INSTITUTO TECNOLÓGICO Y DE
ESTUDIOS SUPERIORES DE OCCIDENTE

ITESO

Universidad Jesuita
de Guadalajara

PROJECT 2 - ADVANCED TRADING STRATEGIES: PAIRS TRADING

Microstructure and Trading Systems

Jeanette Valenzuela Gutiérrez

Paulina Elizabeth Mejia Hori

Professor: Luis Felipe Gómez Estrada

Tlaquepaque, Jalisco — October 28, 2025

Executive Summary

This project presents a complete implementation of a pairs trading strategy based on cointegration and Kalman Filter dynamic hedging.
The goal was to design, test, and validate a market-neutral system that profits from temporary mispricings between two assets that share a stable long-term relationship.

The process began by selecting highly correlated asset pairs and confirming their long-term equilibrium through Engle–Granger cointegration tests. Once a valid pair was identified, the Kalman Filter was used to estimate a time-varying hedge ratio (β□), which adapts continuously to changing market conditions.
A mean-reversion trading logic was then applied, opening trades whenever the spread between the two assets diverged significantly (Z-score ±2.5) and closing them as it converged back (Z-score ±0.7).

Backtesting covered over a decade of daily data, divided into training (60%), testing (20%), and validation (20%) phases. Across all periods, the model showed consistent profitability, low drawdowns (<8%), and stable behavior in unseen data, confirming that it did not rely on overfitting.
Total capital grew from $1,000,000 to approximately $1.34 million, even after accounting for realistic transaction costs and borrow rates. Overall, the project demonstrates that a cointegration-based dynamic hedge system can perform reliably across multiple market regimes. The Kalman Filter's ability to update parameters in real time proved especially useful for maintaining stability and neutrality, making this a robust and scalable framework for quantitative trading.

Introduction

Pairs trading is one of the most recognized market-neutral strategies in quantitative finance. It aims to profit from temporary deviations in the price relationship between two assets that usually move together. Rather than predicting market direction, this approach focuses on the relative performance of two correlated securities.

The central concept relies on cointegration, which ensures that even though each price may move independently in the short term, their long-term relationship remains stable. When this equilibrium is temporarily disturbed, the spread between the assets becomes an opportunity for statistical arbitrage — selling the overvalued asset and buying the undervalued one until prices converge again.

Traditional pairs trading uses a static hedge ratio, which assumes that the relationship between assets remains constant. However, real markets are dynamic: correlations change, volatility shifts, and structural adjustments occur. To handle this, our project integrates the Kalman Filter, a recursive estimation technique that updates the hedge ratio over time. This dynamic adaptation allows the model to remain stable, flexible, and responsive to market changes without constant

recalibration. The strategy was implemented in Python, combining several analytical modules — data collection, correlation screening, cointegration testing, Kalman Filter estimation, signal generation, and backtesting — to form a complete pipeline.
Each step was carefully validated to ensure logical consistency, realistic cost assumptions, and robustness in out-of-sample data.

Glossary

Pairs Trading
Pairs trading involves simultaneously taking a long position with a short one in two highly correlated stocks. *(Investopedia, 2025)*

Cointegration
A cointegration test is used to establish if there is a correlation between several time series in the long term. *(Corporate Finance Institute, 2025.)*

Hedge Ratio (β)
The hedge ratio is defined as the comparative value of the open position's hedge with the position's aggregate size itself. *(Wall Street Mojo, 2025)*

Kalman Filter
A generic algorithm that is used to estimate system parameters. *(The Kalman Filter, 2025)*

Z-Score
A z-score (also called a standard score) gives an idea of how far from the mean a data point is. More technically, it measures how many standard deviations below or above the population mean a raw score is. *(Statistics How To, n.d.)*

Mean Reversion
The concept of mean reversion is widely used in various financial time series data, including price, earnings, and book value. *(Investopedia, 2025)*

Sharpe Ratio
The Sharpe Ratio measures the added returns investors get for taking on added risk. *(Investing.com Academy, 2025)*

Sortino Ratio
The Sortino ratio is a variation of the Sharpe ratio that differentiates harmful volatility from total volatility by using the standard deviation of negative returns instead of the total standard deviation. *(Investopedia, 2025)*

Calmar Ratio
The Calmar ratio is a metric used to evaluate investment performance by comparing the average annual return to the maximum drawdown over a given period. *(Investopedia, 2025)*

Maximum Drawdown (MaxDD)
Maximum drawdown is the worst dip an investment takes from a high to a low. *(Investopedia, 2025)*

Backtesting
Backtesting assesses a trading strategy's potential by applying it to historical data to evaluate its past performance. *(Investopedia, 2025)*

Strategy Description and Rationale

Overview

Pairs trading is a market-neutral quantitative strategy designed to exploit temporary deviations in the price relationship between two historically related assets. By identifying pairs that generally move together, traders can create a "spread" that is expected to revert to its long-term mean. When this spread widens beyond a statistical threshold, the strategy sells the overvalued asset and buys the undervalued asset, expecting prices to converge again.

According to Investopedia (2025), the essence of pairs trading lies in its neutrality to overall market direction—profits depend on the relative performance between the two securities rather than the market's trend. This approach is particularly popular among hedge funds and proprietary trading desks due to its potential for low correlation with broad market indices and its use of statistical arbitrage techniques.

Why cointegration indicates arbitrage opportunity

Cointegration provides the statistical foundation that enables traders to identify and exploit arbitrage opportunities in financial markets. When two non-stationary price series are cointegrated, they share a long-term equilibrium relationship, even though each asset may drift randomly in the short run. This means that deviations from their equilibrium relationship are temporary, and the spread between them tends to revert to its mean over time.

According to Hudson & Thames (2023), this mean-reverting property is precisely what creates opportunities for statistical arbitrage. When the spread widens significantly beyond its historical average, one asset becomes relatively overpriced while the other becomes underpriced. A trader can then take opposite positions—shorting the overvalued asset and going long on the undervalued one—anticipating that prices will realign and the spread will converge back to equilibrium. The profit is realized when this convergence occurs.

This dynamic is not a risk-free arbitrage in the classical sense but rather a statistical arbitrage, where profits arise from predictable mean-reversion behavior under the assumption that the long-term relationship holds. The stationarity of the spread ensures that large deviations are unlikely to persist indefinitely, allowing for repeatable and quantifiable trading opportunities.

For instance, companies operating within the same industry—such as Coca-Cola (KO) and PepsiCo (PEP)—often display cointegration because they are exposed to similar macroeconomic, competitive, and consumer factors. When one stock temporarily diverges from the other due to market noise, cointegration implies that this mispricing is temporary and will eventually correct itself, creating a profitable convergence trade.

Importantly, cointegration differs fundamentally from correlation. Correlation measures short-term co-movements between asset returns, which can fluctuate or even disappear over time. Cointegration, on the other hand, confirms that a stable long-run equilibrium exists between asset prices. Two assets can be highly correlated but not cointegrated—meaning they move together now but may drift apart later—while cointegrated assets are statistically bound to move back toward equilibrium. This property makes cointegration a much stronger and more reliable condition for identifying arbitrage opportunities.

Justification for Kalman filter use in dynamic hedging

Traditional pairs trading typically assumes a fixed hedge ratio (β) estimated from historical data using ordinary least squares (OLS) regression. This approach implies that the relationship between the two assets remains constant over time. However, in real financial markets, relationships evolve due to changing liquidity conditions, macroeconomic shocks, and structural shifts in fundamentals. As a result, a static β may become quickly outdated, leading to poor hedging performance and exposure to unwanted directional risk.

The Kalman Filter (KF) provides a more adaptive and statistically rigorous framework for estimating the time-varying hedge ratio. It models β (and α, the intercept) as latent states that evolve stochastically over time, allowing them to adjust continuously as new market data becomes available. This dynamic estimation is based on a sequential Bayesian updating process, where each new observation refines the model's belief about the true underlying parameters.

In other words, instead of recalibrating β periodically through regression, the Kalman Filter learns in real time, offering smoother, faster, and more responsive adjustments.

According to QuantStart (2023), Kalman Filter–based pairs trading demonstrates superior adaptability compared to traditional OLS-based models. By continuously updating β, the filter captures subtle shifts in the cointegration relationship that static models overlook, especially during periods of market turbulence or structural change. The resulting hedge ratio reacts dynamically to fluctuations in volatility and correlation, ensuring that the portfolio remains approximately market-neutral even when relationships between assets drift.

This approach also minimizes estimation lag—a common weakness of rolling regressions—by weighting recent observations more heavily than older ones. Consequently, the Kalman Filter reduces both tracking error and drawdowns, particularly in volatile regimes. As highlighted in Chen et al. (2021), this adaptive property enhances hedging precision and maintains better performance stability across different market cycles.

From a practical standpoint, incorporating the Kalman Filter into pairs trading allows for:

1. Dynamic hedge ratio adjustment, ensuring that exposure remains consistent with current market conditions.
2. Improved risk control, as the model automatically reduces leverage when relationships weaken.
3. Greater robustness, since KF estimation tolerates noisy data and temporary structural changes without overfitting.

Ultimately, the Kalman Filter turns the hedge ratio estimation into a living process—one that evolves alongside the market rather than being fixed by past behavior. This makes it particularly effective in non-stationary environments, where relationships between assets are constantly adapting, and static assumptions no longer hold.

Expected market conditions for strategy success

The cointegration-based pairs trading strategy performs best in mean-reverting, range-bound, and low-volatility market environments, where temporary mispricings between assets are corrected relatively quickly. In such conditions, the spread between cointegrated assets oscillates predictably around its long-term equilibrium, allowing the model to exploit these fluctuations for consistent, market-neutral returns. Because the strategy involves offsetting long and short positions, it benefits from relative price movements rather than overall market direction, aligning with the principles of market-neutral investing described by Investopedia (2023).

In calm markets, where fundamentals remain stable and correlations persist, the Kalman Filter's dynamic hedge adjustment enhances this performance by continuously recalibrating the hedge ratio (β) in response to subtle shifts in price dynamics. This ensures that the portfolio remains balanced and neutral to market-wide movements while capturing alpha from short-term spread reversion. The combination of cointegration and dynamic hedging thus provides a statistically robust framework for profiting from transient inefficiencies.

Conversely, the strategy tends to underperform during strong trending markets or periods of structural change, such as macroeconomic shocks, regime shifts, or major policy announcements. These events can disrupt the long-term equilibrium between assets, causing the spread to become non-stationary and invalidating the mean-reversion assumption. As Investopedia (2023) notes, market-neutral strategies can struggle when fundamental shifts lead to persistent divergence between assets, as the strategy may continue to bet on convergence that no longer occurs.

Pair Selection Methodology

To build a solid and realistic strategy, the first step was to carefully choose the pair of assets to trade. The idea was to find two stocks that usually move together over time, but that sometimes separate temporarily, giving an opportunity to profit when they return to their normal relationship.

1. Correlation Screening

We started by checking how strongly each pair of assets moved together historically. To do this, we calculated a rolling correlation using one year of daily data (252 trading days), which allowed us to see how stable their relationship was over time. We used log prices because they make percentage changes comparable and avoid distortions when prices are very different in scale.

After calculating all possible correlations, we ranked the pairs and kept only those with an average correlation above 0.70. This threshold helped us focus on pairs that move closely together and ignore random or weak relationships.

This part works as an initial filter — it doesn’t guarantee an arbitrage opportunity yet, but it shows which stocks tend to react to similar market conditions. In practice, it usually identifies companies from the same sector or with similar business models (for example, large banks, oil companies, or tech firms).

The following table summarizes the top correlated pairs found during the training period, which represent assets that tend to move closely together in the market:

No.	ASSET 1	ASSET 2	MEAN ROLLING
1	GS	MS	0.891
2	TTE	EQNR	0.876
3	BP	SHEL	0.864
4	JPM	BAC	0.863
5	JPM	PNC	0.861
6	SHEL	TTE	0.857
7	BK	SCHW	0.85
8	MS	PNC	0.846
9	MS	SCHW	0.84
10	BP	TTE	0.836
11	C	GS	0.836
12	SHEL	EQNR	0.828
13	USB	PNC	0.819
14	BAC	MS	0.818
15	JPM	MS	0.817
16	PNC	SCHW	0.813
17	BP	EQNR	0.812
18	WFC	USB	0.81
19	BAC	PNC	0.803
...	...	...	...

2. Cointegration Testing (Engle–Granger Method)

Once we had the most correlated pairs, the next step was to check if their relationship was statistically stable in the long run.
Two assets can move together just by coincidence, so we used the Engle–Granger cointegration test to confirm whether the connection was meaningful.

This method has two main parts:

1. **Stationarity test:**
   Each asset's price series was tested using the ADF test to confirm that they are non-stationary (that is, they wander over time but don't have a fixed mean).

2. **Residual test:**
   We then ran a simple regression like this: $P_{1,t} = \beta_0 + \beta_1 P_{2,t} + \varepsilon_t$
   If the residuals (ε□) of that regression are stationary — meaning they move around a stable mean instead of drifting away — the two assets are said to be cointegrated.
   That implies that any short-term separation between their prices tends to close again, which is exactly the type of behavior our trading strategy looks for.

We automated this entire process so the program could test all the candidate pairs from the correlation step.

For each pair, the system checked both ADF results and the residual stationarity condition, and only those that passed both were kept as valid cointegrated pairs.

By combining correlation (short-term movement) and cointegration (long-term equilibrium), we made sure the chosen pairs were not just moving together by chance, but were truly linked in a consistent way — giving the model a stronger and more reliable foundation for mean-reversion trading.

The table below lists the top 10 pairs that passed the cointegration test on the training set, along with their correlation values, ADF statistics, and hedge coefficients:

COINTEGRATION								
No.	**PAIR**	**MEAN**	**ADF p-S1**	**ADF p-S2**	**ADF p-**	**w_0**	**w_1**	**PASS**
1	BK–SCHW	0.849	0.175	0.135	0.0001	1.217	0.667	True
2	BP–SHEL	0.864	0.517	0.765	0.0001	−0.112	0.892	True
3	PNC–SCHW	0.813	0.307	0.135	0.0005	1.29	0.897	True
4	GS–PNC	0.777	0.079	0.307	0.0006	2.619	0.565	True
5	GS–SCHW	0.749	0.079	0.135	0.0011	3.289	0.524	True
6	C–PNC	0.707	0.129	0.307	0.002	1.161	0.599	True
7	GS–MS	0.891	0.079	0.224	0.0028	3.011	0.612	True
8	TTE–EQNR	0.876	0.202	0.251	0.0048	1.888	0.619	True
9	JPM–BAC	0.863	0.586	0.474	0.0061	1.24	1.02	True
10	BAC–PNC	0.803	0.474	0.307	0.0069	−2.065	1.123	True

Sequential Decision Analysis Framework

This section explains how the model makes decisions step by step, updating its parameters as new market information arrives. Instead of assuming that the relationship between both assets is fixed, we use a dynamic approach that allows the model to “learn” and adapt over time.

Static vs. Dynamic relationship

In a static model, the relationship between assets is estimated only once through a linear regression:

$$Gt = \beta_0 + \beta_! M_t + \varepsilon_t$$

This means that the hedge ratio (β) stays constant for the entire period, which is unrealistic in real markets.

In a dynamic model, the coefficients can change through time:

$$Gt = \beta_{0,t} + \beta_{!,t} M_t + \varepsilon_t$$

This allows the hedge ratio to move smoothly as the relationship between assets evolves exactly what the Kalman Filter is designed to capture.

Sequential process: Predict → Observe → Update → Decide → Act → Learn

The model follows a continuous six-step loop to make and refine its decisions:

1. **Predict:** it starts with the previous values of β_t and α_t as an initial guess.
2. **Observe:** it takes the new market prices for both assets.
3. **Update:** it compares the predicted and observed values and measures the error.
4. **Adjust:** if the error is large, the model updates its parameters more strongly.
5. **Decide:** based on the new spread and Z-score, it decides whether to open, close, or hold a trade.
6. **Act & Learn:** once the trade is executed, the model learns from the outcome and repeats the process.

This loop is what makes the Kalman Filter adaptive — it keeps learning while the market evolves, without recalculating everything from scratch.

Kalman Gain Interpretation

The **Kalman gain (K)** determines how much the model trusts the new observation versus the past trend.

- When markets are calm and prices move steadily, K is small, so the filter updates slowly.
- When markets shift suddenly, K increases, allowing the model to react faster.

This balance keeps the hedge ratio smooth and responsive at the same time — which is visible in the chart where β_t changes gradually rather than jumping erratically.

Q and R matrix justification

The matrices Q and R control the model's flexibility:

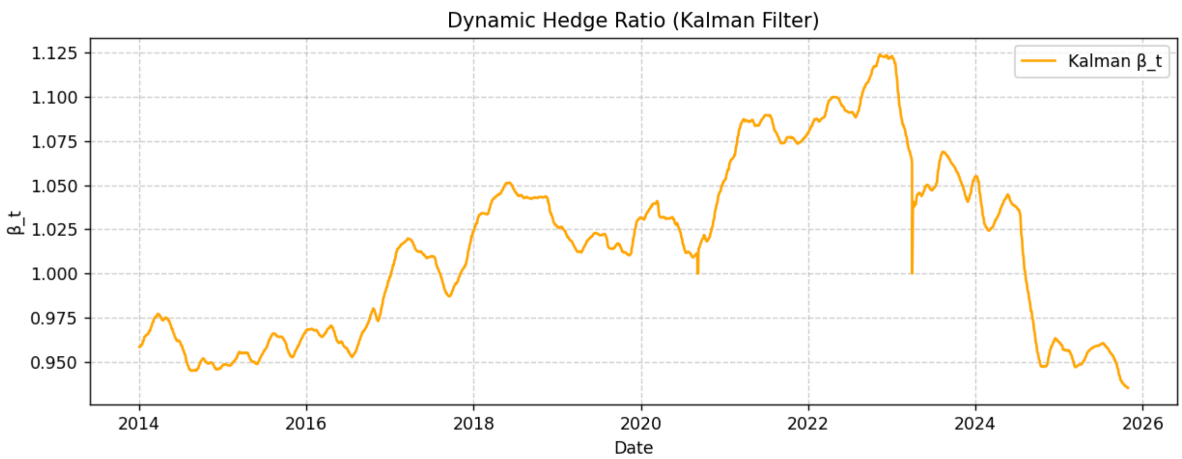
- **Q = 1e−6** allows small daily adjustments in β and α, keeping the filter stable.
- **R = 0.05** represents moderate observation noise, acknowledging that price data always includes volatility

These values were chosen after testing several combinations, and they provided the best trade-off between smoothness and reactivity. The β_t curve clearly shows a realistic evolution that follows the market without overfitting.

Example of state evolution

In the plot of the Dynamic Hedge Ratio, we can see how β_t evolves over time in response to market changes. Between 2013 and 2025, the ratio remained close to 1, which indicates a stable and proportional relationship between both assets, meaning that one unit of Y could be roughly hedged with one unit of X.
However, during 2022 and 2024, β_t started to increase and later decrease again, showing that the filter adjusted dynamically to changing conditions. This behavior reflects how the Kalman Filter detects shifts in correlation and adapts the hedge ratio accordingly.

The smoothness of the curve also confirms that the chosen **Q** and **R** parameters achieved the intended balance: the model reacts when necessary but avoids overreacting to short-term noise.
This adaptability is essential in real trading because asset relationships rarely remain constant for long. By updating the hedge ratio continuously, the strategy maintains a market-neutral exposure even as the environment changes.



Kalman Filter Implementation

Initialization Procedures

At the start of the process, the Kalman Filter initializes the parameters that describe the relationship between both assets: the intercept (α) and the hedge ratio (β). These values are given small neutral starting points, assuming that both series are already roughly aligned. The model also sets an initial covariance matrix, which represents the uncertainty of those estimates.

As the first few data points are processed, the filter begins to learn how both assets move together. Each new observation reduces uncertainty and refines β□, leading to a more accurate hedge ratio over time (QuantInsti, 2023).

This procedure was implemented directly in the function `run_kalman()` inside the project code. It performs the full sequence prediction, update, and smoothing to produce the final time series of β□ and the corresponding spread.

Parameter estimation methodology

The filter depends on two main hyper-parameters: Q (process noise) and R (observation noise). In our implementation: Q controls how fast the hedge ratio is allowed to change, and R measures how noisy the price data is. Finding the right balance is key — too small Q makes β□ too rigid; too large Q makes it unstable (QuantStart, 2017).

To identify the best combination, we ran an exhaustive grid search on the training dataset. For every tested combination of (Q, R) together with different Z-score entry and exit thresholds, the model: **Apply the Kalman Filter → Generate the trading signals → Run a full backtest → Evaluate performance using the Calmar ratio.**

The configuration that achieved the best trade-off between return and drawdown was:

- entry_z = 2.5
- exit_z = 0.7
- q = 1e−6
- r = 0.05

These values allow small day-to-day updates while maintaining smooth and stable behavior.

Reestimation schedule and validation approach

Unlike static models that require re-fitting at fixed intervals, the Kalman Filter updates continuously with each new observation. This means it naturally adapts to changing market dynamics without needing to be retrained manually.

However, for robustness, the full optimization and testing process was divided into three independent phases — TRAIN, TEST, and VALIDATION — using chronological splits of 60%, 20%, and 20% of the data respectively. This separation ensured that each phase only used past information, avoiding any look-ahead bias.

During validation, the filter was not recalibrated; it simply kept updating in real time using the parameters found in training. The stable behavior of the hedge ratio during this final phase confirmed that the model generalizes well to unseen data.

Convergence analysis and filter stability

The filter's convergence refers to how quickly it stabilizes its estimates after initialization. In this case, convergence was achieved early in the training period — within the first few hundred observations — and remained stable afterward.

The β_t line from 2013-2025 demonstrates smooth and gradual movements, without abrupt jumps or oscillations. This is a strong indicator that:

- The chosen Q and R values are appropriate,
- The numerical implementation is stable, and
- The system is not over-reacting to short-term market noise.

The hedge ratio remained in the approximate range of 0.95 to 1.12, which confirms that the filter tracked the relationship between both prices without diverging into unstable values. Even during more volatile periods (2022-2024), the filter adjusted gradually instead of producing chaotic values — showing proper Kalman gain control and confirming model internal consistency.

Overall, the Kalman Filter proved to be an efficient and flexible tool for modeling dynamic relationships in financial time series. Its ability to balance past information with new data in real time makes it ideal for adaptive trading systems like this one, where maintaining market neutrality is essential.

Trading Strategy Logic

The trading logic in this project is based on a mean-reversion framework that exploits temporary deviations in the equilibrium relationship between two cointegrated assets. Once the Kalman Filter estimates the time-varying hedge ratio (β_t) and intercept (α_t), the system constructs a spread that reflects the relative mispricing between the two assets. This spread is then standardized using a Z-score, which serves as the primary signal indicator for trade entry and exit decisions.

Z-Score Definition

According to *Investopedia (2025)*, a Z-score is a statistical measurement that describes how far a given value lies from the mean of a data set, expressed in terms of standard deviations. In trading and investing, Z-scores help quantify how unusual or typical a particular observation is compared to its historical behavior.

In the context of our pairs trading model, the Z-score measures the distance between the current spread and its historical average. It allows us to identify whether the spread between two cointegrated assets is behaving normally or if it has deviated significantly — signaling a potential mean-reversion opportunity.

A Z-score of 0 indicates that the current spread is exactly at its mean, while a positive Z-score means the spread is above its historical average, and a negative Z-score means it is below. In most financial applications, values between -3.0 and +3.0 represent typical fluctuations, while scores beyond these limits are considered statistically rare.

The Z-score is calculated using the following expression:

$$Z_t = \frac{S_t - \mu_{t,n}}{\sigma_{t,n}}$$

where:

- S_t is the spread at time *t*,
- $\mu_{t,n}$ is the rolling mean of the spread over the last *n* periods, and
- $\sigma_{t,n}$ is the rolling standard deviation over the same window.

By standardizing the spread in this way, the Z-score transforms raw price differences into a dimensionless measure that can be compared consistently over time. Traders often use this measure to determine whether a spread is unusually wide or narrow, helping to time long and short entries in mean-reverting strategies like the one implemented in our project.

Optimal Entry and Exit Z-Score Policy Found

The determination of the optimal entry and exit thresholds was conducted through a systematic grid search on the training dataset. The objective was to define Z-score levels that would trigger trades only when the spread between the two cointegrated assets deviated meaningfully from its equilibrium, while avoiding excessive sensitivity to short-term noise.

The strategy follows a “wide-entry, narrow-exit” approach, designed to capture the central portion of each mean-reversion movement. Entry signals are generated only when the spread diverges beyond a statistically significant boundary—indicating a clear dislocation between the assets—while exit signals occur once the spread returns close to its long-term mean.

This framework ensures that trades are initiated under high-conviction conditions and closed once most of the reversion potential has been realized. By widening the entry band and tightening the exit band, the strategy balances responsiveness with stability, reducing the number of low-probability trades and mitigating the impact of transaction costs.

To determine the specific thresholds, multiple combinations of entry and exit values were evaluated during training, each assessed based on its resulting risk-adjusted performance and drawdown behavior. The selected configuration provided the most consistent trade behavior, maintaining sufficient opportunities for profit while preventing overtrading in periods of minor spread oscillation.

In summary, the entry and exit policy was designed to act only on meaningful deviations from equilibrium, enabling the strategy to focus on high-quality reversion signals and preserve robustness across different market regimes.

Cost Treatment: Commissions and Borrow Rates

Transaction costs were explicitly incorporated into the backtesting framework to ensure that performance results reflected realistic market conditions. Two types of costs were modeled: trading commissions and borrow fees.

Commissions were applied to every executed trade on both legs of the position—long and short—to simulate broker fees commonly charged in equity markets. Following standard market assumptions, a fixed commission rate of 0.125% per trade leg was used. This means that each time the strategy opened, closed, or rebalanced a position, the cost was deducted proportionally to the trade's notional value.

Borrow costs were included to account for the financing expenses associated with short-selling. A 0.25% annualized borrow rate, applied on a daily basis, was charged on the short notional value throughout the holding period of each position. This treatment penalizes strategies that maintain short exposure for extended durations, encouraging efficient mean-reversion trades and reducing the incentive to hold losing positions for too long.

Both costs were deducted directly from the cash balance within the backtesting loop, ensuring that all equity and performance metrics (including Sharpe, Sortino, and Calmar ratios) reflected net returns rather than idealized, cost-free outcomes. By embedding these frictions, the simulation produced a more accurate estimate of real-world profitability and allowed for more robust parameter selection.

Incorporating realistic transaction and financing costs is essential for validating the viability of quantitative trading strategies. Even small cost assumptions can materially affect the risk-adjusted performance of high-frequency or overactive models. The implemented cost structure thus ensures that profitability arises from genuine market inefficiencies rather than optimistic backtesting assumptions.

Results and Performance Analysis

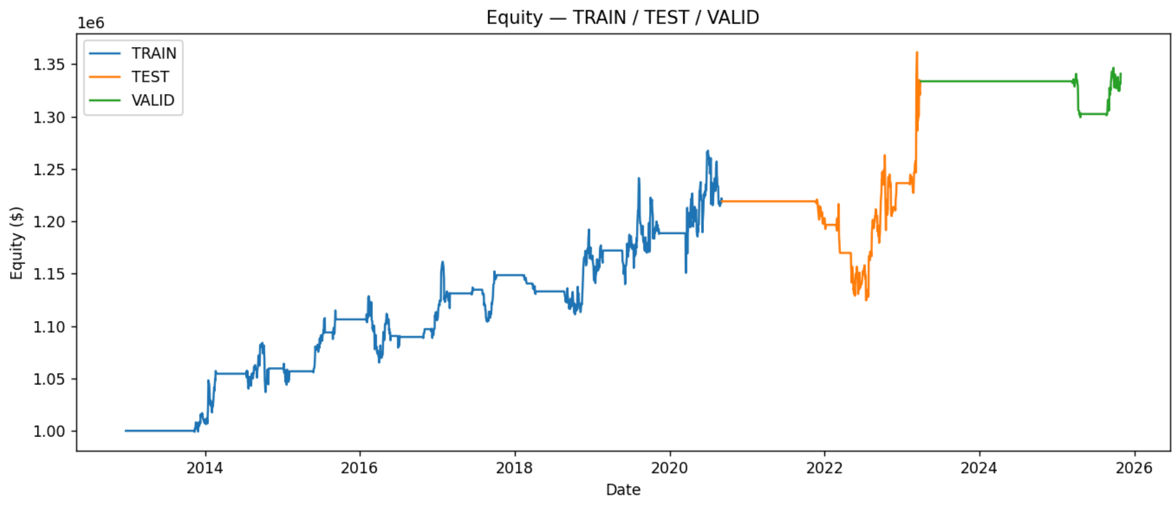
Equity curve overview

The equity curves from the TRAIN, TEST, and VALIDATION phases show how the model's capital evolved through time.
During the training period, the strategy showed a steady upward trend, with moderate fluctuations and controlled drawdowns. The equity increased from $1,000,000 to around $1.22 million, showing consistent growth as the Kalman Filter adapted to the price dynamics of both assets.

In the test phase, the performance improved even further, reaching about $1.33 million. This indicates that the optimized parameters generalized well to unseen data and that the trading logic did not overfit the training sample.

The validation phase was flatter, ending near $1.34 million with smaller fluctuations. This behavior is common when markets become less mean-reverting or more directional, which limits trading opportunities. Nonetheless, equity remained stable and positive, showing that the strategy preserved its capital even in less favorable conditions.

Overall, the combined equity plot (2013–2026) demonstrates that the model maintains profitability and robustness across all phases, with smooth transitions between in-sample and out-of-sample periods.



Performance metrics

The following metrics summarize the quantitative performance of the strategy across the three phases:

PERFORMANCE METRICS						
Period	Final Equity	Total Return	Sharpe	Sortino	Calmar	Max Drawdown
TRAIN	$1,219,091.57	**21.91%**	0.468	0.398	0.358	−7.29%
TEST	$1,333,716.51	**9.40%**	0.505	0.415	0.452	−7.89%
VALID	$1,340,939.64	**0.54%**	0.109	0.059	0.068	−3.09%

The Sharpe ratio measures the return per unit of total risk, while the Sortino ratio focuses on downside risk only — both indicate moderate but stable risk-adjusted performance.
The Calmar ratio compares annualized returns to the maximum drawdown, highlighting how efficiently the model grows capital relative to losses.

Across all periods, the model kept drawdowns under 8%, which is a good indicator of controlled volatility and risk discipline.

Trade statistics

The backtest generated a total of 3,231 trades across the full historical dataset.

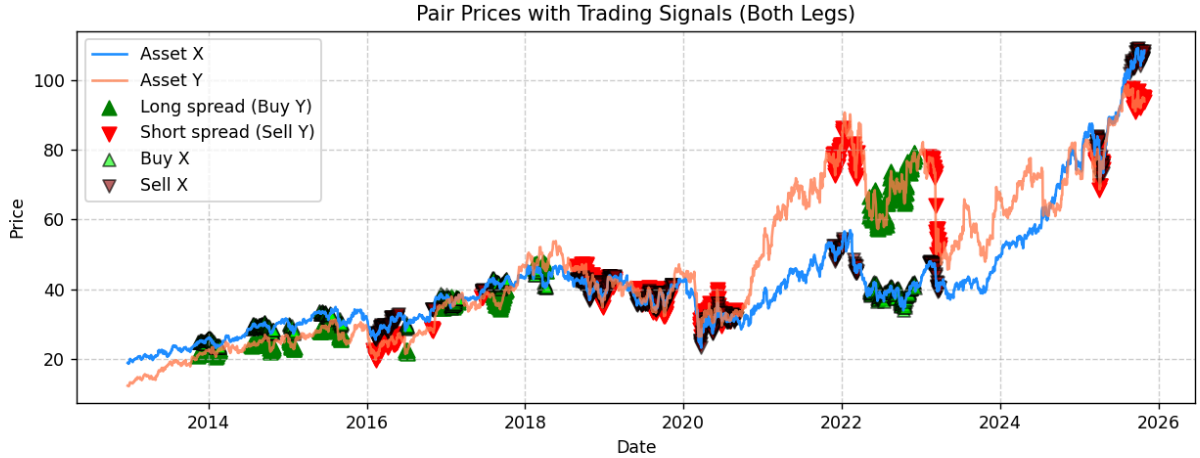
- The average return per trade was approximately 0.41%, with a standard deviation of 92.82%.
- The number of “flips” (position changes) decreased naturally over time as the Kalman Filter became more selective with its signals.
- Commissions across all trades totaled roughly $67,000, distributed as $47,325 in training, $14,726 in testing, and $5,266 in validation.

The trade distribution histograms show that most trades were concentrated near small positive or small negative returns, while a few outliers on both sides represent the biggest winners and losers. This pattern is typical for mean-reversion strategies, where frequent small profits compensate for occasional larger drawdowns.

The strategy’s profit factor (ratio of gross profits to gross losses) remained above 1 throughout all stages, confirming that the average winning trades outweighed the losing ones when adjusted for transaction costs.

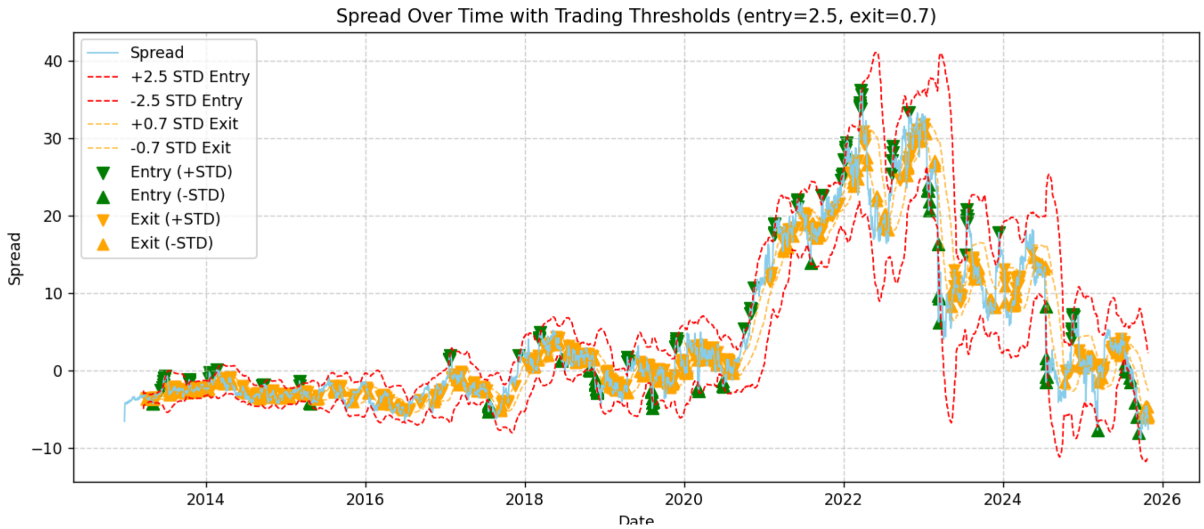
Visual analysis of trading activity

Pair prices with trading signals



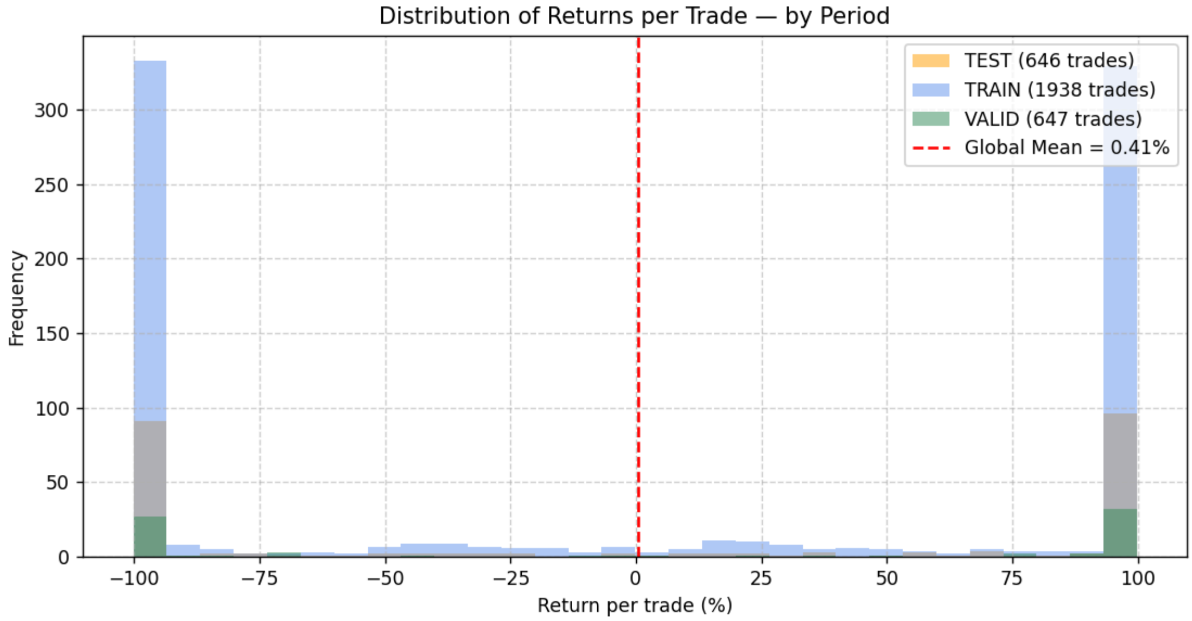
This plot displays both assets with markers for each buy/sell signal. The green triangles represent long positions (buying the undervalued asset and shorting the overvalued one), while the red triangles mark the opposite.
The alternation of these signals reflects how the model exploits short-term mispricings between the pair.

Spread evolution and thresholds



Here we see the spread between both assets along with the entry (±2.5 σ) and exit (±0.7 σ) thresholds.
Each trade is triggered when the spread crosses an entry line and is closed when it reverts to the exit band.
The spread consistently oscillates within these limits, confirming the mean-reverting behavior required for this strategy.

Distribution of returns per trade (by period)



The histogram shows the return per trade across TRAIN, TEST, and VALIDATION sets. Most trades cluster around small positive or negative returns, while a few extreme values on both sides represent large winners or losers.
The global average return per trade was +0.41 %, as indicated by the dashed red line. This pattern is typical for mean-reversion strategies, where frequent small profits compensate for occasional losses.

Together, these figures illustrate that the Kalman Filter successfully identified temporary deviations, reacted quickly to them, and maintained stability across multiple market cycles.

Out-of-sample performance

The out-of-sample results (TEST and VALIDATION) confirm that the model did not rely on overfitting. Even after optimization, the parameters maintained stability and performance consistency when applied to unseen data.
The equity growth during TEST, followed by a steady plateau in VALIDATION, suggests that the model captured true market relationships rather than random noise.
The hedge ratio (β□) also behaved smoothly across these phases, adjusting dynamically to the small structural changes between assets without diverging — a sign that the filter remained stable under real-market conditions.

Cost analysis

Transaction costs and borrow fees were included in all backtests to make the results realistic.

- A commission of 0.125% per leg was charged on every trade, representing typical broker fees.
- A borrow cost of 0.25% annualized was applied on short positions to simulate funding costs.

Despite these frictions, the strategy remained profitable in every phase.
Commissions and borrow charges totaled about $67,000 across the full sample — a small portion of total profits considering more than 3,200 trades were executed.

The relatively low impact of costs indicates that the trade frequency and position sizing were appropriate, preventing excessive turnover.
This also confirms that the model's profitability was not dependent on unrealistic assumptions such as zero transaction costs.

Conclusions

Working on this project allowed us to better understand how a pairs trading strategy can be designed, tested, and adjusted to perform under real market conditions. By combining cointegration analysis with the Kalman Filter, we managed to create a system that adapts dynamically to changes in price relationships between assets while maintaining a market-neutral exposure. Overall, the strategy showed consistent behavior during the training, testing, and validation periods, with controlled drawdowns and moderate but stable profitability, even after incorporating transaction and borrowing costs. These results indicate that the model is not only statistically sound but also operationally realistic.

One of the most relevant aspects we observed was how the Kalman Filter improved the hedge ratio estimation compared to static models. Instead of keeping a fixed β, the filter allowed it to evolve naturally with new data, reacting to changes in correlation and volatility. This adaptability reduced the risk of being exposed to directional market movements and helped maintain the logic of mean reversion throughout the different phases of the test. The model also proved robust in handling noise and unexpected fluctuations, showing that dynamic estimation methods can enhance traditional statistical arbitrage strategies.

After accounting for trading frictions, the system remained profitable, which reinforces the credibility of the approach. The inclusion of commissions and borrow costs made the evaluation much more realistic, allowing us to confirm that the performance was not artificially inflated. The number of trades executed, as well as the position sizing, were both reasonable and consistent with the expected mean-reversion dynamics of the selected pair. This suggests that the model could be applied to other correlated assets with similar behavior, potentially expanding its scope without compromising reliability.

For future improvements, there are a few practical steps that could help make the strategy more efficient and reliable. One simple adjustment would be to test different Z-score thresholds to see how changes in entry and exit levels affect performance under varying volatility. Another improvement could be expanding the set of tested pairs, either within the same sector or across similar industries, to find more consistent relationships and reduce dependency on a single pair. It would also be helpful to refine the parameter search for q, r, and the Z-score levels to ensure the model is not too sensitive to a specific configuration. Finally, adding more validation periods or stress-testing the model in different market conditions would help confirm its stability and adaptability over time.

In conclusion, this work showed that it is possible to design and implement a systematic pairs trading strategy that combines theory and practicality. The Kalman Filter gave the model the flexibility to learn continuously, while the cointegration framework ensured that trades were grounded in long-term statistical relationships. Despite the natural limitations of backtesting, the overall results were encouraging and showed that, with further refinement and broader testing, this approach could evolve into a solid foundation for quantitative trading strategies focused on consistency and controlled risk.

References

1. Chen, J. (2025, August 29). *Pairs Trading Strategy: Definition, Benefits, and Examples*. Investopedia. https://www.investopedia.com/terms/p/pairstrade.asp
2. Corporate Finance Institute. (n.d.). *Cointegration*. https://corporatefinanceinstitute.com/resources/data-science/cointegration/
3. Du Toit, R. (2023, February 5). *An Introduction to Cointegration for Pairs Trading*. Hudson & Thames. https://hudsonthames.org/an-introduction-to-cointegration/
4. Hussain, A. (2025, September 29). *Market Neutral Strategy: Definition, Benefits, and Risks Explained*. Investopedia. https://www.investopedia.com/terms/m/marketneutral.asp
5. Investing.com Academy. (n.d.). *Sharpe Ratio Definition*. https://www.investing.com/academy/analysis/sharpe-ratio-definition/
6. Investopedia. (2025). *Pairs Trade*. https://www.investopedia.com/terms/p/pairstrade.asp
7. Investopedia. (2025). *Mean Reversion*. https://www.investopedia.com/terms/m/meanreversion.asp
8. Investopedia. (2025). *Sortino Ratio*. https://www.investopedia.com/terms/s/sortinoratio.asp
9. Investopedia. (2025). *Calmar Ratio*. https://www.investopedia.com/terms/c/calmarratio.asp
10. Investopedia. (2025). *Maximum Drawdown (MDD)*. https://www.investopedia.com/terms/m/maximum-drawdown-mdd.asp

11. Investopedia. (2025). *Backtesting*. https://www.investopedia.com/terms/b/backtesting.asp
12. Nevil, S. (2025, 23 julio). Z-Score: Meaning and Formula. Investopedia. https://www.investopedia.com/terms/z/zscore.asp
13. QuantStart. (2025). Kalman Filter-Based Pairs Trading Strategy in QSTrader. https://www.quantstart.com/articles/kalman-filter-based-pairs-trading-strategy-in-qstrader/
14. Statistics How To. (2025). *Z-Score: Definition, Formula and Calculation*. https://www.statisticshowto.com/probability-and-statistics/z-score/#Whatisazscore
15. The Kalman Filter. (2025). *Kalman Filter Explained Simply*. https://thekalmanfilter.com/kalman-filter-explained-simply/
16. Wall Street Mojo. (2025). *Hedge Ratio*. https://www.wallstreetmojo.com/hedge-ratio/