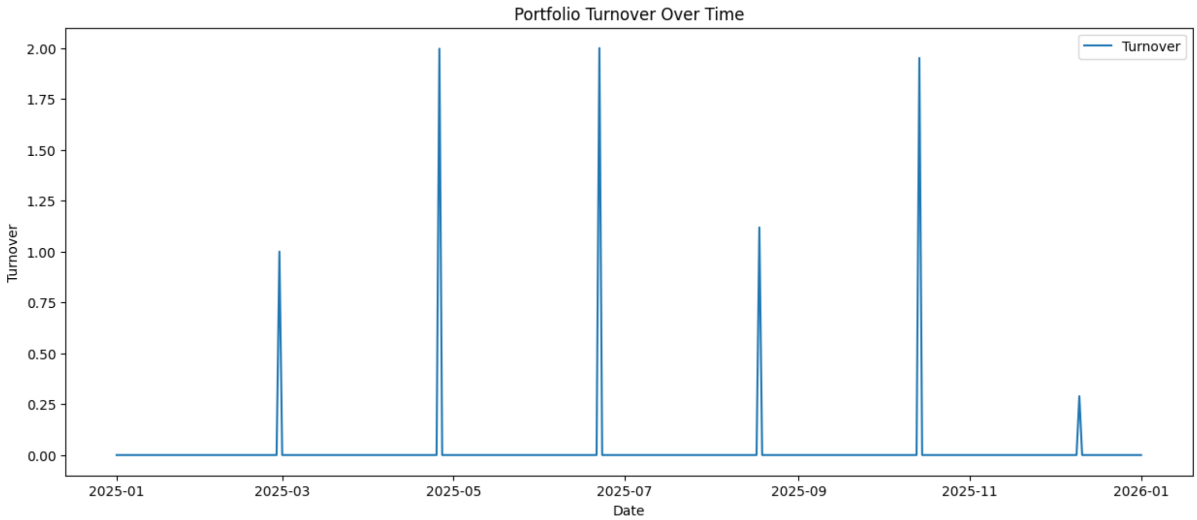
Performance Metric	Strategy
Alpha	0.00104734
Beta	0.9992
Information Ratio (IR)	0.432190
Alpha t-stat	0.520136

While the incorporation of an **equal-volatility** component improved our annualised return and Sharpe Ratio, we can still our strategy's annualised volatility increasing rather than decreasing (which contradicts our idea that equal-volatility portfolio could translate to lower volatility across the portfolio).

4.5 Turnover and Net Returns of Strategy

It is crucial that we evaluate our strategy on the basis of constrained backtesting, such as the incorporation of execution costs (slippages and commission fees, so as to develop a more realistic view of the real trading landscape.

Assuming all orders are executed at the market price, and using an estimate value of 0.2% transaction cost for market orders, we can see the true performance of our strategy below.



The spikes happened over a period of time because of the holding period feature in our strategy, where portfolio rebalancing only happens on a certain date.

Performance Metric	Strategy
Average Turnover	0.0228326
Gross of Transaction Cost Returns	0.426538
Net Return after Transaction Cost	0.3144565
Annualised Sharpe after Transaction Cost	0.415265

This implies that our strategy generated an average of 31.4% after factoring in transaction costs over the backtesting period. Our strategy has a relatively low turnover of 0.0228, which means that on average, approximately 2.3% of the portfolio is traded per period.

4.6 Strategy Conclusion

There could be several explanations as to why our **cross-sectional equal-volatility rolling mean** momentum strategy's performance may not be that superior.

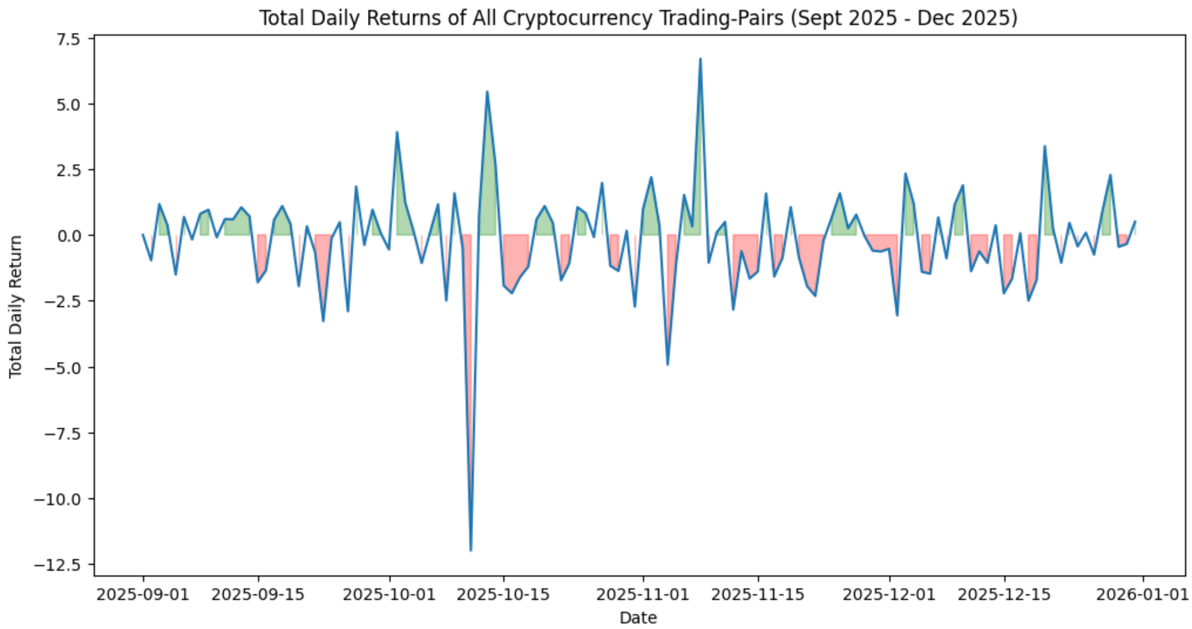
1. Long-only Portfolio

By only going LONG on our cryptocurrencies, we fail to potentially capture returns when the cryptocurrencies' prices are falling, since momentum goes both ways. Furthermore, being fully-invested means that our portfolio will only benefit when the cryptocurrencies rise, and suffer when they decline (since we are still having a long position for the duration of the holding period).

2. Holding-Period Constraint

We see that on average, our optimal holding period is approximately 2-months, while some academic research showed that popular backtesting strategies have holding periods of approximately 1-month. This is not to say that our strategy will underperform, but having a long holding period, especially when the market is volatile, will not be very ideal.

A good explanation can be found in the chart below, where we see consistent spikes up and down across the duration of analysis. This means that having a long holding period would mean we may potentially miss out on profits from certain spikes in the upward direction, and even noting the possibility of rebalancing / reallocating our portfolio in a market downturn.



We can also see once again, that a Long-only portfolio will not benefit as much as one that is both Long and Short. Therefore, we will be exploring our next momentum strategy that seeks to trading on crytpocurrencies in market upturns and downturns.

5. Quantitative Strategy 2

5.1 Strategy: Moving Average Cross-Over Equal Volatility Strategy

Our second strategy is called the **Moving Average Cross-Over Equal Volatility Strategy**, a momentum-based strategy that plays on both the long and short position of assets in our portfolio. Essentially, we will be employing two simple moving averages with different lookback periods. Our signals will be generated when the short term moving average (STMA) crosses the long term moving average (LTMA). The rules for asset positioning will be as follow:

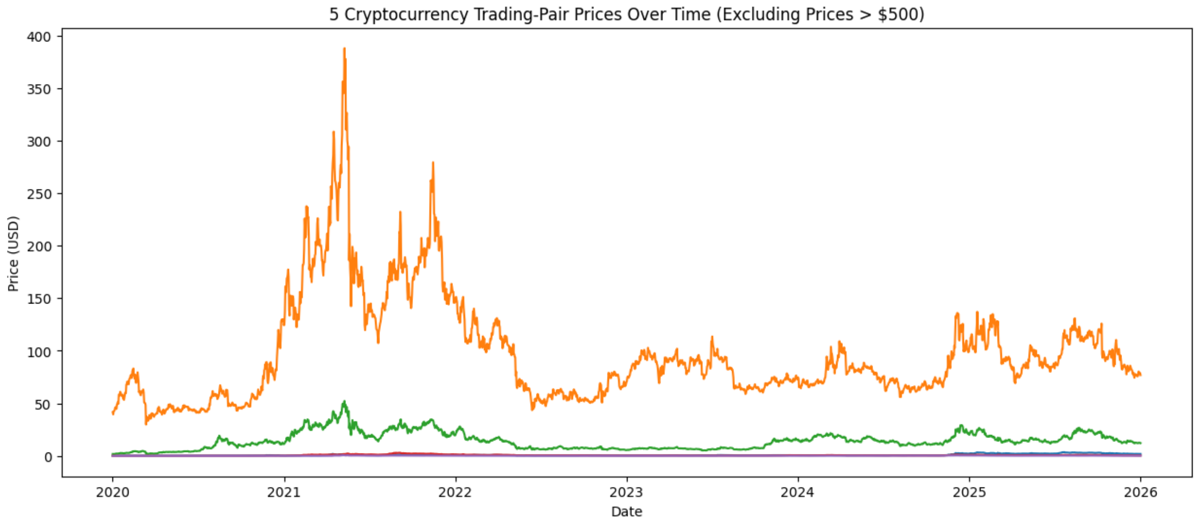
- If STMA crosses above LTMA, we will LONG the asset
- If STMA crosses below LTMA, we will SHORT the asset (*strategy improvement section*)

This is similar to the first strategy above, in that we momentum in assets over the short term to medium term. When the asset performs well in the short term (i.e its STMA goes up), we can expect its outperformance to last for some time before retracing, and vice versa.

5.2 Approach

For this strategy, we will just be focusing on 8 cryptocurrency trading pairs that are widely traded on the exchanges, and contain sufficient data across the duration of analysis.

BTCUSDT, ETHUSDT, XRPUSDT, LTCUSDT, LINKUSDT, ADAUSDT, BNBUSDT, DOGEUSDT



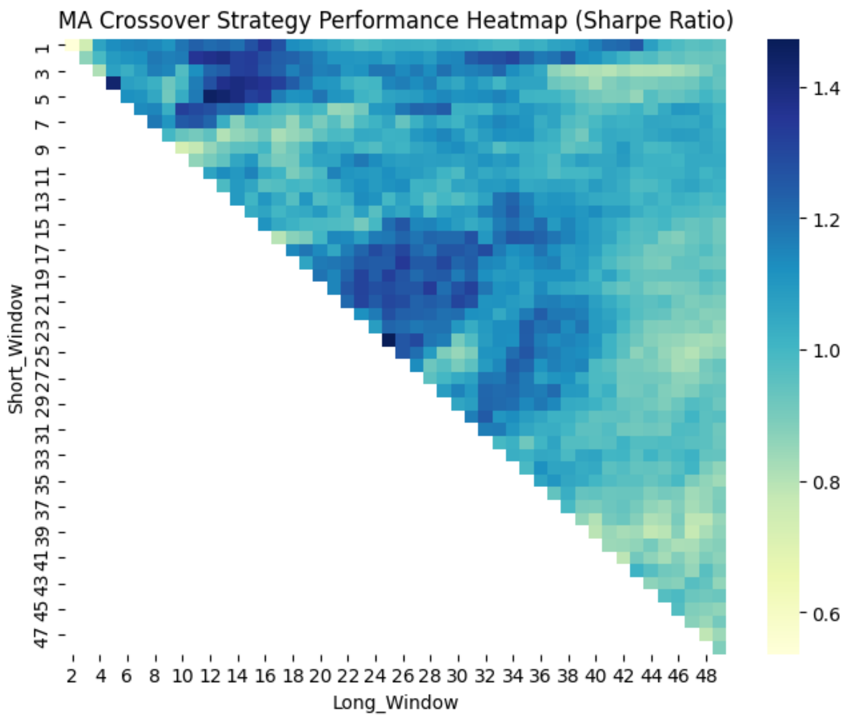
Once again, we can visualise the cryptocurrency prices over the duration of our analysis above.

For our first iteration of this MA cross-over strategy, we will only LONG assets. To begin, we need to identify the ideal lookback window (rw) for our STMA and LTMA, which we will denote as $(STMA_i, LTMA_i)$. Instead of having a fixed holding period, the duration of being LONG an asset is contingent on the MA crossover happening. For our first iteration, we will be assuming equal dollar-weight all LONG positions, and portfolio rebalancing happens every time an MA cross-over (i.e signal) is generated. This means that any LONG position that we currently have in our portfolios will be completely reevaluated during rebalancing, which again could be challenging given that we are always trying to execute at market prices (i.e market order instead of limit order).

For our grid-search, we will be iterating over a period of 50 days for our STMA and LTMA. For each pair $(STMA_t, LTMA_t)$, where t refers to the number of days, we will be calculating the portfolio returns of our LONG-only portfolio and derive its Sharpe Ratio, assuming the risk-free return r_f is 0.

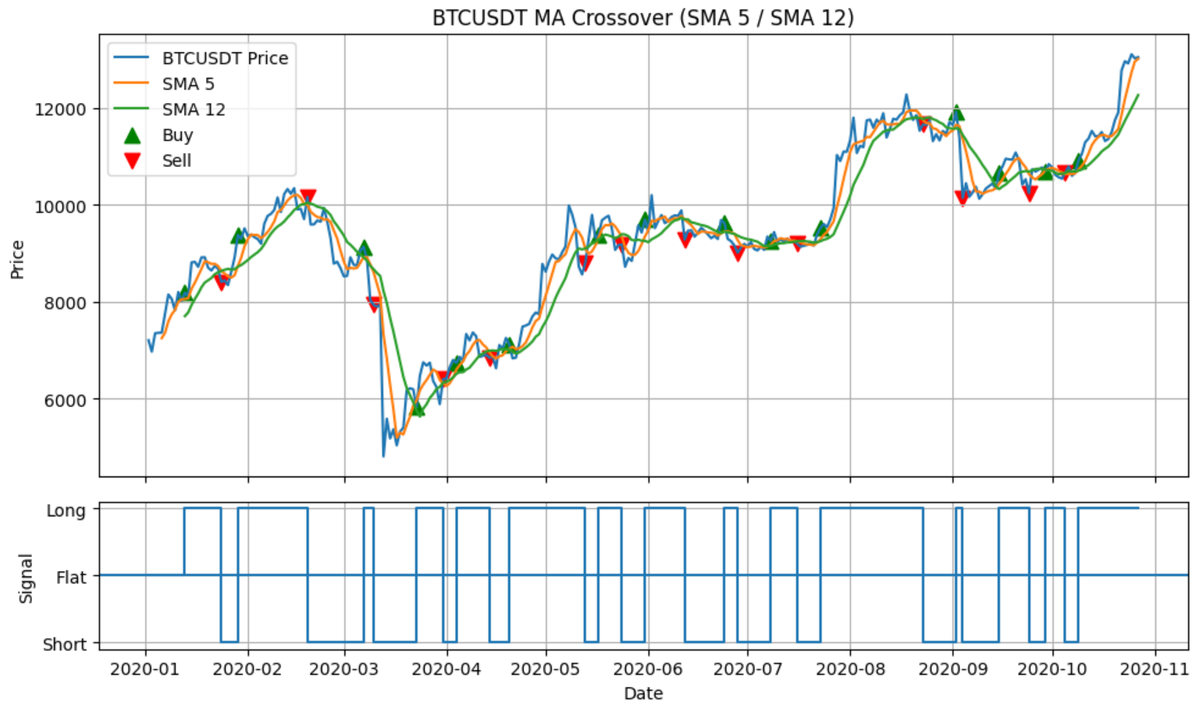
$$Sharpe\ Ratio = \frac{E(r_p)}{\sigma_p}$$

Iterating over each pair (rw_t, hp_t) and plotting the heatmap of its corresponding Sharpe Ratio, we will get the following diagram below.



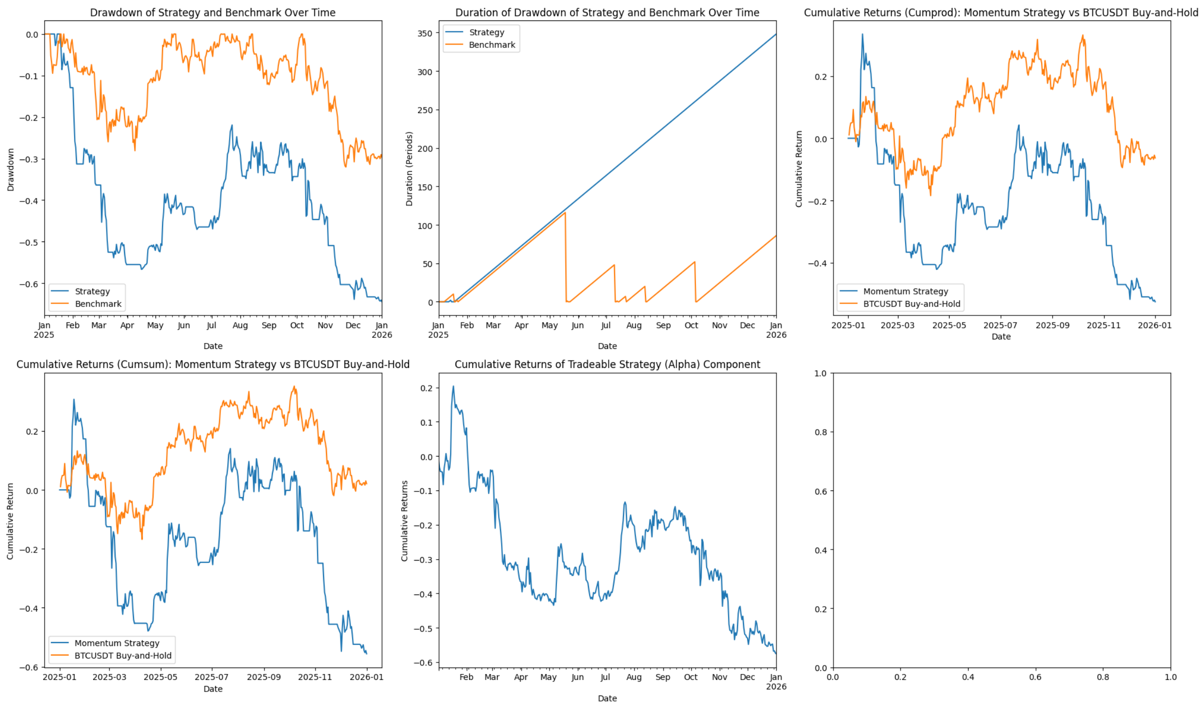
From the heatmap that we plot, we can conclude that our Sharpe Ratio optimal pairs for STMA and LTMA are 24 and 25 respectively. However, since we would ideally like to have a broader range for STMA and LTMA, we have decided to opt for the second Sharpe Ratio optimal pairs with values $STMA_i = 5.0,\ LTMA_i = 12$.

Just for some context, we can visualise one pair of STMA and LTMA over one cryptocurrency across a period of time in the chart below



5.3 Performance Metrics

Running the Sharpe Ratio Optimal values ($STMA_i = 5.0,\ LTMA_i = 12$) for our long-only **Moving Average Cross-Over Strategy**, we got the following results shown below.



Interestingly enough, like our very first rolling mean momentum strategy, this **Moving Average Cross-Over Strategy** significantly underperformed the benchmark in the out-of-sample data.

Performance Metric	Strategy	Benchmark
Annualised Return	−0.318434	0.014711
Annualised Volatility	0.508840	0.346389
Sharpe Ratio	−0.75831	0.042162
Hit Rate	0.346995	0.500000
Max Drawdown (%)	64.467202	32.022515
Max Drawdown Duration (periods)	348.000000	116.000000

Our strategy generated negative returns with a higher volatility and a longer maximum drawdown (period and depth).

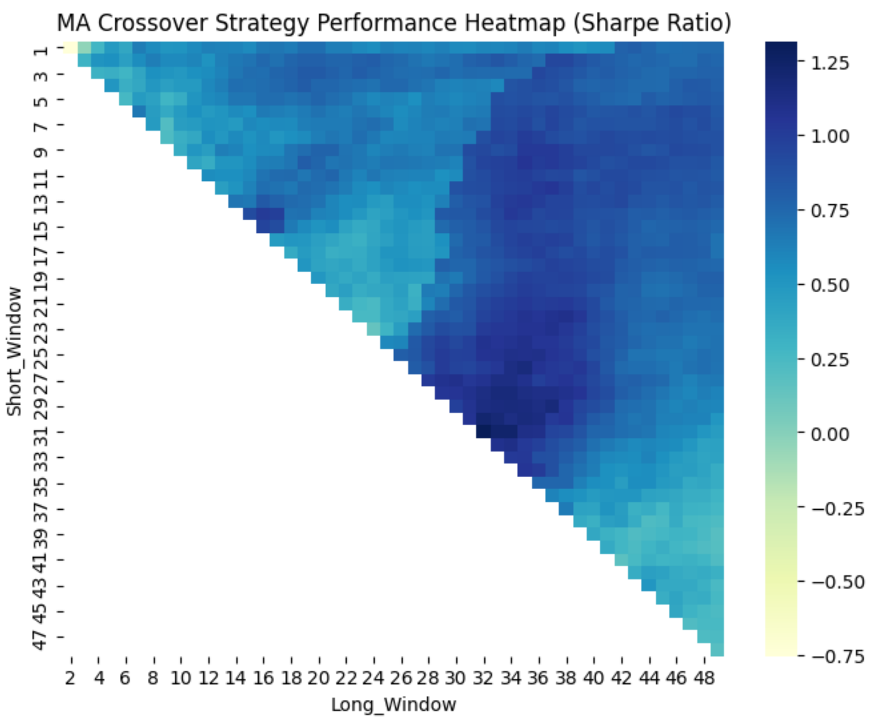
Performance Metric	Strategy
Alpha	−0.001595
Beta	0.9275
Information Ratio (IR)	−1.017913
Alpha t-stat	−1.22504

For our long-only **Moving Average Cross-Over Strategy**, it has an abysmal Sharpe Ratio of -0.758, which underperformed the benchmark's value of 0.0421. Let us consider two improvements that can be made to the MA Cross-over strategy.

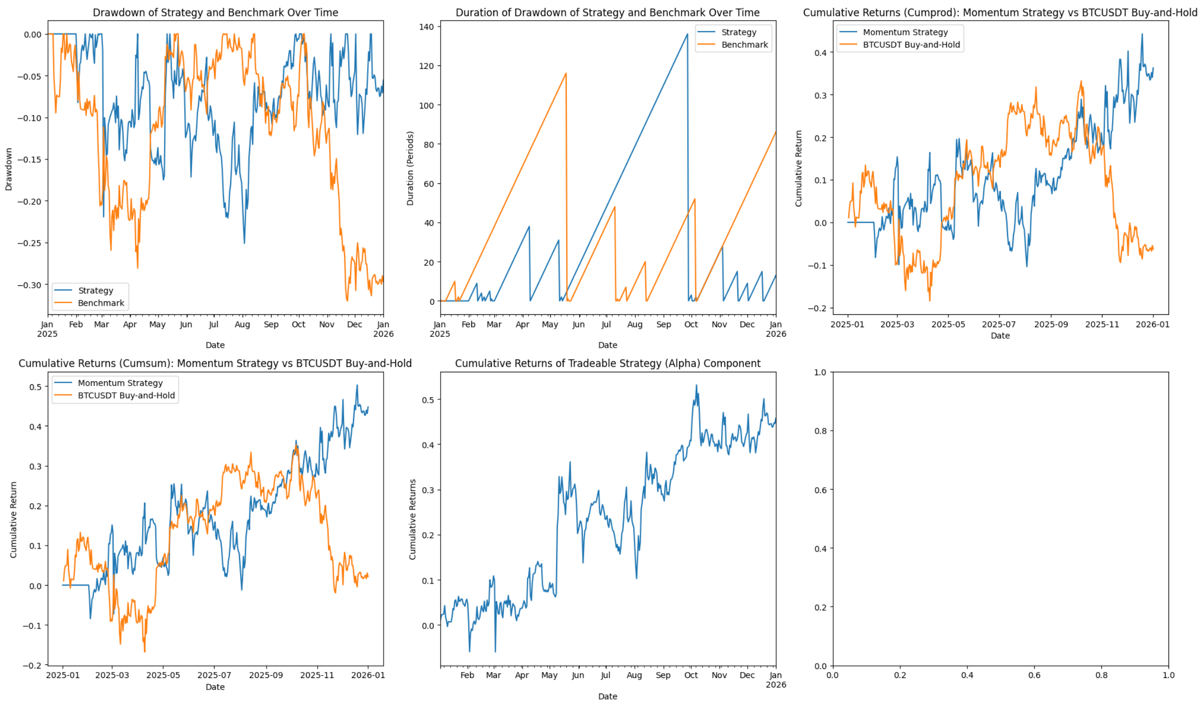
5.4 Strategy Improvements

1. Short Component

Firstly, we can include a SHORT component, from which we will short the cryptocurrency when the STMA crosses below the LTMA. Like the first iteration, portfolio rebalancing will happen whenever the MA crossover happens. Running the grid search on our training data yielded the following heatmap with the optimal short and long windows being ($STMA_i = 31,\ LTMA_i = 32$).



Rerunning on our test data yielded a completely different picture, where we see our strategy significantly improving as compared to a long-only portfolio.



Performance Metric	Strategy	Benchmark
Annualised Return	0.360902	0.014711
Annualised Volatility	0.433643	0.346389
Sharpe Ratio	0.711037	0.042162
Hit Rate	0.491803	0.500000
Max Drawdown (%)	25.120015	32.022515
Max Drawdown Duration (periods)	136.000000	116.000000

Our strategy generated positive excess returns with a lower volatility compared to the first iteration, and a lower maximum drawdown (%).

Performance Metric	Strategy
Alpha	0.0012619
Beta	−0.4752
Information Ratio (IR)	0.79255
Alpha t-stat	0.95383

Interestingly enough, we see that our Beta is now negative, and this reflects the nature of having a SHORT component in our portfolio strategy, as this component can improve the portfolio's return when the market is experiencing a downturn. Both our Alpha and Information Ratio have also improved with this SHORT component.

2. Equal Volatility Portfolio

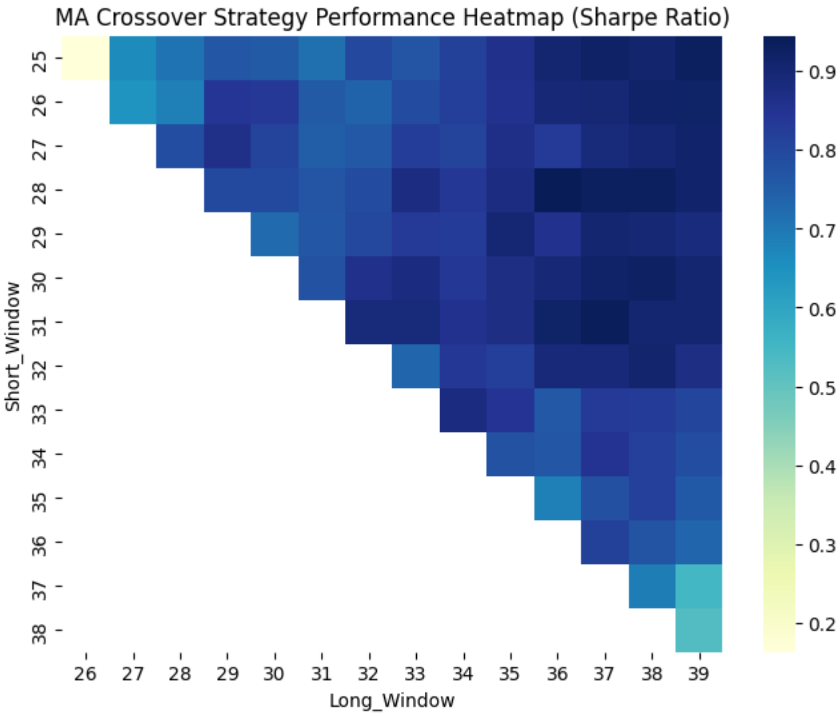
Next up, we can incorporate the equal volatility weighting criterion into our portfolio allocation strategy component. This approach is done in the following manner, where each asset is scaled inversely to its volatility, which creates a risk-parity-like allocation.

$$w_i \propto \frac{1}{\sqrt{\sigma_{ii}}}$$

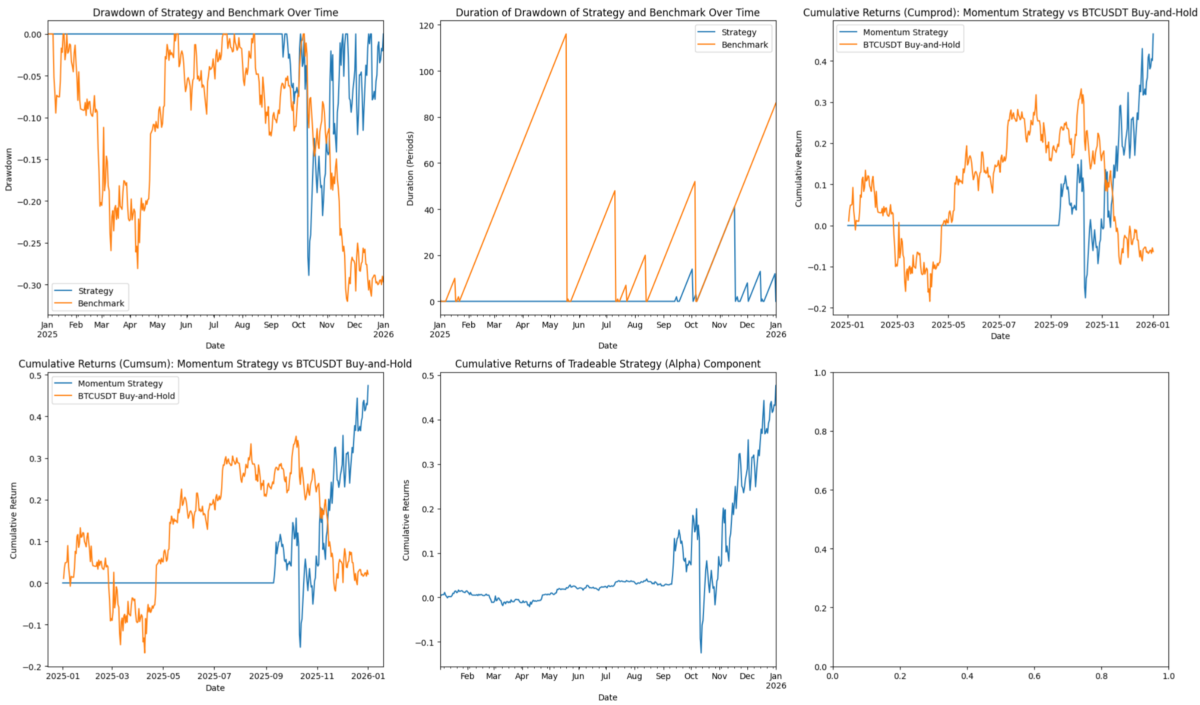
This method balances risk contribution rather than dollar exposure, giving lower-volatility assets more weight. More formally, we can define the normalised form of the inverse volatility weights formula (shown in the Python code below) as follow:

$$w_i = \frac{\frac{1}{\sqrt{\sigma_{ii}}}}{\sum \frac{1}{\sqrt{\sigma_{ii}}}}$$

As the computation with this new approach took up more computing power, we limited the lookback period range to be between 25 and 40. This yielded the heatmap shown below:



Backtesting the Sharpe ratio optimal pair ($STMA_i = 28,\ LTMA_i = 36$) yielded a slightly lower portfolio Sharpe on the test data, but the 2nd most optimal Sharpe ratio pair ($STMA_i = 31,\ LTMA_i = 37$) showed an improvement of our annualised Sharpe and volatility as noted in the table and chart below.



If you notice the flat line at the start of our strategy returns for the test data, this is because we used a look-back period of 252 days (i.e 252-day rolling window) to estimate our asset volatility. Having a sufficiently long and stable historical sample such as a full trading year smooths short-term noise and produces more robust volatility estimates for equal-volatility weighting.

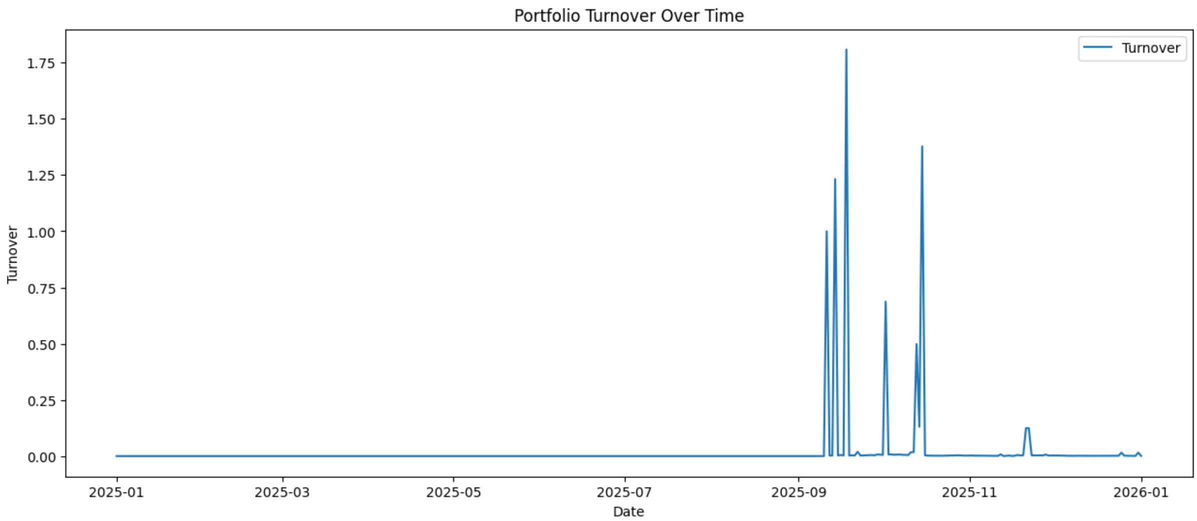
Performance Metric	Strategy	Benchmark
Annualised Return	0.385679	0.014711
Annualised Volatility	0.347875	0.346389
Sharpe Ratio	0.938273	0.042162
Hit Rate	0.185792	0.500000
Max Drawdown (%)	28.909824	32.022515
Max Drawdown Duration (periods)	41.000000	116.000000

Our strategy generated positive excess returns with a volatility comparable to our benchmark, and a lower maximum drawdown (%).

Performance Metric	Strategy
Alpha	0.00130529
Beta	−0.1243
Information Ratio (IR)	0.952866
Alpha t-stat	1.1467629

As shown above, our Alpha and Information ratio also improved with the inclusion of the equal-vol component in our portfolio allocation.

5.5 Turnover and Net Returns of Strategy



The spikes only occurred after a period of time, as our improvised strategy, which takes a look-back period of 252 for computation of asset volatilities, meant that we will only trade after 252 days from the start of the analysis period.

Performance Metric	Strategy
Average Turnover	0.0200406
Gross of Transaction Cost Returns	0.9382732
Net Return after Transaction Cost	0.459389
Annualised Sharpe after Transaction Cost	0.91066

This implies that our strategy generated an average of 45.9% after factoring in transaction costs over the backtesting period. Our strategy has a relatively low turnover of 0.0200, which means that on average, approximately 2..0% of the portfolio is traded per period. Furthermore, our annualised Sharpe after factoring in transaction costs still outperformed the benchmark significantly on the test data.

5.6 Strategy Conclusion

Overall, we see that contrary to the first momentum strategy, the inclusion of a SHORT component can aid in generating returns during market downturns. Furthermore, the hedging effect from having both a LONG and SHORT component means that our portfolio volatility can be minimised (since we are depending our returns on the asset moving only in one direction)

6. Limitations

While the above momentum strategies have generally outperformed the benchmark over the training and testing period (after factoring in improvements), our strategy's resulting Sharpe Ratio is still relatively poor as they are below 1.0, indicating weak risk-adjusted performance. This can be attributed to several structural and methodological limitations, which I have listed below:

1. High Return Volatility and Regime Sensitivity

Momentum strategies are inherently regime-dependent and tend to perform poorly during sideways or mean-reverting market conditions. Sudden trend reversals increase return volatility, which depresses the Sharpe Ratio even when cumulative returns remain positive.

2. No Dynamic Risk or Exposure Control

The strategies does not adjust leverage, position sizing, or exposure based on drawdown conditions. While we may have an equal volatility component in our portfolio allocation strategy, we notice that our strategies still have relatively high drawdowns. As a result, the portfolio maintains similar risk levels across calm and turbulent periods so long as the positions are being held, leading to inefficient risk allocation and lower risk-adjusted returns.

3. Simplistic Signal and Weighting Assumptions

Signals are derived from relatively simple rolling-window momentum indicators and positions are allocated using equal-volatility weighting. While robust and interpretable, this approach does not capture cross-asset correlations, nonlinear effects, or signal confidence, limiting diversification benefits and Sharpe improvement.

7. Future Improvements

While these two momentum strategies have their potential to generate Alpha and a decent Sharpe relative to the benchmark, there are some future considerations which can be expanded to the above strategies:

1. Enhanced Risk Allocation Framework

Move beyond equal-volatility weighting by incorporating correlation-aware methods such as risk parity or minimum-variance optimisation. This improves diversification and reduces concentration risk, particularly during market stress.

2. Downside-Risk-Oriented Optimisation

Optimise for downside risk measures such as maximum drawdown, CVaR, or Sortino Ratio instead of Sharpe alone. This aligns the strategy more closely with capital preservation objectives. For instance, if a drawdown hits a certain level, we can initiate our stop-loss procedure and exit our losing trade to prevent the drawdown from worsening or prolonging.

3. Dynamic Lookback and Holding Periods

Allow rolling windows and holding periods to adapt based on market conditions instead of remaining fixed. This improves responsiveness while avoiding overfitting to a single parameter choice.

8. Conclusion

This project investigated a range of momentum-based trading strategies with systematic portfolio construction and risk controls. The results show that, while the proposed strategies are able to outperform the benchmark in terms of cumulative returns across both training and testing periods, their risk-adjusted performance remains modest, as reflected by Sharpe Ratios below 1.0.

The findings highlight the inherent trade-offs in momentum investing: strong return potential during trending regimes, but elevated volatility and drawdowns during market reversals. Despite enhancements such as equal-volatility weighting and transaction cost adjustments, the strategy's performance is constrained by regime sensitivity, static risk allocation, and simplified signal construction.

Overall, this study demonstrates that momentum strategies can be effective as a return-generating component, but require more advanced risk management, regime awareness, and signal diversification to achieve consistently strong risk-adjusted performance. These results provide a solid foundation for future extensions toward more robust, production-ready systematic trading frameworks.