

15-11-2025

004 Pairs Trading with Cointegration & Kalman Filters

MARKET MICROSTRUCTURE AND TRADING SYSTEMS

José Armando Melchor Soto
Rolando Fortanell Canedo
ITESO

Contents

1. Introduction	3
2. Strategy Overview	4
3. Pair Selection Methodology.....	6
3.1 Correlation	4
3.2 Cointegration	4
3.2.1 Engle–Granger method.....	4
3.2.2 Johansen's method	4
3.5 Statistical evidence	4
3.6 Results	
3.6.1 Price Relationships	
3.6.2 Spread Evolution	
3.6.3 Table	
3.7 Final Pair Selection.....	4
3.7.1 Asset.....	4
3.7.2 Economic Sector	4
3.7.3 Industry	4
3.7.4 Company Overview	4
4. Sequential Decision Analysis Framework.....	10
4.1 Trading Signals	10
5. Kalman Filter Implementation	11
5.1 MLP Architecture.....	11
6. Trading Strategy Logic	13
6.1 Structure.....	13
7. Results and Performance Analysis	17
7.1 Performance Metrics	18
7.2 TRAIN (0.60)	18
7.2.1 Equity Curve	18
7.2.2 Trading Performance Train	18
7.3 TEST (0.20)	18
7.3.1 Equity Curve	18
7.3.2 Trading Performance Train	18
7.4 VALIDATION (0.20)	18

7.4.1	Equity Curve	18
7.4.2	Trading Performance Train	18
7.5	TEST + VALIDATION	18
8.	Conclusions	25
9.	Bibliography	26

Introduction

This project focuses on the development of a statistical arbitrage strategy based on Pairs Trading using cointegration and Kalman filters. The objective is to design, implement, and evaluate a market-neutral quantitative trading system capable of generating consistent risk-

adjusted returns by exploiting mean reversion opportunities between pairs of cointegrated assets. The project applies the Sequential Decision Analysis (SDA) framework proposed by Powell to model the Kalman filters as dynamic decision processes that adapt to changing market conditions in real time.

The strategy systematically identifies pairs of equities with strong long-term equilibrium relationships by performing statistical tests for cointegration. Initially, rolling correlations are computed to screen candidate pairs with correlation coefficients above 0.7. Subsequently, Engle-Granger and Johansen cointegration tests are applied to confirm the existence of a stable long-run relationship, ensuring that the residual spread between assets is stationary.

Once cointegrated pairs are identified, two Kalman filters are implemented. The first estimates the dynamic hedge ratio, allowing for continuous portfolio rebalancing as market dynamics evolve. The second filter generates trading signals by monitoring the spread deviations and detecting mean reversion opportunities. Both filters are formulated as state-space models within the sequential decision process, incorporating prediction, observation, updating, and learning steps. Additionally, a Vector Error Correction Model (VECM) is employed to capture short-term adjustments toward equilibrium, stabilizing the strategy through adaptive signal generation.

The Backtesting framework simulates realistic trading conditions, including a 0.125% transaction cost per trade, 0.25% annualized borrowing rate for short positions, and daily interest accrual. The portfolio allocates 80% of total capital (\$1,000,000) equally across both assets, with dynamic rebalancing of hedge ratios daily. To ensure robustness and avoid overfitting, a walk-forward analysis is conducted, dividing the 15-year historical dataset into training, testing, and validation periods (60%, 20%, 20%, respectively).

Finally, the report provides a comprehensive analysis of the strategy's design, implementation, and performance evaluation. Results are assessed using standard performance metrics such as Sharpe Ratio, Sortino Ratio, Calmar Ratio, and Maximum Drawdown, as well as trade-level statistics including win rate, average trade return, and profit factor. The analysis is complemented by graphical representations of the spread evolution, hedge ratio dynamics, first eigenvector trajectory, and distribution of returns per trade. The overall goal is to demonstrate the feasibility and profitability of combining cointegration-based pair selection with Kalman filter-driven dynamic hedging under a sequential decision-making framework.

Strategy Overview

The strategy implements a market-neutral pairs trading approach that exploits mean reversion between cointegrated equity pairs. Pairs are selected through correlation screening and validated using the Engle-Granger and Johansen cointegration tests to ensure a stable long-term relationship. Once identified, a Kalman filter dynamically estimates the hedge ratio, adapting to time-varying market conditions, while a second filter generates trading signals

based on deviations of the spread from its equilibrium level. A Vector Error Correction Model (VECM) supports the detection of short-term adjustments. Trading positions are opened when the standardized spread (Z-score) crosses defined thresholds and closed upon reversion, maintaining zero net market exposure. The strategy accounts for transaction costs, borrowing rates, and daily rebalancing, with performance evaluated through Backtesting and Walk-forward analysis using standard risk-adjusted metrics such as Sharpe, Sortino, and Calmar Ratios.

Pair Selection Methodology

The pair selection methodology serves as the foundation of the pairs trading strategy, as it determines which assets exhibit a stable and exploitable long-term relationship. This process focuses on identifying asset pairs that move together over time and whose price relationship remains consistent enough to allow for mean reversion opportunities. To achieve this, different statistical techniques are applied, starting with correlation analysis and followed by cointegration testing, to ensure that the selected pair not only shows a strong historical association but also a valid long-run relationship with equilibrium. Once the pair passes these preliminary criteria, more detailed statistical and econometric tests are conducted to confirm its suitability for our trading strategy.

Correlation

In the initial stage of the pair selection process, the goal is to identify asset combinations that exhibit similar price movements over time. For this purpose, the correlation coefficient between assets is calculated to measure the strength and direction of their linear relationship. A correlation value close to +1 indicates that the two assets tend to move in the same direction with high consistency, which is a desirable characteristic for pairs trading strategies.

A high correlation suggests that the assets share common market or sectoral factors influencing their prices, making them more likely to maintain a stable relationship. This step is crucial because a strong correlation increases the probability that the pair will also be cointegrated, meaning their price difference (spread) will revert to a long-term equilibrium over time. Therefore, selecting pairs with high historical correlation serves as an essential first filter before performing formal cointegration tests, ensuring the strategy focuses only on assets that move closely together under similar market conditions.

Cointegration

Engle–Granger method

The simple cointegration method, also known as the Engle–Granger approach, is applied to evaluate whether two asset price series maintain a stable long-term equilibrium relationship. The process begins by cleaning the dataset to remove any missing values and then selecting one asset as the dependent variable (y) and the other as the independent variable (x). A linear regression model is estimated using the Ordinary Least Squares (OLS) method to obtain the interception (α) and the slope coefficient (β), which describe the relationship between both assets.

Once the regression is estimated, the residuals are calculated as the difference between the observed and predicted values. These residuals represent the spread between the two assets. The means of the spread is computed as a reference for its equilibrium level. To verify the existence of cointegration, a stationarity test is applied to the residuals using the Augmented Dickey-Fuller (ADF) test. The key decision rule is based on the p-value obtained from this test:

- If ADF p-value < 0.05 , the null hypothesis of a unit root is rejected, meaning the residuals are stationary, and therefore the two series are cointegrated.
- If ADF p-value ≥ 0.05 , the residuals are non-stationary, indicating that no cointegration exists.

Through this method, the OLS regression combined with the ADF test provides a simple but effective way to detect cointegration between two assets, forming the foundation for constructing a mean-reverting spread used in pairs trading.

Johansen's method

The Johansen method is a multivariate and more robust approach used to confirm the existence of cointegration between non-stationary time series. It is based on the Vector Error Correction Model (VECM), which captures both short-term dynamics and long-term equilibrium relationships. The procedure involves performing an eigenvalue decomposition of the system to obtain eigenvectors that represent possible equilibrium relationships. The first eigenvector defines the strongest and most stable long-term relationship between the assets.

To assess statistical significance, the trace statistic is compared against its critical value. If the trace statistic exceeds the critical value commonly evaluated at the 95% confidence level the null hypothesis of no cointegration is rejected, confirming that at least one cointegrating vector exists. Overall, this method provides a more comprehensive and symmetric validation of cointegration than the Engle–Granger approach, reinforcing the reliability of the pair selection process.

Statistical Evidence

The GOOGL-HD pair exhibits one of the strongest statistical relationships within the tested universe. With a historical correlation of 0.9719, the two assets show highly synchronized price movements over time, despite belonging to different sectors. This high degree of co-movement suggests they respond similarly to broader macroeconomic and consumer-driven market dynamics. The ADF p-value of 0.0140 is below the 0.05 threshold, allowing us to reject the null hypothesis of a unit root in the residuals. This indicates that the spread between GOOGL and HD is stationary according to the Engle–Granger method, confirming the presence of a stable mean-reverting relationship. The Johansen test supports this finding. The trace statistic (22.0955) exceeds the 95% critical value (15.4943), validating the existence of at least one cointegrating vector. The Johansen strength metric (6.60)—computed as the difference between the trace statistic and its critical threshold—shows that GOOGL-HD forms one of the most stable long-term equilibrium relationships in the entire set. In summary, both the Engle–Granger and Johansen methods consistently identify the GOOGL-HD pair as statistically cointegrated. Its combination of very high correlation, strong cointegration metrics, and a stable long-run equilibrium makes GOOGL-HD one of the most economically coherent and reliable candidates for implementing a pairs-trading strategy.

Results

Price Relationships

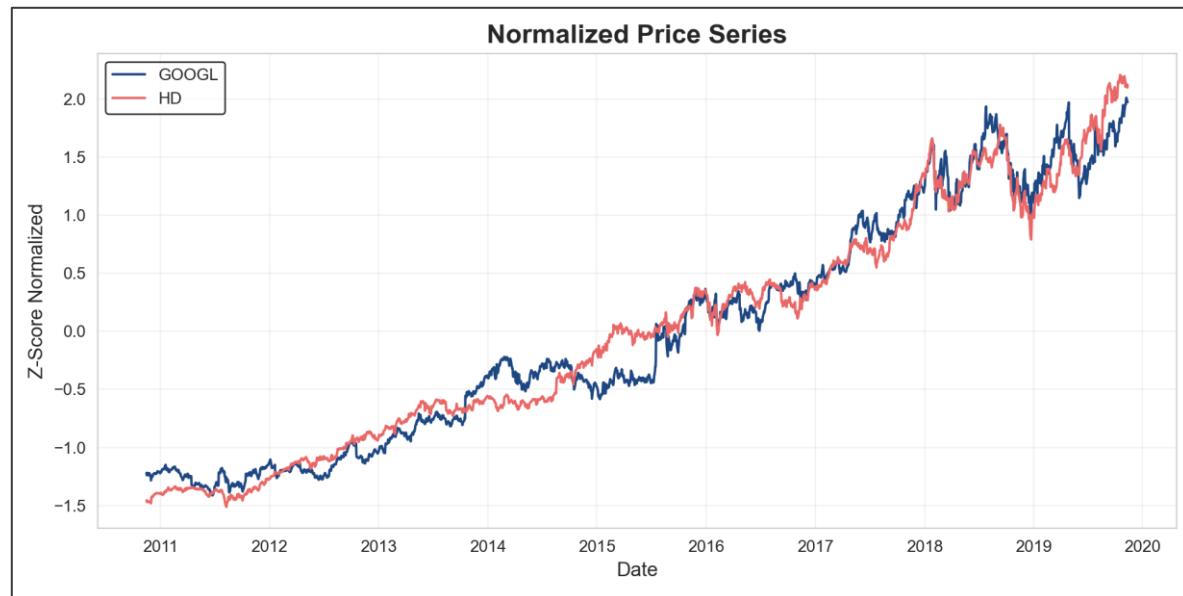


Figure 1. Normalized Prices

Spread Evolution

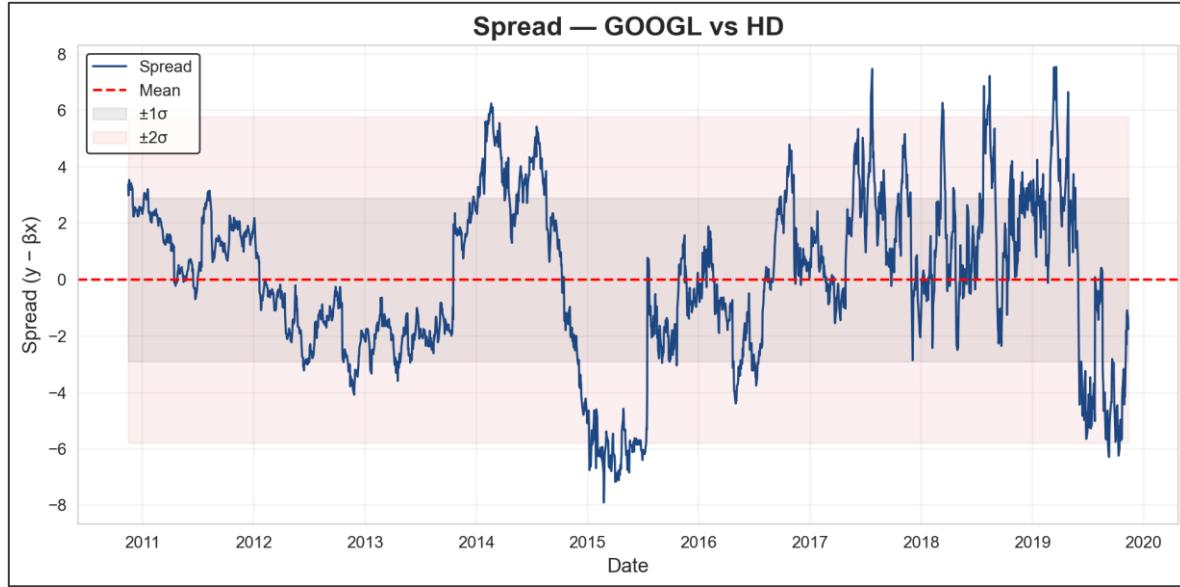


Figure 2. Spread Evolution

The spread evolution between GOOGL and HD shows clear mean-reverting behavior, supporting the cointegration relationship identified earlier. Throughout the period, the spread repeatedly moves away from the mean and then returns, crossing the $\pm 1\sigma$ and $\pm 2\sigma$ bands on multiple occasions. These deviations highlight well-defined trading opportunities, as extended moves beyond the outer bands are often followed by strong reversions back toward equilibrium. Although certain periods, such as 2014–2015, show more extreme divergence, the spread consistently returns to its long-term mean, confirming the stability of the relationship and validating the foundation for a pairs-trading strategy.

Pair Selection Table

PAIR SELECTION													
Asset1	Asset2	Correlation	ADF_pvalue	ADF_Cointegrated	Johansen_stat	Johansen_crit_95	Johansen_Cointegrated	Eigenvector_1	Eigenvector_2	Beta1_norm	Beta2_norm	Johansen_strength	
PFE	BA	0.822454	0.002646	TRUE	16.341964	15.4943	TRUE	0.434068	-0.013796	-31.463382	1	0.847664	
ABBV	CAT	0.825756	0.00569	TRUE	15.781355	15.4943	TRUE	0.096196	-0.060082	-1.601082	1	0.287055	
AMZN	LOW	0.93362	0.006294	TRUE	16.684159	15.4943	TRUE	0.076033	-0.099864	-0.761364	1	1.189859	
JNJ	HON	0.924371	0.008352	TRUE	16.717767	15.4943	TRUE	0.129474	-0.082418	-1.570947	1	1.223467	
GOOGL	HD	0.979173	0.010406	TRUE	22.095528	15.4943	TRUE	0.314046	-0.088092	-3.564971	1	6.601228	
TMO	LOW	0.943926	0.012147	TRUE	15.703646	15.4943	TRUE	0.036814	-0.113524	-0.324284	1	0.209346	
GOOGL	JNJ	0.930247	0.01241	TRUE	18.711339	15.4943	TRUE	0.158284	-0.134662	-1.175416	1	3.217039	
MSFT	TMO	0.989625	0.013112	TRUE	25.437893	15.4943	TRUE	0.146502	-0.08112	-1.805983	1	9.943593	
GS	ABBV	0.755151	0.017087	TRUE	15.923694	15.4943	TRUE	0.047111	-0.069279	-0.680023	1	0.429394	
AMD	TMO	0.942597	0.027163	TRUE	17.17805	15.4943	TRUE	0.130225	-0.034074	-3.821807	1	1.68375	
JNJ	ABBV	0.85283	0.027328	TRUE	15.646011	15.4943	TRUE	0.085573	-0.10408	-0.822184	1	0.151711	

Table 1. Pair Selection

Final Pair Selection

Asset 1	Asset 2
Alphabet Inc (GOOGL)	Home Depot Inc (HD)
Economic Sector	Economic Sector
Communication Services	Consumer Discretionary
Industry	Industry
Interactive Media & Services	Home Improvement Retail
Company Overview	Company Overview

Alphabet Inc. (GOOGL) is one of the world's leading technology companies, primarily driven by digital advertising, cloud services, and information platforms such as Google Search and YouTube. Its performance is closely tied to consumer trends and corporate marketing spending, making it a highly liquid and widely used asset for quantitative research and statistical trading strategies.

Home Depot (HD) is the largest home improvement retailer in the United States. Its business is influenced by consumer spending, housing market conditions, and construction activity. With strong liquidity and relatively stable behavior, HD is a suitable asset for statistical analysis and pairs trading strategies, especially those involving macro-sensitive consumer dynamics.

Economic Relationship

Even though GOOGL and Home Depot belong to different sectors, they react to the same big economic forces. Both depend heavily on the strength of the U.S. consumer and overall economic conditions. When the economy is doing well, companies spend more on advertising (which benefits GOOGL), and consumers invest more in home-improvement projects (which benefit Home Depot). When the economy slows down, both advertising budgets and home-improvement spending usually fall at the same time.

Also, since both are large components of major U.S. indices like the S&P 500, they tend to move together when market sentiment shifts or when there are strong inflows or outflows from index funds. Because of all this, GOOGL and HD end up showing very similar long-term movements, which explains why their economic relationship is stronger than it might seem just from looking at their sectors.

Sequential Decision Analysis Framework

Before describing the Sequential Decision Analysis framework, it is important to clarify the operational structure exactly as it was defined in the configuration of the strategy. The commission rate was set as COM = 0.00125, meaning every transaction (entry or exit, for both legs of the pair) applies a 0.125% cost over the traded notional. The strategy is also restricted by INVEST = 0.80, so now of opening a position, only 80% of the available capital is allowed to be deployed. Short positions incur financing costs based on BR = 0.25% annualized, and since the model evolves day by day, this becomes a daily charge of BR/252 applied to the notion of the short leg. The signals for the strategy rely on Z-score boundaries: ENTRY_Z = 1.0 determines when a spread deviation is large enough to justify opening a long or short spread, EXIT_Z = 0.5 determines when the spread has reverted sufficiently and STOP_Z = 3.5 acts as a safety mechanism to close positions when the deviation becomes extreme. All rolling calculations, both for spread statistics and for cointegration validation, use TDays = 20.

With these parameters in place, the strategy follows the Sequential Decision Analysis (SDA) structure. The model observes daily prices and continually updates its internal estimate of the equilibrium between the two assets. At each step it evaluates risk, applies operational costs, and decides whether to open, maintain, or close positions.

The central element of SDA is the system state, which I represent as

$$St = (\beta_t, s^t, P_t^\beta, P_t^\alpha, spread_{t-20:t}, cap_t, pos_t)$$

Here, β_t is the dynamic hedge ratio estimated by the Kalman filter and governs the equilibrium through

$$Y_t = w_{0,t} + \beta_t X_t + \epsilon_t$$

This β_t is the key parameter that determines the hedge between both legs and is recalculated every single day, meaning the model always trades using the current relationship between the assets. The second component, s_t , is the smoothed version of the cointegration error. Instead of relying on the raw spread

$$s_t = Y_t - \beta_t X_t$$

smoothed value stabilizes the signal and avoids reacting to microstructure noise, which is especially important given that entry decisions use fixed thresholds (ENTRY_Z). The matrices P_t^β and P_t^α reflect the uncertainty in the estimated parameters from both Kalman filters; when these matrices shrink, the model is effectively learning and gaining confidence in the underlying dynamics.

The term $\text{spread}_{t-20:t}$ corresponds to the last 20 smoothed-spread observations, the same window defined by TDays, and is used to compute the rolling mean μ_{20} and standard deviation σ_{20} . These generate the Z-score

$$z_t = \frac{\hat{s}_t - \mu_{20}}{\sigma_{20}},$$

which directly triggers entries, exits, and stop-loss conditions.

The capital component, cap_t , represents the remaining cash after incorporating gains or losses, the commissions deducted using COM, and the borrow costs associated with short exposure using BR/252. Given the constraint INVEST = 0.80, the maximum position size now of entry is computed as

$$n_t = \lfloor \frac{0.80 \cdot \text{cap}_t}{|Y_t| + |\beta_t X_t|} \rfloor$$

Ensuring that position sizing always respects both capital availability and the current hedge ratio. The variable pos_t records the open legs of the trade: for each active position the system stores whether it is long or short, the number of shares, the entry price, and the ticker ("Y" or "X" leg), exactly as handled internally in the backtest.

Given this state, the system chooses between

$$x_t \in \{\text{open long}, \text{open short}, \text{close}, \text{hold}, \text{no entry}\}$$

A long-spread entry means buying asset Y and shorting X when the spread is below equilibrium (negative Z-score), while a short-spread entry does the opposite when the spread is above equilibrium (positive Z-score). The conditions are structured around the thresholds from the configuration: the model opens a position when $|z_t| > \text{ENTRY_Z}$; it closes it once $|z_t|$ falls below EXIT_Z ; and regardless of signal direction, a stop-loss is triggered if $|z_t| > \text{STOP_Z}$. Beyond this, the model requires cointegration to be statistically valid: the rolling ADF test on the smoothed spread must satisfy

$$p\text{-value}_{ADF,t} \leq 0.05$$

Otherwise, entries are blocked until the spread becomes stationary again.

Formally, the SDA transition is expressed as

$$S_{t+1} = S^M(S_t, x_t, W_{t+1})$$

where W_{t+1} contains the new observations for day $t + 1$: the updated prices (Y_{t+1}, X_{t+1}) , the new spread, the ADF result for the rolling window, and the filtered value \hat{s}_{t+1} . The parameters governed by the Kalman filters evolve under the state equation

$$w_{t+1} = Fw_t + q_t, q_t \sim \mathcal{N}(0, Q),$$

and the observation equation

$$y_t = H_t w_t + r_t, r_t \sim \mathcal{N}(0, R),$$

where w_t holds quantities such as β_t or the equilibrium-correction parameter α_t , and H_t corresponds to the observable price information.

On the financial side, capital evolves by marking positions to market, closing or maintaining legs, and subtracting the transaction costs

$$C_t^{(\text{COM})} = \text{COM} \times \text{trade amount},$$

alongside the daily borrow charges

$$C_t^{(\text{BORROW})} = \text{short value}_t \times \frac{BR}{252}.$$

These components directly update cap_t , which then affects future position sizing and decisions. The equity is rebuilt each day by combining updated capital with the current value of long and short legs in pos_t . Through this structure, the model follows the full SDA loop: observing new data, updating its internal state, planning, acting on that decision, and learning for the next iteration, exactly the behavior intended in a Sequential Decision Analysis framework, and fully aligned with the parameters defined in the configuration and the mechanics used in the backtest.

Kalman Filter Implementation

Kalman Filter 1. Dynamic Hedge Ratio

Within the model, the first Kalman filter is the one that keeps learning, step by step, the dynamic hedge ratio β_t . This parameter describes the equilibrium relationship between the two assets at each moment, and the structure behind it is

$$Y_t = w_{0,t} + \beta_t X_t + \epsilon_t.$$

Both the intercept and β_t change over time, and the filter simply updates them every day based on the new price information. The idea is that the relationship between the assets is not fixed; sometimes they move almost one-to-one, sometimes they diverge more, and the filter adjusts that proportion automatically.

The Kalman framework assumes that the parameters evolve smoothly according to

$$w_{t+1} = F w_t + d_t,$$

and that the prices we observe each day relate to those internal parameters through

$$Y_t = H_t w_t + r_t.$$

Here, $H_t = [1, X_t]$ links observation to the internal parameters.

In the backtest, this estimate of β_t is essential for three reasons. First, it defines the dynamic spread used by the strategy:

$$s_t = Y_t - \beta_t X_t.$$

Second, it determines position sizing, because the amount of capital used at entry is constrained by $\text{INVEST} = 0.80$, and the quantity is calculated as

$$n_t = \lfloor \frac{0.80 \cdot \text{cap}_t}{|Y_t| + |\beta_t X_t|} \rfloor.$$

Third, it turns the hedge ratio into a completely adaptive component, updating with the market rather than being a fixed historical estimate. This filter also maintains its uncertainty through the matrix P_t^β , which essentially tells the system how confident it is in the updated value of β_t . Altogether, the first filter forms the structural core of the system's understanding of equilibrium.

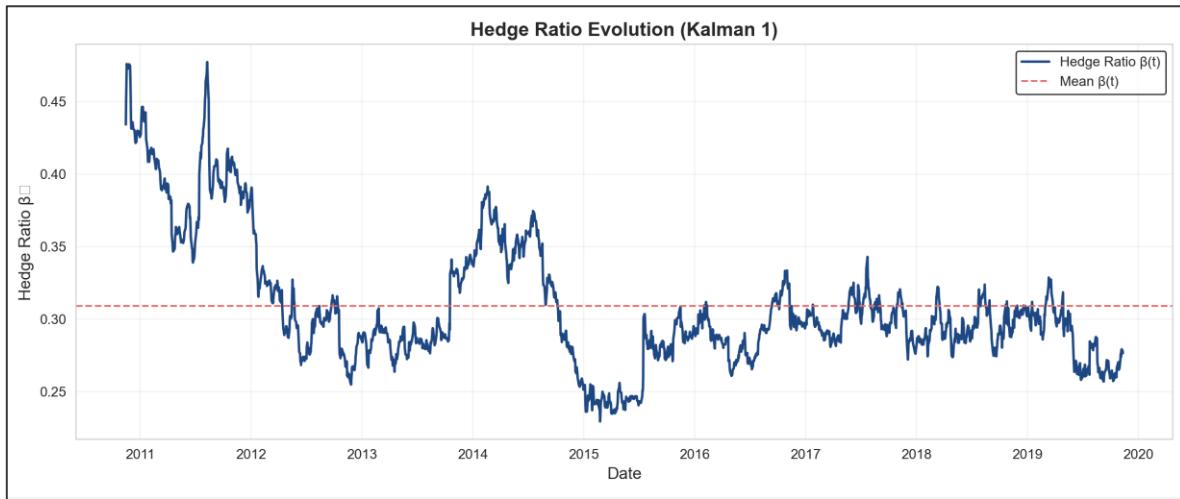


Figure 2. Hedge Ratio Evolution

Kalman Filter 2. Smoothed Spread Estimation

The second Kalman filter works directly on the spread, not to change β_t , but to obtain a smooth and stable version of the cointegration error. Once the first filter provides the dynamic spread, the second filter tries to isolate the “structural” part of that spread and remove daily noise. The output of this filter is \hat{s}_t .

Its internal idea is simple:

$$\alpha_{t+1} = \alpha_t + q_t,$$

so, the equilibrium component evolves slowly, and the daily observation is

$$s_t = \alpha_t + r_t.$$

This means \hat{s}_t becomes a filtered, more reliable version of the spread—precisely the value the system uses to compute the Z-score with your rolling window TDays = 20:

$$z_t = \frac{\hat{s}_t - \mu_{20}}{\sigma_{20}}.$$

This filtered signal is exactly what drives every rule you defined in the backtest:

- Entry when $|z_t| > 1.0(\text{ENTRY_Z})$.
- Exit when $|z_t| < 0.5(\text{EXIT_Z})$.
- Stop-loss when $|z_t| > 3.5(\text{STOP_Z})$.

It also controls when the system is allowed to enter positions, because the rolling ADF test is applied to the smoothed spread, and new entries only occur when

$$p\text{-value}_{\text{ADF},t} \leq 0.05.$$

This ensures the spread is stationary at that moment, which is critical for any mean-reversion strategy. Just like the first filter, this one keeps track of its own uncertainty via P_t^α , letting the system know how stable its signal is. In the sequential structure, this second filter determines the trading signal itself, since without a stable filtered spread the Z-score would be noisy, unreliable, and would generate unnecessary entries and exits. In summary, the first Kalman filter builds the dynamic equilibrium through β_t , and the second Kalman filter produces the clean signal \hat{s}_t . These two components form the core of the learning and decision-making process inside the SDA framework.

Trading Strategy Logic

The trading logic of the strategy follows directly from the idea that the spread between GOOGL - HD exhibits a stable long-run equilibrium but can drift temporarily due to short-term market frictions. Whenever these deviations occur, the strategy attempts to capture the reversion back toward equilibrium, using the dynamic hedge ratio, the filtered spread signal, and the thresholds defined in the configuration. Because all of this is embedded inside Powell's sequential decision structure, each trading decision reflects the updated understanding of the system at that specific moment.

The daily process begins with the construction of the dynamic spread, which uses the hedge ratio estimated by the first Kalman filter. This avoids the typical limitation of static OLS hedging

and ensures that the model always reacts to changes in the relationship between the two assets. The spread is expressed conceptually as:

$$\text{spread}_t = y_t - \beta_t x_t$$

where β_t evolves through the state equation of the Kalman filter. Immediately after this, the second Kalman filter processes the raw spread and produces a smooth error-correction signal. This filtered signal is more stable and removes day-to-day microstructure noise, making it more reliable for generating trading signals.

To determine when a deviation is large enough to justify a trade, the model computes a rolling Z-score using the last 20 filtered observations. The rolling mean μ_t and rolling standard deviation σ_t are estimated over this window, and the Z-score is calculated as:

$$Z_t = \frac{\hat{s}_t - \mu_t}{\sigma_t}$$

This Z-score is the centerpiece of the entire trading logic. When Z_t becomes sufficiently positive, the spread is above equilibrium; when it becomes sufficiently negative, it lies below equilibrium.

A position is opened only when the absolute deviation exceeds the entry threshold, meaning:

- Short spread (sell GOOGL, buy HD) when $Z_t > \text{ENTRY_Z}$,
- Long spread (buy GOOGL, short HD) when $Z_t < -\text{ENTRY_Z}$.

However, not all large deviations are valid opportunities. To avoid entering positions when the long-run relationship temporarily breaks down, the strategy requires the spread to be locally stationary. The model performs an ADF test on the same 20-day window used for the Z-score, and the system allows new positions only when the ADF p-value is below 0.05. This restriction is essential, because large deviations without means reversion would fail the economic logic of the strategy.

Once a position is open, the system monitors the Z-score every day. When the deviation returns toward equilibrium and crosses the exit threshold, the model closes the trade:

$$|Z_t| < \text{EXIT_Z}.$$

Additionally, a stop-loss threshold is imposed to protect the strategy from extreme and unusual divergences. If the Z-score exceeds this safety bound in magnitude (for example, $|Z_t| > 3.5$), the system immediately exits the position, regardless of direction. The strategy also incorporates real-world frictions. Every time the model enters or exits, it pays a commission of 0.125% per leg, which means any round-trip trade (open + close, long and short legs) carries a total cost of 0.50%. Short positions also accumulate a daily financing cost derived from the annual borrow rate of 0.25%, applied proportionally on the exposure of the short leg. These costs directly affect the capital evolution and therefore influence the amount of capital available for the next trade.

Position sizing follows a simple but important rule: the model uses 80% of available capital, split evenly between the two legs to maintain market neutrality. If C_t is the current capital, the capital deployed in each trade is:

$$\text{Capital Used} = 0.80C_t,$$

With 40% allocated to each leg. The number of shares in each asset respects the dynamic hedge ratio, guaranteeing that the structure of the trade reflects the most recent estimate of equilibrium.

The entire process repeats every day following the Sequential Decision Analysis loop. New prices are observed, the Kalman filters update both the hedge ratio and the filtered spread, the Z-score is recalculated, and the system decides whether to open, hold, or close a position. The capital is updated after including profits, losses, commissions, and borrow costs, and the cycle begins again. In this way, the model continuously adapts to market conditions while respecting the economic foundation of cointegration and the operational constraints of realistic trading. In summary, the trading logic combines statistical validation, dynamic state estimation, and disciplined thresholds to structure a fully automated mean-reversion strategy. The approach is coherent with the cointegration evidence, robust under sequential learning, and consistent with market-neutral risk principles.

Results and Performance Analysis

Performance Metrics

To measure the effectiveness of the strategy, the following metrics were used:

- **Sortino Ratio:** It measures how much return is obtained for each unit of downside volatility. A Sortino Ratio greater than 2 is generally considered favorable, as it indicates that the strategy generates strong returns relative to downside risk.

$$\text{Sortino Ratio} = \frac{R_p - R_f}{\sigma_d}$$

Where R_p represents the average portfolio return, R_f is the risk-free rate (assumed to be zero), and σ_d is the standard deviation of negative returns only, not total volatility.

- **Sharpe Ratio:** It measures how much return is obtained for each unit of total volatility. A Sharpe Ratio greater than 1 is generally considered acceptable, indicating strong returns relative to total volatility.

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p}$$

Where R_p represents the average portfolio return, R_f is the risk-free rate (assumed to be zero), and σ_p is the standard deviation of portfolio returns.

- **Maximum Drawdown:** It represents the maximum potential loss that could have occurred during the evaluation period. A lower Maximum Drawdown indicates reduced exposure to significant losses.

$$\text{Max Drawdown} = \frac{P_{\max} - P_t}{P_{\max}}$$

Where P_{\max} denotes the historical peak value of the portfolio up to time t , and P_t is the portfolio value at that same time.

- **Calmar Ratio:** It reflects the relationship between profitability and maximum risk, measuring how much return is obtained for each unit of maximum drawdown. A higher Calmar Ratio indicates stronger returns relative to the maximum drawdown experienced.

$$\text{Calmar Ratio} = \frac{R_A}{\text{Max Drawdown}}$$

Where R_A denotes the annual average return of the portfolio, and Max Drawdown represents the maximum observed decline from a peak to a trough in portfolio value.

- **Win Rate:** It indicates how often the strategy achieves profitable outcomes. A higher Win Rate reflects a greater proportion of successful trades relative to the total number of executed trades.

$$\text{Win Rate} = \frac{\text{Number of winning trades}}{\text{Total number of trades}}$$

Where *Number of winning trades* denotes the total profitable trades, and *Total number of trades* represents all executed trades.

Train (0.60)

Equity Curve



Figure 3. Train - Equity Curve

Trading Performance Train

Tables

TRAIN PORTFOLIO VALUE	
Initial Capital	\$ 1,000,000.00
Final Capital	\$ 881,509.97
Return (%)	-11.85%

TRAIN METRICS	
Metric	Value
Sharpe Ratio	-0.1627
Sortino Ratio	-0.0723
Maximum Drawdown	0.256
Calmar Ratio	-0.0443
Win Rate (%)	4.290%

TRAIN -TRADE STATISTICS	
# Trades	60
Buy	30
Sell	60
Hold	1771
Avg Win	\$ 17,244.40
Avg Loss	-\$ 21,100.00
Profit (\$)	-\$ 118,490.03
Borrow Cost	\$ 1,956.10
Commision Cost	\$ 105,488.68

The training results show noticeably weaker performance compared to the later phases. The strategy finished with a final capital of \$881,509.97, yielding a -11.85% return. Risk-adjusted results were negative, with a Sharpe Ratio of -0.1627, a Sortino Ratio of -0.0723, and a maximum drawdown of 25.6%, reflected in the prolonged decline observed in the equity curve.

The model executed 60 trades (30 buys and 60 sells), along with 1,771 hold periods, indicating constant market activity. However, the Win Rate of 4.29% was too low to sustain profitability, and average losses (-\$21,100.00) exceeded average gains, resulting in a total loss of - \$118,490.03, even before commission and borrow costs.

In summary, the training phase shows that the strategy struggled to maintain capital, faced significant drawdowns, and was unable to generate consistent returns. These results indicate that the model was not yet robust during training and required further adjustment before improving in later stages.

Test (0.20)

Equity Curve

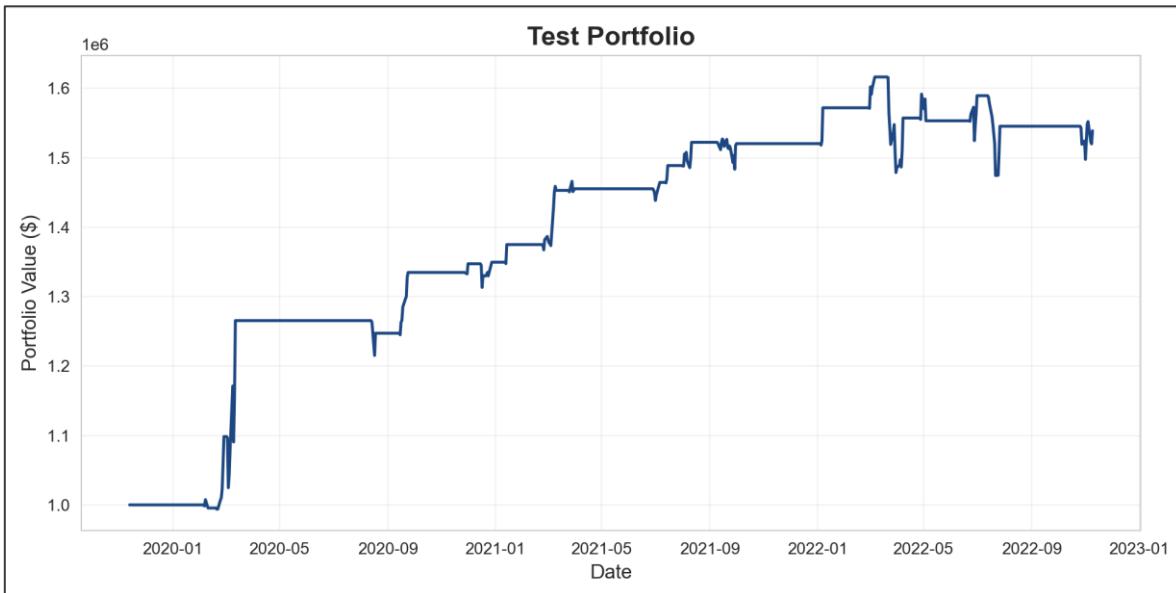


Figure 4. Test - Equity Curve

Trading Performance Test

TEST PORTFOLIO VALUE	
Initial Capital	\$ 1,000,000.00
Final Capital	\$ 1,538,861.00
Return (%)	53.89%

TEST METRICS	
Metric	Value
Sharpe Ratio	1.128
Sortino Ratio	0.7206
Maximum Drawdown	0.0878
Calmar Ratio	1.8876
Win Rate (%)	0.1009

TEST - TRADE STATISTICS	
# Trades	44
Buy	23
Sell	44
Hold	593
Avg Win	\$ 39,667.74
Avg Loss	-\$ 28,759.60
Profit (\$)	\$ 582,116.00
Borrow Cost	\$ 1,517.54
Commission Cost	\$ 118,942.90

The test results show strong and consistent performance throughout the period. The strategy achieved a 53.89% return, finishing with a final capital of \$1,538,861.00. Risk-adjusted metrics were solid, with a Sharpe Ratio of 1.128, a Sortino Ratio of 0.7206, and a maximum drawdown of only 8.78%, reflecting stable behavior and limited downside exposure in the equity curve.

The model executed 44 trades, including 23 buys and 44 sells, along with 593 hold periods, indicating steady market engagement. Although the Win Rate was just 10.09%, the strategy compensated through strong reward-to-risk dynamics: the average winning trade (\$39,667.74) exceeded the average loss (\$28,759.60), generating a total profit of \$582,116.00 even after borrow and commission costs. In summary, the strategy demonstrated robust profitability during the test phase, combining controlled risk, low drawdowns, and effective trade management, while performing well despite a low proportion of winning trades.

Validation (0.20)

Equity Curve

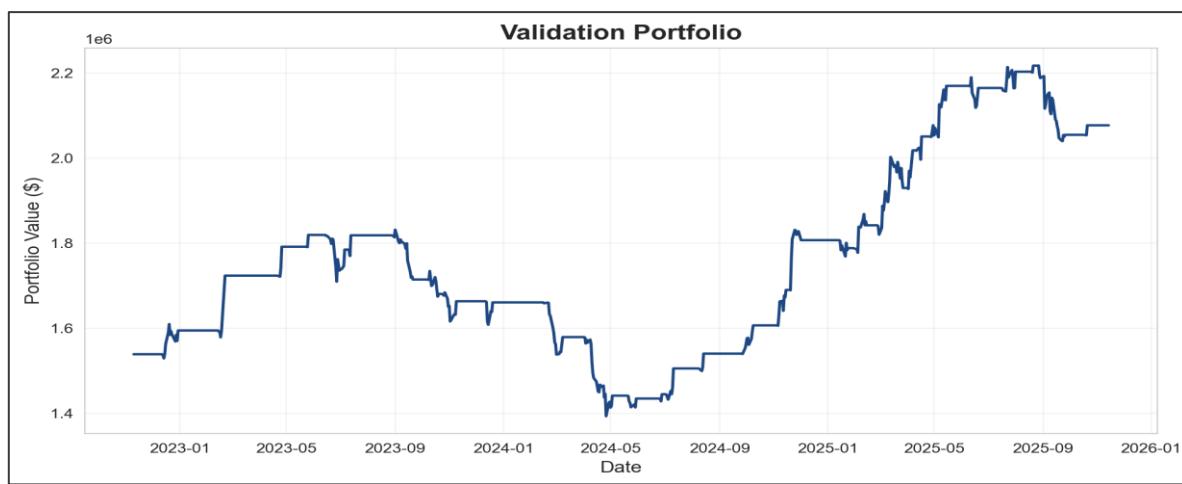


Figure 5. Validation - Equity Curve

Trading Performance Validation

VALIDATION PORTFOLIO VALUE	
Initial Capital	\$ 1,538,861.00
Final Capital	\$ 2,077,107.00
Return (%)	34.98%

VALIDATION METRICS	
Metric	Value
Sharpe Ratio	0.8557
Sortino Ratio	0.9342
Maximum Drawdown	0.2394
Calmar Ratio	0.4763
Win Rate (%)	0.1589

VALIDATION -TRADE STATISTICS	
# Trades	62
Buy	31
Sell	62
Hold	600
Avg Win	\$ 40,131.76
Avg Loss	-\$ 42,607.30
Profit (\$)	\$ 585,169.60
Borrow Cost	\$ 2,759.51
Commission Cost	\$ 181,049.30

The validation results show a solid improvement compared to previous phases. The strategy achieved a 34.98% return, ending with a final capital of \$2,077,107.00. Risk-adjusted metrics were acceptable, with a Sharpe Ratio of 0.8557, a Sortino Ratio of 0.9342, and a maximum drawdown of 23.94%, reflected in the mid-period decline seen in the equity curve.

The model executed 62 trades (31 buys and 31 sales), along with 600 holds periods, showing consistent market participation. Despite a low Win Rate of 15.89%, the average winning trade (\$40,131.76) was strong enough to offset losses, resulting in a total profit of \$585,169.60, even after borrow and commission costs. In summary, the strategy demonstrates that it can generate profitable results with effective risk control, recovering from drawdowns and performing well even with a low percentage of winning trades.

Test + Validation

Equity Curve

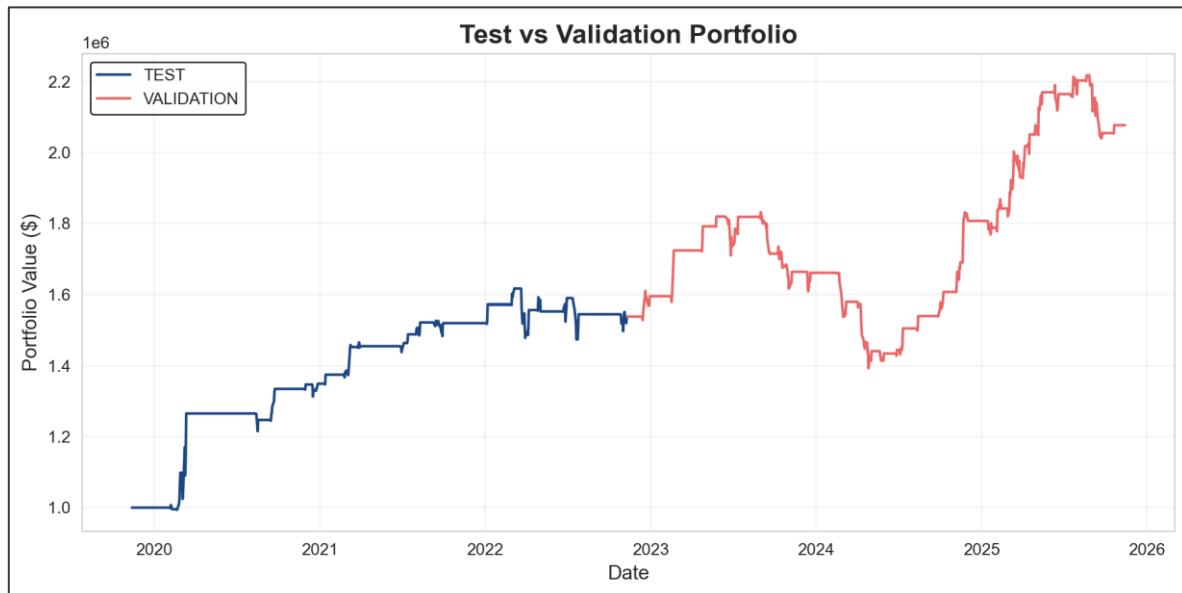


Figure 6. Test + Validation - Equity Curve

The combined test and validation results show a strong and consistent upward trend in portfolio value across both phases. During the test period, the strategy displays steady growth with controlled drawdowns, indicating that the model successfully identified profitable trading opportunities and maintained stability throughout different market environments. In the validation phase, the portfolio continues to increase, reaching new heights despite experiencing temporary pullbacks. The sharp recovery after each decline demonstrates good adaptability and the ability to respond effectively to changing market conditions. The sustained upward trajectory suggests that the model preserved its predictive capacity beyond the test sample. In summary, the strategy exhibits solid profitability, resilience during market

fluctuations, and robust generalization, making it a strong candidate for real-world implementation under similar market regimen.

Portfolio Value

Equity Curve

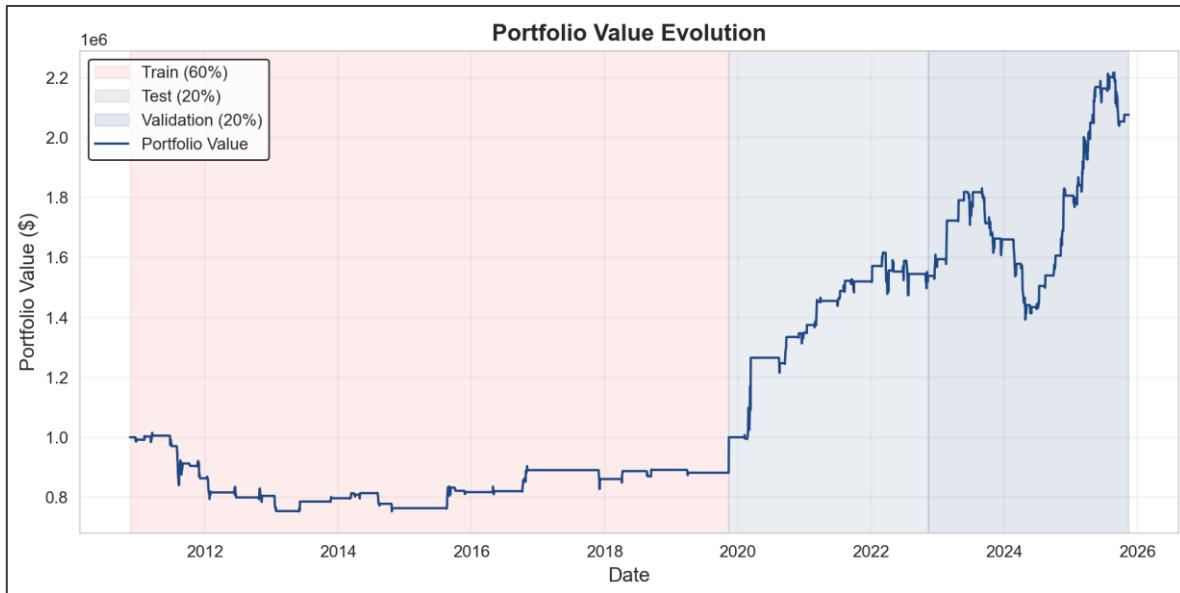


Figure 7. Portfolio Value Evolution - Equity Curve

The full portfolio evolution shows a clear contrast between the training period and the later phases of the strategy. During training, the portfolio experienced a prolonged decline, ending below its initial capital and reflecting the weaker performance observed earlier. However, once entering the test phase, the strategy began to show consistent improvement, with steady growth and controlled drawdowns. This upward trajectory continued into the validation period, where the portfolio reached its highest values, demonstrating strong recovery, resilience, and the ability to capture profitable opportunities across different market conditions. In Summary, the combined progression illustrates how the strategy evolved from an underperforming system during training into a more robust and profitable model in the test and validation stages.

Distribution of Returns per Trade

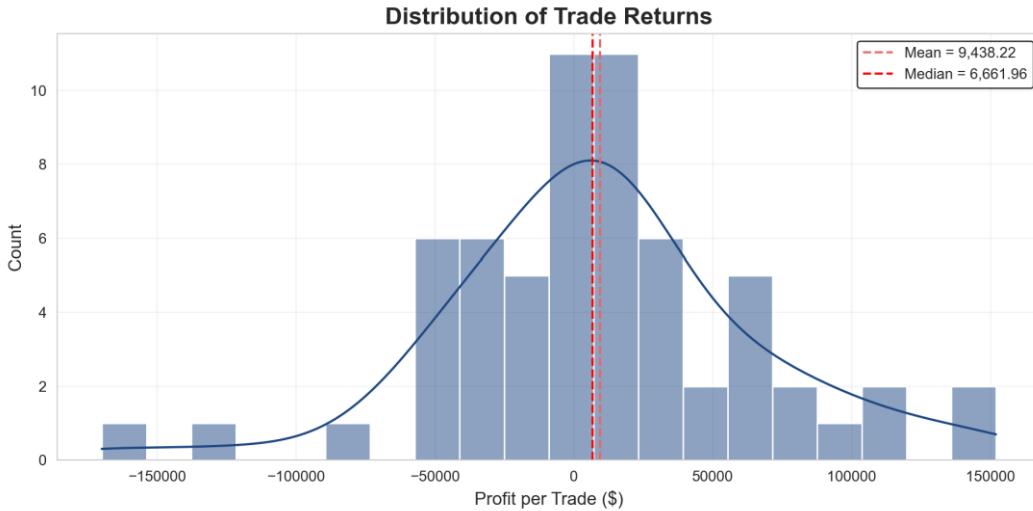


Figure 8. Distribution Returns per Trade.

The data show that most trades fall near the center of the distribution, with many outcomes between $-\$20,000$ and $+\$20,000$, but the right tail includes several large winners above $\$80,000$ and reaching up to $\$150,000$. This skew is reflected in the mean of $\$9,438.22$ and median of $\$6,661.96$, both positive despite the low win rate. In summary, the distribution confirms that a small number of high-profit trades compensate for the more frequent smaller losses, driving the strategy's overall profitability.

Conclusions

In conclusion, the strategy shows that combining cointegration, Kalman filters, and the SDA framework can create an adaptive market-neutral system capable of exploiting stable mean-reversion dynamics. The GOOGL-HD pair displays a strong long-run equilibrium, and although the model performs weakly during training, it becomes consistently profitable in the test and validation periods, even after including transaction costs and borrow fees. This indicates that the strategy can remain profitable under realistic conditions, mainly because a few large winning reversions compensate for many small losses. However, the approach also has limitations, including a low win rate, sensitivity to regime shifts that may weaken cointegration, and vulnerability to execution frictions such as slippage and spread widening. Fixed Z-score thresholds may also lose effectiveness during periods of changing volatility. Potential improvements include using adaptive thresholds, incorporating regime-detection methods, applying machine-learning models to validate entry signals, and expanding the approach to multiple pairs to increase diversification. Overall, the strategy appears viable and potentially profitable, but it requires additional robustness to handle market shifts and real-world trading constraints.

Bibliography

- OpenAI. (2025). ChatGPT (GPT-5-mini) [Large language model].
<https://openai.com/chatgpt>.
- Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business*, 39(1), 119–138.
- Sharpe, W. F. (1994). The Sharpe ratio. *Journal of Portfolio Management*, 21(1), 49–58.
- Sortino, F. A., & Price, L. N. (1994). Performance measurement in a downside risk framework. *Journal of Investing*, 3(3), 59–64.
- Calmar, T. (1991). The Calmar Ratio: A Measure of Return vs. Drawdown. *Futures*, 20(4), 22–25.
- Maginn, J., Tuttle, D., McLeavey, D., & Pinto, J. (2007). *Managing investment portfolios: A dynamic process* (3rd ed.). Wiley.
- Chan, E. (2009). *Quantitative trading: How to build your own algorithmic trading business*. Wiley.
- Chan, E. (2009). *Quantitative trading: How to build your own algorithmic trading business*. Wiley