UNIVERSITY OF BATH

Faculty of Humanities & Social Sciences

Dissertation

Submitted for MSc Economics Programme

Department of Economics

Pairs Trading: Evaluating the Performance of the Distance-based approach on the stocks comprising the S&P500 Index

Rohan Mukhtar

Department of Economics

September 2025

Abstract

This study evaluates the performance of a distance-based pairs trading framework through six distinct strategies that combine relative and absolute mean reversion rules with varying stop-loss thresholds. Three strategies employed absolute mean reversion with 1%, 5%, and 10% stop-loss levels, closing trades upon return to the mean. The other three relative strategies applied the same stop-loss thresholds but exited positions once the z-score reverted to ±0.5. Using daily price data through the end of 2024, we find that performance metrics varied across strategies, with win rates consistently fluctuating between 30 and 40 percent, while the number of trades was strategy dependent. Results demonstrate consistent Sharpe ratios and sectoral composition of trades, underscoring the robustness of the distance-based approach. Among the six, the Absolute 1% mean reversion strategy delivered the strongest results, achieving a cumulative PnL of 23.20, driven by the compounding of steady gains and effective risk management through its tight stop-loss. Beyond profitability, the model successfully captured temporary dislocations in sectoral trades, most frequently within Financials. The paper suggests that distance-based trading can deliver strong returns, which stands in contrast to some recent findings in the literature, and highlights the need for further research and deeper exploration of distance-based strategies in contemporary trading environments.

Contents

Abstract ... 2
Figures ... 4
Tables ... 5
Introduction ... 6
Literature Review ... 10
Methodology ... 13
Results and Analysis ... 17
i. Win Rate ... 17
ii. Average Trading Days ... 18
iii. Sharpe Ratio ... 18
iv. Cumulative Returns ... 19
v. Number of Trades ... 22
vi. Average PnL ... 22
vii. Trade Level Breakdown ... 24
i. Sector Pairs Categorization ... 24
ii. Top Pair by Win Rate ... 26
iii. Top 10 Pairs by Cumulative PnL ... 30
iv. Top 10 Trade Pair Sector Combinations ... 33
v. Individual Sector Trades ... 35
vi. Win Rate of Different Sectors ... 36
Conclusion ... 38
Appendix ... 39
Bibliography ... 50

Figures

Figure 1: Global Market Cap by Region: Authors own illustration using (Companiesmarketcap.com, 2025) 7
Figure 2: S&P500 Index over the Year (Maverick, 2025) 8
Figure 3: Sector-wise breakdown of the stocks comprising the S&P500 (Author’s own illustration using CRSP data) 13
Figure 4: Win Rate using different trading strategies (Author’s own calculations) 17
Figure 5: Sharpe Ration using different trading Strategies (Author’s own calculations) 19
Figure 6: Cumulative Returns over the years (Author’s own calculations) 21
Figure 7: Number of Trades using different trading strategies (Author’s own calculations) 22
Figure 8: PnL Frequency (1% absolute mean reversion): Author’s own calculations 23
Figure 9: PnL Frequency (1% relative mean reversion): Author’s own calculations 24
Figure 10: Sector Trade Categorization (1% Absolute Mean Reversion): Author’s own calculations 25
Figure 11: Sector Trade Categorization (5% absolute mean reversion): Author’s own calculations 25
Figure 12: Sector Trade Categorization (10% absolute mean reversion): Author’s own calculations 26
Figure 13: Top Pair by Win Rate (1% absolute mean reversion): Author’s own calculations 27
Figure 14: Yearly Breakdown of 1% Absolute Mean Reversion Top Pair (Author’s own calculations) 27
Figure 15: Top Pair by Win Rate (1% relative mean reversion): Author’s own calculations 28
Figure 16: Top Pair by Win rate (5% relative mean reversion): Author’s own calculations 29
Figure 17: Top Pair by Win rate (10% relative mean reversion): Author’s own calculations 29
Figure 18: Top 10 Aggregated Traded Sector Combinations: Author’s own calculations 33
Figure 19: Top 10 Sector Combinations for 1% Mean Reversion: Author’s own calculations 33
Figure 20: Top 10 Sector Combinations for 5% Mean Reversion: Author’s own calculations 34
Figure 21: Top 10 Sector Combinations for 1% relative mean reversion: Author’s own calculations 34
Figure 22: Individual Sector Trades (1% Absolute Mean Reversion): Author’s own calculations 35
Figure 23: Individual Sector Trades (5% Absolute mean reversion): Author’s own calculations 35
Figure 24: Individual Sector Trades (10% Absolute mean reversion): Author’s own calculations 36
Figure 25: Heatmap of Top Sector Pairs across the six trading strategies: Author’s own calculations 37

Figure A 1: PnL Frequency (5% absolute mean reversion): Author’s own calculations 39
Figure A 2: PnL Frequency (5% relative mean reversion): Author’s own calculations 39
Figure A 3: PnL Frequency (10% absolute mean reversion): Author’s own calculations 40
Figure A 4: PnL Frequency (10% relative mean reversion): Author’s own calculations 40
Figure A 5: Sector Trade Categorization (1% relative mean reversion): Author’s own calculations 41
Figure A 6: Sector Trade Categorization (5% relative mean reversion): Author’s own calculations 41
Figure A 7: Sector Trade Categorization (10% relative mean reversion): Author’s own calculations 42
Figure A 8: Top Pair by Win Rate (5% absolute mean reversion): Author’s own calculations 42
Figure A 9: Top Pair by Win Rate (10% absolute mean reversion): Author’s own calculations 43

Figure A 10: Top 10 Sector Combinations for 5% relative mean reversion: Author’s own calculations 44
Figure A 11: Sector Combination for 10% relative mean reversion: Author’s own calculations................ 44
Figure A 12: Sector Combinations for 10% Mean reversion strategy: Author’s own calculations 45
Figure A 13: Individual Sector Trades (1% relative mean reversion): Author’s own calculations 45
Figure A 14: Individual Sector Trades (5% relative mean reversion): Author’s own calculations 46
Figure A 15: Individual Sector Trade for 10% relative mean reversion strategy: Author’s own calculations 46

Tables

Table 1: Trading Windows Breakdown 14
Table 2: Trading Criteria........ 15
Table 3: Average Trading Days using different trading strategies (Author’s own calculations)................ 18
Table 4: Cumulative Returns using different trading Strategies (Author’s own calculations).................... 20
Table 5: Top Pair by Cumulative PnL (1% relative mean reversion): Author’s own calculations 30
Table 6: Top Pair by Cumulative PnL (1% absolute mean reversion): Author’s own calcualtions 31
Table 7: Top Pair by Cumulative PnL (5% absolute mean reversion): Author’s own calculations 31
Table 8: Top Pair by Cumulative PnL (10% absolute mean reversion): Author’s own calculations 32

Table A 1: Top Pair by Cumulative PnL (5% relative mean reversion) 43
Table A 2: Top Pair by Cumulative PnL (10% relative mean reversion): Author’s own calculations 43
Table A 3: Top 10 Sector Combinations by Win Rate (1% relative mean reversion strategy): Author’s own calculations 47
Table A 4: Top 10 Sector Combinations by Win Rate (5% relative mean reversion strategy): Author’s own calculations 47
Table A 5: Top 10 sector combinations by win rate (10% relative mean reversion): Author’s own calculations 48
Table A 6: Top 10 sector combinations by win rate (1% absolute mean reversion): Author’s own calculations 48
Table A 7: Top 10 sector combinations by win rate (5% absolute mean reversion): Author’s own calculations 49
Table A 8: Top 10 sector combinations by win rate (10% absolute mean reversion): Author’s own calculations 49

Introduction

Trading as a concept involves the exchange of goods and services between parties. Earliest form of trading was done through the Barter System, where goods were exchanged directly without the use of money (Kindleberger, "The Great Transformation" by Karl Polanyi, 1974). However, the barter system was complicated and limited the exchange of goods because of the lack of the common value of measure (Jevons, 1876). The earliest forms of organized marketplaces first emerged in the Ancient Middle East, around 3000 BC. This helped traders to trade goods and services at a common place. Trade then expanded through land and sea routes like the Silk Road (Braudel, 1982).

Modern trade has entirely transformed, largely due to the pivotal role played by financial institutions. Early banking systems, such as the Medici Bank in Renaissance Florence, introduced mechanisms for credit, bills of exchange, and currency transfer that allowed merchants to expand their operations beyond local markets (Roover, 1948). These innovations supported the rise of international commerce and reduced the risks associated with long-distance trades that existed before. The 17^{th} century was a pivotal time, as it witnessed the establishment of formal stock exchanges, beginning with the Amsterdam Stock Exchange in 1602, which facilitated the trading of shares in the Dutch East India Company (Neal, 1990). Similar institutes appeared such as the London Stock Exchange formalizing share trading by the late 18^{th} century (Michie, 1999). Alongside these developments, Central Banks emerged as a major institution that helped as a stabilizing force in the world of global finance, with the Bank of England (founded in 1694) providing a mechanism for monetary policy and regulation (Kindleberger, A Financial History of Western Europe, 1993). Collectively, these institutions formed the basis for the transformation from the traditional system to the modern interconnected financial system that we see today.

The evolution of trade from goods and commodities into financial instruments marked a decisive turning point in economic history. Stock trading, unlike the exchange of physical goods, involves the buying and selling of company shares, bonds, and other securities, enabling businesses to raise capital and investors to share in profits and risks associated with such investments (Goetzmann & Rouwenhorst, 2005). Chicago Board of Trade founded in 1848, institutionalized futures contracts and standardized markers, expanding the scope of what could be traded beyond tangible goods (Cronon, 1991). By the late 19th and early 20th centuries, stock exchanges in New York, London, and other financial centers had established trading floors where brokers matched buyers and sellers, embedding securities trading at the heart of global capitalism (Sylla, 1998). This framework paved the way for subsequent digital innovations, as technological advances evolved the core practice of matching buyers and sellers of assets into electronic and eventually online systems. The launch of NASDAQ in 1971 as the first electronic exchange and subsequent online brokerage services in the 1990s carried forward the same logic of trading, but at speeds and scales never previously imagined (Macey & O'Hara, 1999).

The rapid expansion of the internet in the late 20th century accelerated the transformation of financial markets, allowing trading to move fully into digital spaces. Electronic platforms enabled orders to be processed at unprecedented speed, linking global markets in real time and vastly increasing trading volumes (Domowitz, 2002). Improvements in computing power supported the development of algorithmic and high-frequency trading in the 2000s, where trades could be executed in fractions of a second, enhancing market efficiency but also raising concerns about volatility and systemic stability, which in turn created opportunities for traders to profit from rapid market movements (Kirilenko, Kyle, Samadi, & Tuzun, 2017). Parallel to these developments, the emergence of block chain technology introduced decentralized forms of trading through cryptocurrencies, creating entirely new asset classes and bypassing traditional financial intermediaries (Tapscott & Tapscott, 2016). These innovations have reshaped trading into a global, continuous, and technologically intensive activity, maintaining the core principle of exchange but within an environment far more integrated and complex than any previous era (Chuen, 2015).

Today, global trading operates on an unprecedented scale, with the total market capitalization of publicly listed companies worldwide reaching approximately US$124 trillion as of 2025. The United States dominates this landscape, accounting for nearly 60% of global equity markets, driven largely by its two powerhouse exchanges, the New York Stock Exchange (US$25 trillion) and NASDAQ (US$20.6 trillion). (Companiesmarketcap.com, 2025)

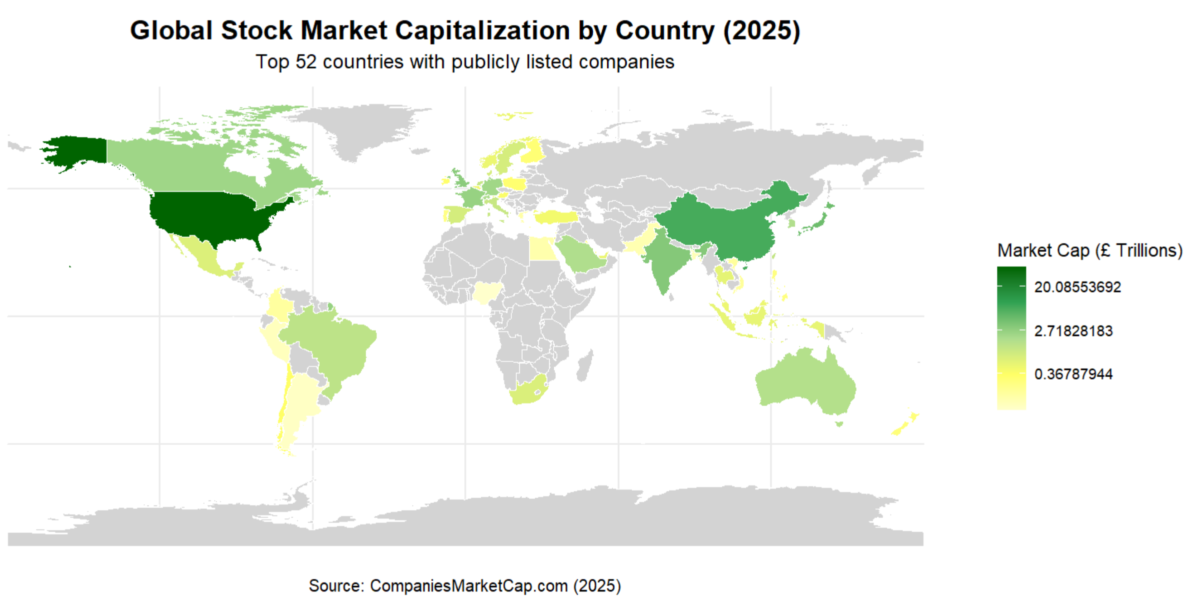

Figure 1: Global Market Cap by Region: Authors own illustration using (Companiesmarketcap.com, 2025)

As of 2025, the top ten countries by stock market capitalization are overwhelmingly led by the United States, which boasts a market cap of £48.33 trillion, followed by China (£7.59 trillion), Japan (£4.55 trillion), India (£3.30 trillion), and the United Kingdom (£3.05 trillion), with the remaining five; Canada, France, Germany, Switzerland, and Saudi Arabia; ranging between £1.74 trillion and £2.53 trillion. A choropleth map (Figure 1) further illustrates the geographic concentration of global market capitalization.

On an average trading day, over $300–$360 billion worth of shares change hands across U.S. markets, reflecting both the liquidity and depth of American financial institutions. In 2024, the U.S. equity markets saw robust activity, with an average daily trading value of approximately $516.5 billion; comprised of $300.3 billion through traditional exchanges, $68.8 billion via ATS platforms, and $147.4 billion over-the-counter (FINRA, 2024). At the same time, market dynamics experienced meaningful growth: average daily volume rose about 10.2% year-over-year to 12.2 billion shares, while the notional value of these trades surged 18.1% to reach $607.7 billion/day, underscoring accelerating investor engagement (Cboe, 2025). Within this ecosystem, the S&P 500 has become the most widely followed benchmark for U.S. equities. Since its inception in 1957, the index has delivered an average annual return of over 10% nominally (6.7% after inflation), meaning that a $100 investment in 1957 would be worth roughly $82,000 today in nominal terms and around $7,100 in real terms. (Maverick, 2025)

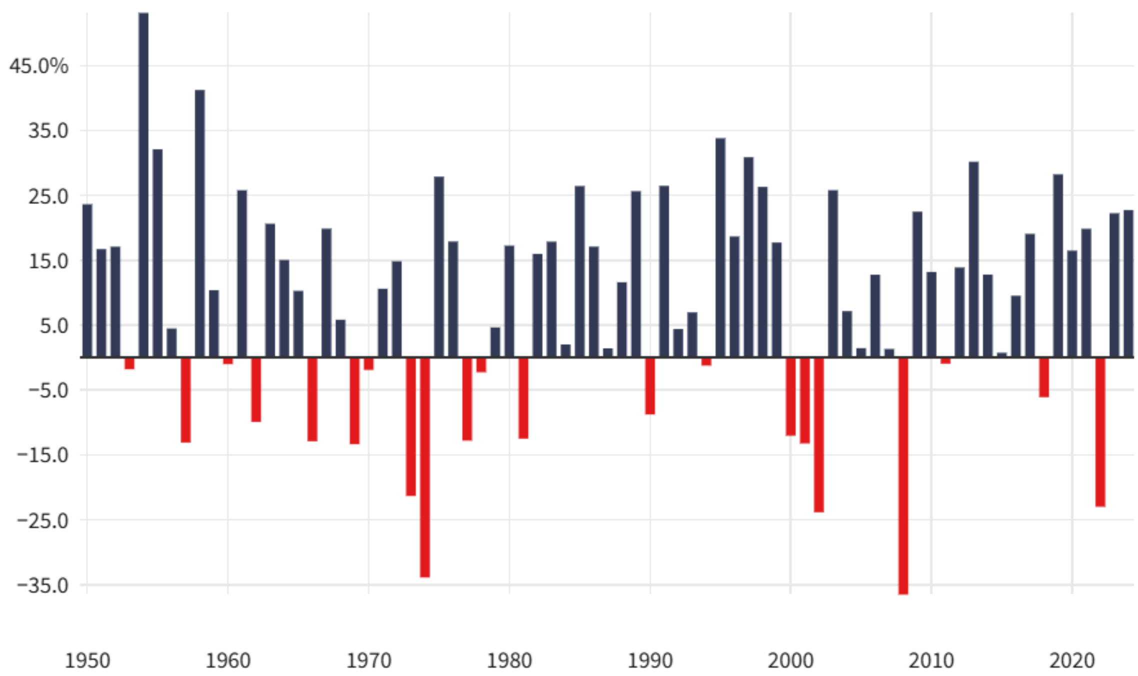

Figure 2: S&P500 Index over the Year (Maverick, 2025)

Figure 2 illustrates the annual returns of the S&P 500 over the past 70 years, highlighting its overall consistent growth despite occasional significant downturns. The large negative returns in specific years correspond to major economic or geopolitical shocks that disrupted financial markets globally. For example, the sharp decline in the 1970s reflects the oil embargo and stagflation, the early 2000s dip aligns with the dot-com bubble burst, the steep fall in 2008 coincides with the global financial crisis, and the 2020 drop is associated with the Covid-19 pandemic. These events caused heightened market volatility and substantial short-term losses, but the broader trend demonstrates the resilience and long-term upward trajectory of the S&P 500 index.

Given this transformation of trading into a highly liquid, technology-driven global system, the S&P 500 naturally emerges as the focal point for this study. The index comprises of about 500 of the largest publicly listed U.S. companies, spanning all major sectors such as technology, healthcare, energy, consumer goods, and financials, and collectively accounts for nearly 80% of U.S. equity market capitalization. This breadth ensures that the index not only reflects the performance of individual firms but also serves as a proxy for the wider U.S. economy and, by extension, global financial markets. Its depth and liquidity make it an ideal environment to study modern trading activity, where price discovery is rapid and transaction costs are minimal. Importantly, the S&P 500 has been central to empirical studies on algorithmic and high-frequency trading, with research showing that automated strategies in its constituent stocks have significantly influenced liquidity, volatility, and efficiency (Hendershott, Jones, & Menkveld, 2011). More recently, its securities have been widely used in machine learning–based trading models, reflecting their role as the primary testbed for innovation in financial research (Gu, Kelly, & Xiu, 2020). For these reasons, the S&P 500 provides both the practical relevance and the methodological robustness to anchor an exploration of trading in the contemporary financial landscape.

While the S&P 500 has long served as the basis for testing trading strategies, the performance of distance-based approaches has been mixed. Earlier studies in the 2010s reported encouraging results, but later evidence shows a marked decline in profitability. (Gatev, Goetzmann, & Rouwenhorst, 2006) note that returns weakened significantly toward the end of their sample, while (Do & Faff, 2010) find that the strategy's effectiveness deteriorated further in subsequent decades. This raises the question of whether distance-based methods can still deliver value in today's highly competitive and algorithm-driven markets. This paper addresses that gap by re-examining the distance-based trading approach on S&P 500 stocks over the recent period, assessing whether the strategy continues to provide profitable opportunities or whether its success has been eroded over time.

Literature Review

Pairs trading is a market-neutral strategy in which investors simultaneously take long and short positions in two historically correlated assets, aiming to profit from temporary deviations in their relative prices. The core idea is that the spread between these assets will eventually revert to its historical mean, allowing traders to capitalize on mispricing without taking on directional market risk (Gatev, Goetzmann, & Rouwenhorst, 2006); (Vidyamurthy, 2004)). This strategy belongs to the broader class of statistical arbitrage methods and has been widely applied across equities, ETFs, commodities, and other asset classes. Its market-neutral nature makes it particularly appealing in volatile or sideways-trending markets, as profits depend primarily on the convergence of paired assets rather than overall market movements.

Pairs trading is implemented through several methodological frameworks, the most widely studied being distance-based, cointegration-based, and correlation/coupled approaches. The distance-based approach involves selecting historically correlated pairs and monitoring the price spread or ratio between them. Trades are entered when the spread exceeds a pre-defined threshold, such as one or two standard deviations from the mean, and exited when convergence occurs (Gatev, Goetzmann, & Rouwenhorst, 2006). This method is straightforward and computationally simple, making it a common choice for practitioners. However, its performance is highly sensitive to parameter choices, formation periods, and transaction costs. Moreover, as the strategy became widely adopted, profits have diminished due to market crowding and reduced arbitrage opportunities (Do & Faff, 2010); (Rad, Low, & Faff, 2016).

In contrast, the cointegration-based approach relies on identifying pairs whose prices share a long-run equilibrium relationship, resulting in a stationary spread. By employing econometric tests such as Engle-Granger or Johansen cointegration tests, traders ensure that deviations from equilibrium are temporary and predictable, making the strategy more robust to trending prices (Sen, 2022); (Rad, Low, & Faff, 2016) . Cointegration methods are generally more statistically rigorous than distance-based approaches and are less prone to spurious signals, although they require larger datasets and more computational effort.

Other approaches include correlation-based selection and copula methods, which model linear or nonlinear dependencies between assets, respectively. These methods can provide more stable but often smaller profits, particularly in markets where traditional pairs are crowded or historical relationships are unstable (Rad, Low, & Faff, 2016); (Tadi & Witzany, 2023).

While distance-based pairs trading offers simplicity and ease of implementation, cointegration-based methods are generally considered more robust, particularly in trending or volatile markets. Empirical studies suggest that cointegration captures the long-run equilibrium relationships between asset prices, reducing the risk of spurious trades that can arise from purely statistical correlations (Sen, 2022); (Rad, Low, & Faff, 2016). (Chen, Chen, Chen, & Li, 2017) finds that abnormal returns in equity pairs often stem from short-term reversals and momentum effects, highlighting the limitations of relying solely on distance measures. Methodological advancements have further strengthened cointegration approaches: (Holý & Černý, 2021) derive optimal entry and exit rules under Ornstein–Uhlenbeck processes, balancing profitability and risk, while (Sarmento & Horta, 2020) employ machine learning techniques, such as clustering and LSTM forecasting, to improve pair selection and reduce drawdowns. (Qureshi & Zaman, 2024) propose a graph-matching approach that minimizes overlap among selected pairs, enhancing portfolio diversification and risk-adjusted returns. Comparative studies indicate that while distance-based strategies can still generate profits in certain periods, cointegration-based methods consistently demonstrate higher Sharpe ratios and stability, especially in highly liquid markets like the S&P 500 ((Gatev, Goetzmann, & Rouwenhorst, 2006); (Rad, Low, & Faff, 2016). Hybrid strategies that combine initial distance screening with cointegration verification have also been suggested to leverage the simplicity of distance methods while maintaining the statistical rigor of cointegration (Vidyamurthy, 2004).

Measuring the profitability of pairs trading strategies typically involves calculating returns based on daily price data, under the assumption that prices follow a continuous trading process and that spreads between paired assets are observable and tradable at market prices. Returns are often computed as the cumulative profit from long and short positions divided by the initial investment or expressed as annualized excess returns relative to a benchmark index. The standard approach assumes that trades are executed at closing prices, though in practice bid–ask spreads, slippage, and execution delays can significantly affect realized performance. Transaction costs are especially important, as pairs trading involves frequent rebalancing and potentially short holding periods; neglecting them can lead to substantial overestimation of profitability. For instance, (Bowen, Hutchinson, & O'Sullivan, 2010) demonstrate that while raw pairs trading returns may appear attractive, incorporating realistic transaction costs such as bid–ask spreads and execution delays can considerably reduce net returns, sometimes eliminating arbitrage opportunities altogether. The study examines the characteristics of high-frequency pairs trading using a sample of FTSE100 constituent stocks for the period January to December 2007. The authors demonstrate that the excess returns of the strategy are highly sensitive to transaction costs and execution speed. Specifically, they find that when transaction costs are set at 15 basis points, the excess returns are reduced by more than 50%. Additionally, implementing a one-period delay in execution eliminates returns altogether. The study also observes that the majority of returns occur in the first hour of trading and that excess returns show weak exposure to traditional risk factors. Similarly, (Rad, Low, & Faff, 2016) find that after accounting for transaction costs, the profitability of pairs trading strategies diminishes significantly, underscoring the importance of evaluating both gross and net performance. Recent studies emphasize the necessity of incorporating transaction costs into performance assessments, with Sharpe ratios and drawdown measures providing a fuller picture of risk-adjusted outcomes.

Although cointegration-based strategies are often considered more robust, several studies indicate that distance-based pairs trading can remain effective, particularly when carefully calibrated to market conditions. (Ma & Ślepaczuk, 2022) find that in the Hong Kong stock market, distance-based methods yielded competitive risk-adjusted returns, sometimes approaching those of cointegration strategies. Similarly, (Hsia, 2021) demonstrates that variations of the distance approach, when adapted to specific price dynamics and trading frequency, can generate profitable outcomes while maintaining simplicity and low computational overhead. These findings suggest that the distance method may be underappreciated, particularly in contexts where computational efficiency and rapid implementation are important, and when transaction costs and execution constraints are carefully managed (Bowen, Hutchinson, & O'Sullivan, 2010).

Prior studies on distance-based pairs trading generally report modest profitability, with win rates typically falling in the 40–50% range. (Do & Faff, 2010) showed that while the strategy can yield abnormal returns, the proportion of profitable trades often hovers closer to the mid-40s once transaction costs are considered. (Huck & Afawubo, 2015) similarly found that success rates decline toward 40% in more efficient markets, highlighting the sensitivity of outcomes to parameter choices and market conditions. (Gatev, Goetzmann, & Rouwenhorst, 2006) documented higher win rates in early U.S. data, but subsequent evidence suggests diminished effectiveness in more recent periods.

These findings indicate that distance-based pairs trading, when carefully implemented and calibrated, can yield robust and competitive returns while maintaining simplicity and computational efficiency. Empirical evidence shows that, under realistic assumptions about transaction costs and trading frequency, the distance method remains effective across different markets and asset classes (Ma & Ślepaczuk, 2022); (Hsia, 2021); (Bowen, Hutchinson, & O'Sullivan, 2010). This positions the distance-based strategy as a practical and well-supported choice for the empirical analysis conducted in this study.

Methodology

This section details the research methodology, which applies a distance-based approach to daily stock prices of S&P 500 constituents over the period from 1 January 2019 to 31 December 2024.The data is obtained from the Centre for Research in Security Prices (CRSP). (CRSP, 2024) .

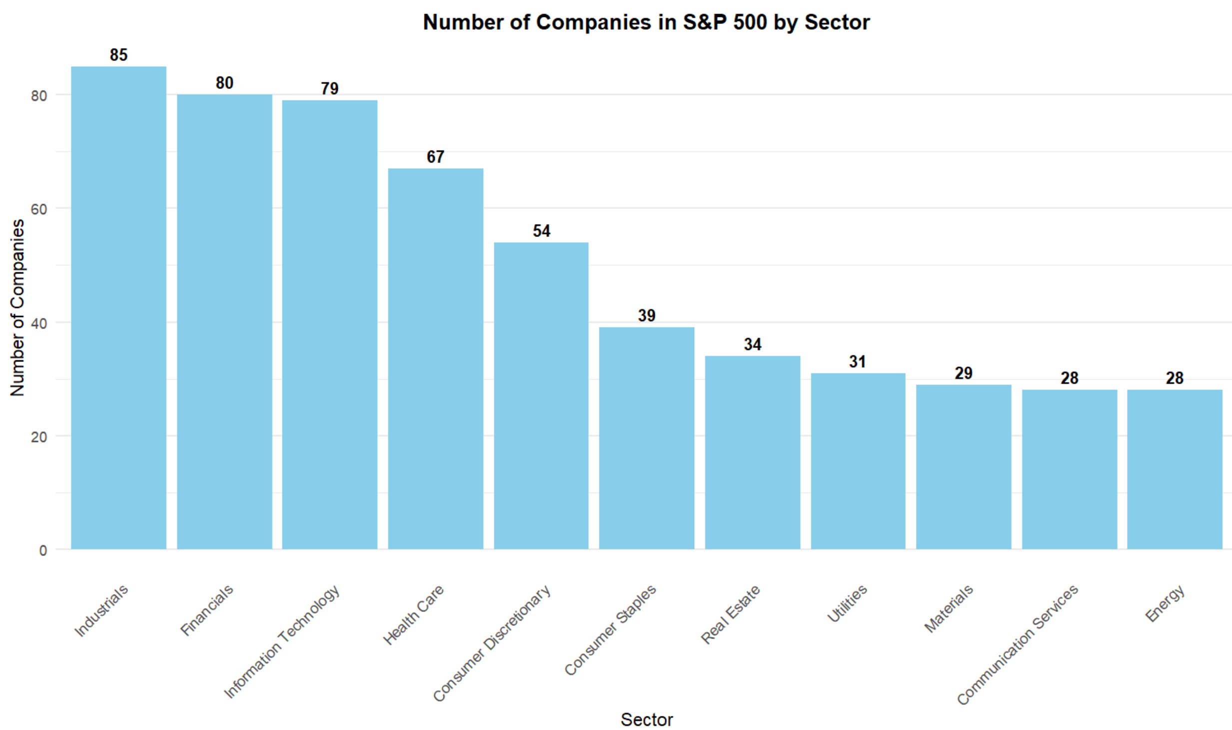

Figure 3: Sector-wise breakdown of the stocks comprising the S&P500 (Author's own illustration using CRSP data)

The data that we obtained for the 554 stocks were mainly composed of companies from the Industrials, Financials and Information Technology Sector. These three sectors made up about 45% of the companies in the S&P500 Index.

The trading strategies are formed on the basis of existing literature and by trying out some new strategy parameters using a trial-and-error approach to find the optimal cumulative returns. We use 2-year formation windows, where we identify the pair of stocks that have the least distance between them. The windows are based on the methodology used by (Gatev, Goetzmann, & Rouwenhorst, 2006). In total, we identify the top 50 pairs for every window. Then, we have a one-year trading window where we short or long these stocks based on a strategy.

Formation Period (2 Years)	Trading Period (1 Year)	Pairs Selected	Strategy Applied
2019 – 2020	2021	Top 50 pairs	Long/Short based on distance strategy
2020 – 2021	2022	Top 50 pairs	Long/Short based on distance strategy
2021 – 2022	2023	Top 50 pairs	Long/Short based on distance strategy
2022 – 2023	2024	Top 50 pairs	Long/Short based on distance strategy

Table 1: Trading Windows Breakdown

Table 1 shows the formation periods, trading period for all the four trading windows.

The strategy used in this paper is based on the direct difference between the scaled prices of the two stocks in a pair:

$$S_{\{i,j,t\}} = \frac{P_{i,t}}{\bar{P}_i} - \frac{P_{j,t}}{\bar{P}_j}$$

Where $P_{i,t}$ and $P_{j,t}$ are daily prices of stocks i and j, and $\bar{P}_i$ and $\bar{P}_j$ are their historical means over the formation window. The z-score is then calculated using a 20-day lookback period:

$$Z_{i,j,t} = \frac{S_{i,j,t} - \mu s_{i,j,t}}{\sigma s_{i,j,t}}$$

For this research, we employed a total of six distinct pairs trading strategies, all built upon a common framework but differentiated by their exit criteria and stop-loss levels. All strategies share the same entry points: we initiate a long position on the spread if its z-score falls below -1.5, betting on the spread to widen, and we go short if the z-score rises above 1.5, anticipating the spread to narrow.

The strategies are then categorized into two groups based on their exit conditions. The relative mean reversion strategies exit a long trade when the z-score rises above -0.5 and a short trade when it falls below 0.5. The absolute mean reversion strategies, on the other hand, require the z-score to fully revert to its mean of 0 before exiting either a long or short position. Both types of exit criteria were tested with three different stop-loss levels: 1%, 5%, and 10%. Furthermore, to manage risk and time in the market, all six strategies are subject to a maximum trading duration of 15 days. While the stop-loss is based on daily closing prices rather than real-time data, it effectively serves to trigger an exit at the end of any day where a loss greater than the specified percentage has occurred.

Strategy	Exit Rule	Stop-Loss Level	Max Holding Period
1	Relative mean reversion (exit at Z = −0.5 for long, Z = 0.5 for short)	1%	15 days
2	Relative mean reversion	5%	15 days
3	Relative mean reversion	10%	15 days
4	Absolute mean reversion (exit only when Z = 0)	1%	15 days
5	Absolute mean reversion	5%	15 days
6	Absolute mean reversion	10%	15 days

Table 2: Trading Criteria

The foundation of the distance-based pairs trading strategy rests on the principle of mean reversion. The core idea is that the relative price of one stock in a pair, when it deviates significantly from its historical mean relationship with its partner, is likely to revert back to that mean over time. The successful implementation of this approach is contingent upon two key assumptions. First, the normalized price spread of the pairs must exhibit stationarity over the trading period, which is the underlying condition that allows for predictable mean-reverting behavior. Second, the chosen distance metric must accurately measure the co-movement between stock prices. By pre-selecting the top 50 pairs with the lowest distance scores, the paper focuses trades on the pairs that have historically demonstrated the strongest mean-reverting tendencies.

The paper's backtesting framework operates under certain simplifying assumptions common in academic studies. Daily closing prices are utilized, which abstracts away from intraday volatility and assumes that positions can be opened and closed without market friction, such as trading costs or slippage. Returns for each pair are calculated as:

$$R_{i,j,t} = \begin{cases} \dfrac{P_{i,t+1} - P_{i,t}}{P_{i,t}} - \dfrac{P_{j,t+1} - P_{j,t}}{P_{j,t}}, \text{if long i short j} \\ \dfrac{P_{j,t} - P_{j,t+1}}{P_{j,t}} - \dfrac{P_{i,t} - P_{i,t+1}}{P_{i,t}}, \text{if short i long j} \end{cases}$$

Cumulative returns and risk-adjusted performance metrics, such as the Sharpe ratio, are computed as:

$$R_{cum} = \sum R_{i,j,t}, \qquad \text{Sharpe Ratio} = \frac{\bar{R}_{daily} \cdot \sqrt{252}}{\sigma(R_{daily})}$$

Maximum drawdown for each strategy is calculated as:

$$\text{MDD} = max_t \frac{Peak_t - Trough_t}{Peak_t}$$

In addition to raw cumulative returns, standardized returns are frequently employed in the literature to enable comparisons across strategies and datasets of different scales. Standardization typically involves expressing returns relative to their standard deviation, effectively transforming them into a z-score series that highlights excess performance adjusted for volatility (Gatev, Goetzmann, & Rouwenhorst, 2006). This approach ensures that performance evaluations are not biased toward strategies that yield higher absolute returns but at the cost of disproportionately higher risk.

This paper adopts a framework consistent with established literature. A back testing procedure was implemented in R to identify the top 50 stock pairs within each two-year formation window, based on the minimum sum of squared price differences (the distance metric).

$$D_{(i,j)} = \sum_{t=1}^{T} \left(\frac{P_{i,t}}{\overline{P_i}} - \frac{P_{j,t}}{\overline{P_j}}\right)^2$$

Where $P_{i,t}$ and $P_{j,t}$ denote daily closing prices of stocks i and j, and $\overline{P_i}$and $\overline{P_j}$ are their respective historical means over the formation window. The 50 stock pairs with the lowest distance values are retained for trading in the subsequent one-year window. These selected pairs were subsequently traded across their respective one-year trading windows in accordance with the six strategies outlined earlier. All trades were executed at daily closing prices, and performance was evaluated using cumulative returns, standardized returns, Sharpe ratios, win rates, and maximum drawdowns. This approach aligns with common academic practices while also introducing parameter variations to assess the robustness of the distance-based methodology.

Moving beyond a single strategy, the central objective of this paper is to conduct a comparative analysis of six distinct pairs trading strategies. This analysis systematically evaluates and contrasts performance based on combinations of relative and absolute exit thresholds, and stop-loss levels of 1%, 5%, and 10%. Strategies are compared across total returns, Sharpe ratio, win rates, and top-performing pairs.

Results and Analysis

The success of the strategies are dependent upon the different parameters set in the trading strategies and are evident in the overall returns, Sharpe ratio, total number of trades, average days of trades, win rate etc.

i. Win Rate

We will start our discussion with the overall win rate of the strategies.

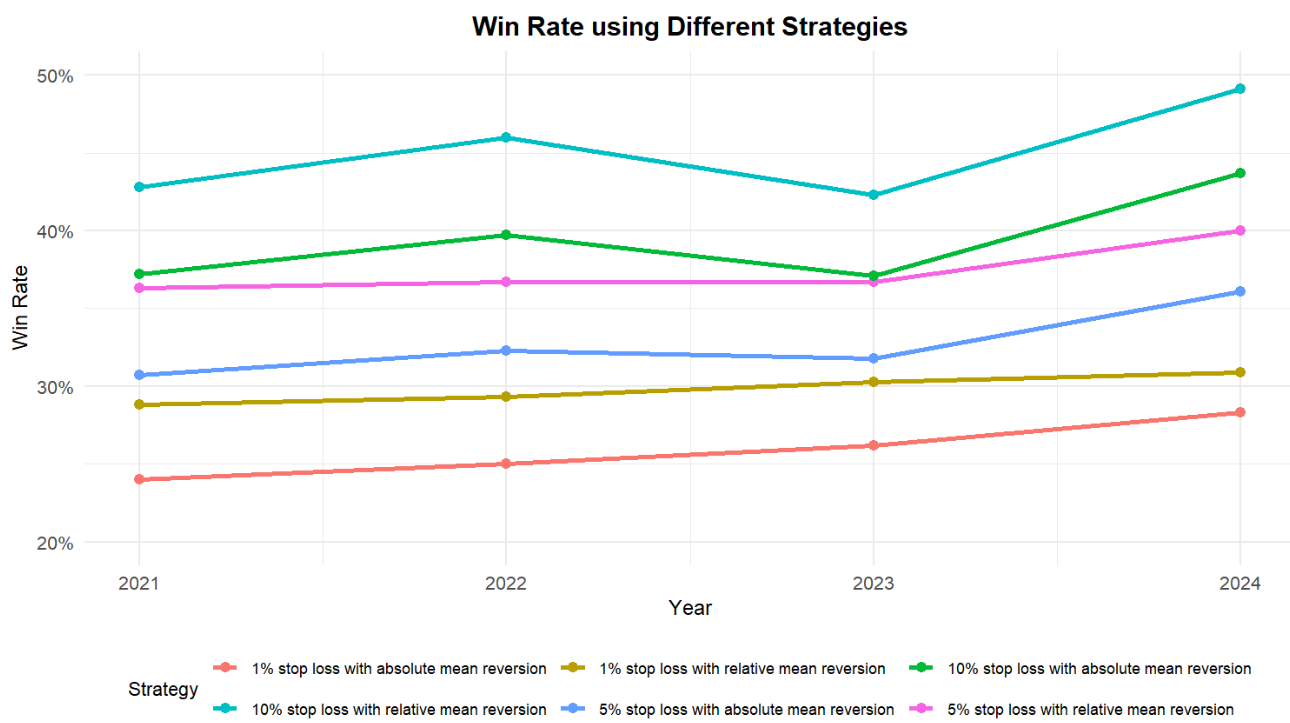

Figure 4: Win Rate using different trading strategies (Author's own calculations)

Figure 4 suggests that the win rate of almost all the trading strategies is consistent with our literature. Existing literature suggests that the win rate of trading using the distance-based approach is at about 40% or less. We can see the same trend in our trading strategies. We also note that the win rate of the strategies decreases as the stop loss decreases. For example, the lowest win rate is for the 1% stop loss with absolute mean reversion. We also note that on average, the relative mean reversion strategy has a better win rate than the absolute ones. The highest win rate is for the 10% stop loss strategy with relative mean reversion.

ii. Average Trading Days

Now, we will see how the average trading days for a trade opened differed for the different trading strategies.

Year	1% Absolute mean reversion	5% Absolute mean reversion	10% Absolute mean reversion	1% Relative mean reversion	5% Relative mean reversion	10% Relative mean reversion
2021	4.403509	5.536466	6.783956	3.564181	4.409305	5.358447
2022	4.31521	5.460203	7.032013	3.531901	4.465395	5.640675
2023	4.477332	5.430506	6.343774	3.594047	4.365211	5.08656
2024	4.864017	6.250791	7.620874	3.94274	5.087423	6.187069

Table 3: Average Trading Days using different trading strategies (Author's own calculations)

From table 3, we can see a couple of trends. First, we see that the average trading days decrease as the stop loss decreases. We also see that the relative mean reversion trading strategies have on average less trading days for a strategy in comparison to the absolute mean reversion trading strategies. This could be because it might take more time for a pair of stocks to converge or diverge based upon the type of trade opened. This is also because with a lesser stop loss, there is a greater chance of the stop loss being hit and this decreases the win rate of the strategies. This, in turn, also decreases the average number of trading days. The highest average trading days for any strategy is for the 10% absolute mean reversion strategy in 2024 and the lowest average trading days are in 2022 for the 1% relative mean reversion strategy.

iii. Sharpe Ratio

Now, we shall discuss the Sharpe Ratio of the different trading strategies that we have used to back test our data. Sharpe ratio is basically the evaluation of our strategies in a way which compares it to a risk-free asset, after adjusting for its returns. It effectively calculates the risk-adjusted returns and is an important part of literature when measuring the effectiveness of trading strategies.

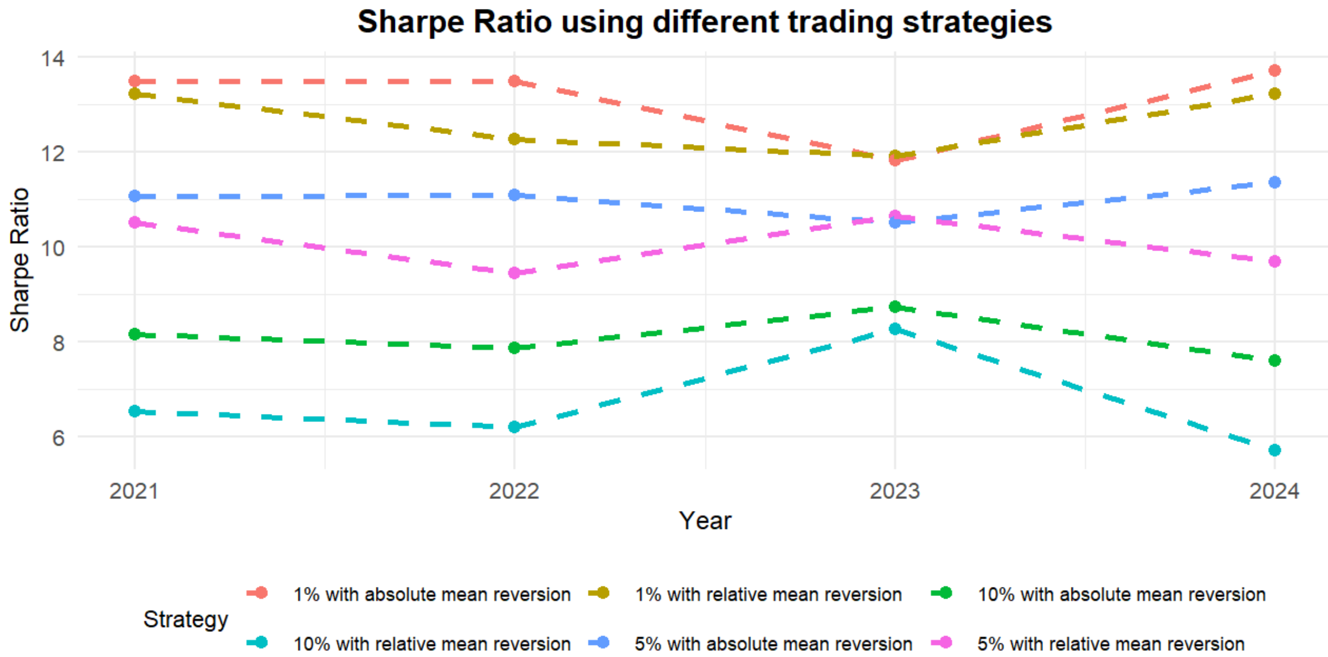

Figure 5: Sharpe Ration using different trading Strategies (Author's own calculations)

We can see from Figure 5 that the Sharpe ratio for all the trading strategies has remained above the threshold of 6 for almost all of the years, which is considered a benchmark for a good trading strategy (Insert Caption). We also see a lot more similarities in trend between the strategies. We all see that the Sharpe ratio for all the strategies in 2023 converges between the range of 8 to 12. We can also see that the best strategy in term of this parameter is the one with 1% stop loss and absolute mean reversion. The highest Sharpe Ratio for any strategy for any year was in 2023 for the 1% absolute mean reversion strategy where the Sharpe ratio was about 13.7. This was also consistent with the highest return for any year, which we will discuss later. The strategy which yielded the lowest Sharpe ratio was the 10% stop loss with relative mean reversion, which had a Sharpe Ratio of about 5.72 in 2024. This is also consistent with the return of the strategy that year, which yielded a return of 0.73 units only.

iv. Cumulative Returns

In pairs trading, the success of a strategy is most effectively measured by its cumulative returns. To ensure a fair and direct comparison between different strategies, a standardized approach is used. This method abstracts away from the actual dollar amount of each trade and instead treats every single trade as having a uniform size, which we can call 1 standardized unit. This approach is crucial because it isolates the performance of the strategy's logic from external factors like trade size. The cumulative return is then calculated by summing the profit or loss from each individual trade, measured in these standardized units, over the entire back testing period. If a trade results in a 0.2-unit gain and the next a 0.1-unit loss, the cumulative return after two trades would be 0.1 units. This provides a clear, scalable metric of how the initial 1 unit of capital would have grown over the test period.

Year	1% Absolute mean reversion	5% Absolute mean reversion	10% Absolute mean reversion	1% Relative mean reversion	5% Relative mean reversion	10% Relative mean reversion
2021	2.196935	2.000095	1.572135	1.776118211	1.54	1.0314
2022	4.293924	3.409577	1.999358	2.684274963	1.90	1.776118
2023	12.86537	9.414718	4.585951	6.027148133	4.00	2.901952
2024	23.19996	14.62742	5.354226	8.347195292	4.35	2.119162

Table 4: Cumulative Returns using different trading Strategies (Author's own calculations)

Table 4 shows that the 1% stop loss with absolute relative mean reversion is by far the best trading strategy out of the 6. With a cumulative return of 23.199 units after 4 years, the strategy represents an increase of 22.19 units or a 2219% increase. The strategy that presents the lowest returns out of the 6 is the 10% stop loss strategy with relative mean reversion, which yields a return of 2.11 standardized units or an increase of 1.11 units or an 111% increase over the course of the four years. The absolute mean reversion strategies at different stop losses produce better cumulative returns than the same stop losses compared to relative mean reversion. The 5% absolute mean reversion strategy generates returns of 14.6 standardized units in comparison to the 5% relative mean reversion strategy, which produces returns of 4.35 standardized units over the course of the four years. The 10% absolute mean reversion strategy generates a cumulative return of 5.35 units, which is an increase on the 10% relative mean strategy.

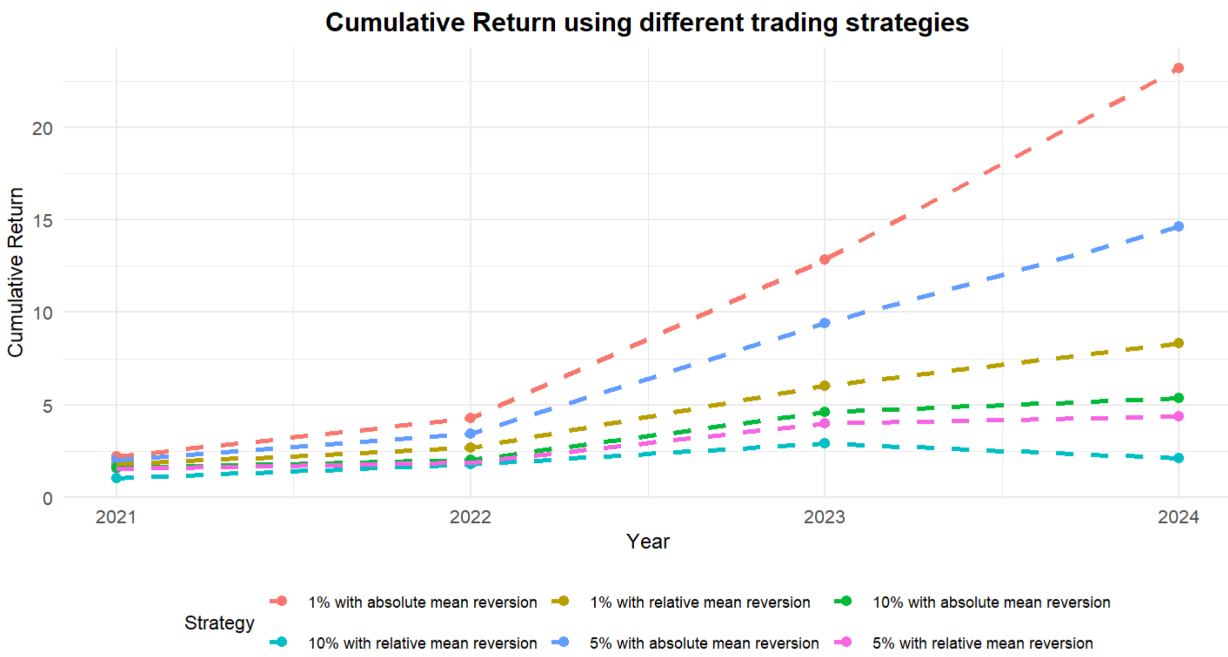

Figure 6: Cumulative Returns over the years (Author's own calculations)

Figure 6 shows the pathway of the cumulative returns of the six different trading strategies that are traded for the four years from 2021 to 2024. The graph shows the power of the compounding effect in the absolute mean reversion strategies which have a higher return for the first year and then the following years build on that. We also see that the difference in returns between the strategies widens as the years pass.

Now, we discuss in detail all the other important components of our trading strategies to provide a more detail view of our results.

v. Number of Trades

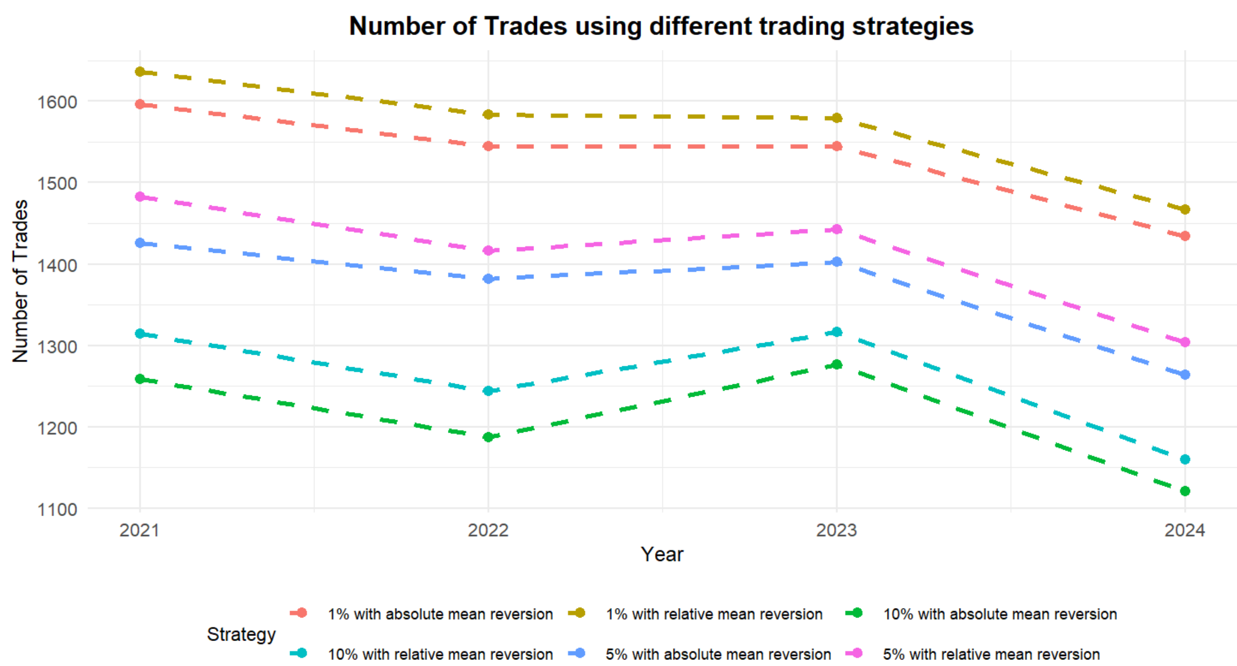

Figure 7: Number of Trades using different trading strategies (Author's own calculations)

The first discussion point is the number of trades for the six trading strategies. Figure 7 shows us that the trend for the total number of trades in a year for the six trading strategies almost remains the same. The year 2021 produces the greatest number of trades with the 1% absolute mean reversion strategy been traded 1596 times during the year. 2024, on the other hand, generates the least number of trades over the course of the year, with the 10% absolute mean reversion strategy only being traded 1121 times.

vi. Average PnL

Another important aspect that determines the success of any trading strategy is the Profit and Loss (PnL). PnL is of vital importance to us as it provides the effect of every individual trade on the overall cumulative returns. A strategy is considered to be successful when it minimizes it loses and maximizes its profits. To see this in full effect, we plot the average PnL of all trades and see where the major chunk of the PnL lies for all the trading strategies.

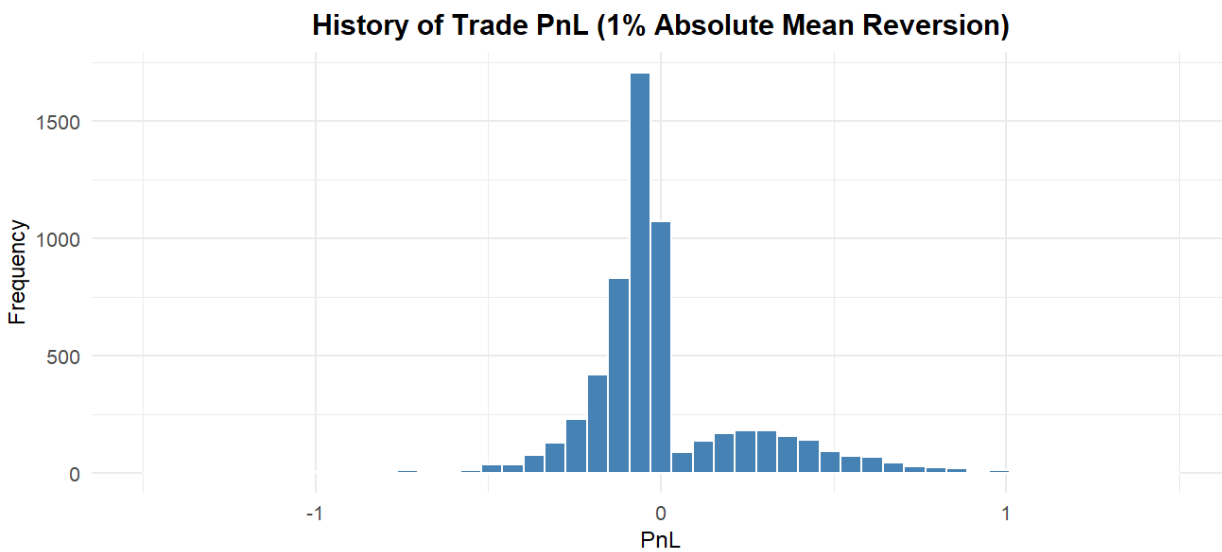

Figure 8: PnL Frequency (1% absolute mean reversion): Author's own calculations

Figure 8 shows the PnL Frequency of the 1% stop loss with absolute mean reversion strategy. We already know that our win rate for most of the strategies lie between the 35%-to-45%-win rate. Hence, it is expected to see a greater frequency of trades that produce a negative PnL. The sign of a good strategy as mentioned above is one which minimizes its losses. We can see from the figure that in the negative PnL section, most of the trades only generate a PnL loss that is less than 0.1 standardized units. On the other hand, we see that on the positive PnL side, we see more of a normal distribution and the profits are spread over and reach as far as PnL of 1, which translates to about a 100% increase as a result of a specific trade.

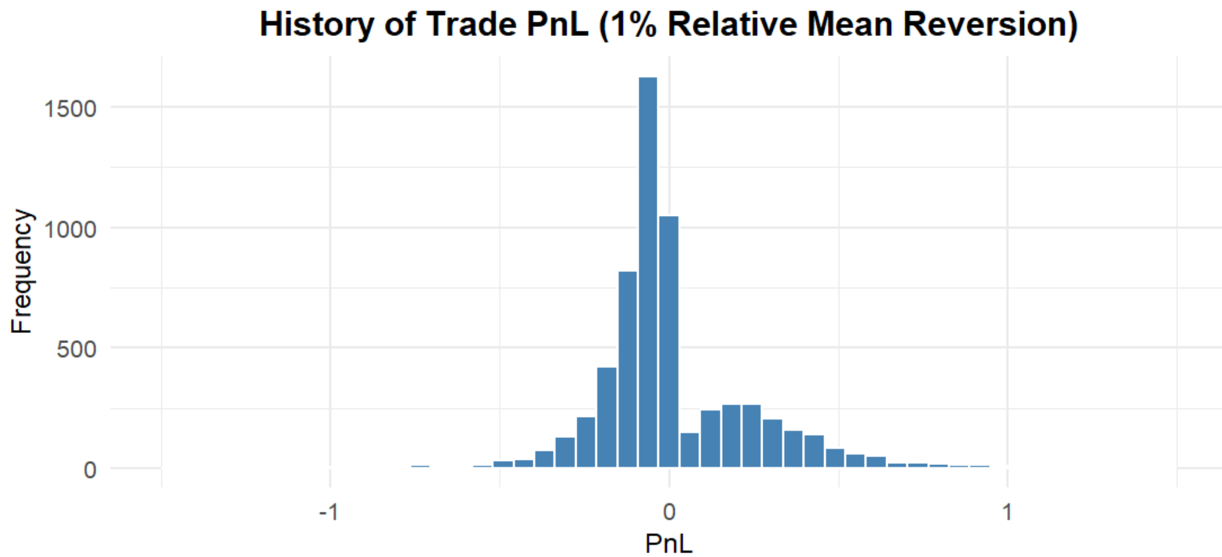

Figure 9: PnL Frequency (1% relative mean reversion): Author's own calculations

Figures 9 shows us a relatively similar picture to the 1% absolute mean reversion strategy, where we see that the win rates are lesser as compared to the 5% and 10% relative and absolute mean reversion trading strategies.

Figure A 1 and Figure A 2 in the appendix section show the PnL frequency for the 5% absolute and relative mean reversion strategies respectfully. We see that in these two cases the frequency of PnL that are in or around the 0 mark as very few. This means that in this case even though our win rate has increased, the losses suffered are bigger contributing to a relatively lesser return that the 1% stop loss strategy.

Figure A 3 and Figure A 4 in the appendix section show the PnL frequency for the 10% absolute and relative mean reversion strategies respectively. Here, the frequency of the negative PnL are lower but again the intensity is greater as we couldn't contain the losses contributing an even lesser return in comparison to the 1% and 5% stop loss trading strategies.

vii. Trade Level Breakdown

Now, we shall see the categorization of the trades that occurred for the six different strategies to gauge an even better view of the stocks being traded and how it affects the results. One such way is to study if trades are taking place between companies of the same sector or other sectors.

i. Sector Pairs Categorization

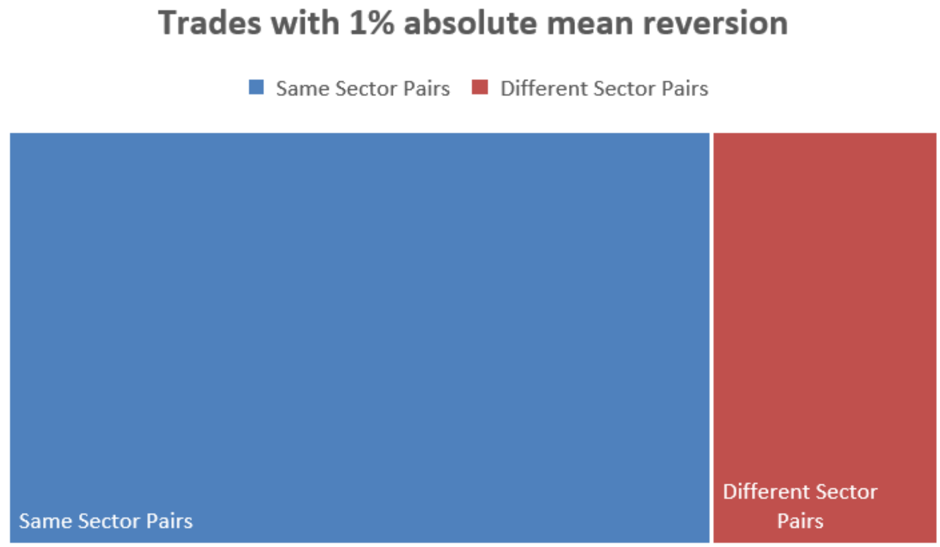

Figure 10: Sector Trade Categorization (1% Absolute Mean Reversion): Author's own calculations

Figure 10 shows that with the 1% stop loss with absolute mean reversion strategies, a total of 6119 trades took place, 4628 trades were same sector trades meaning that both the stocks belonged to the same sector. The remaining 1491 trades were traded between stocks of two different sectors.

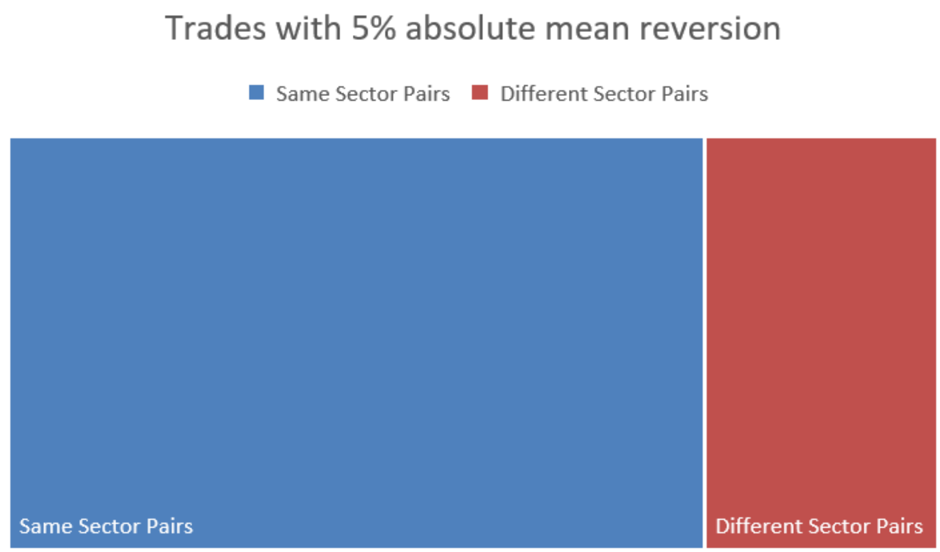

Figure 11: Sector Trade Categorization (5% absolute mean reversion): Author's own calculations

Figure 11 shows that there was a total of 5475 trades that took place as a result of the 5% absolute mean reversion strategy. Out of these, 4097 of these were same sector trades while the remaining were different sector trades.

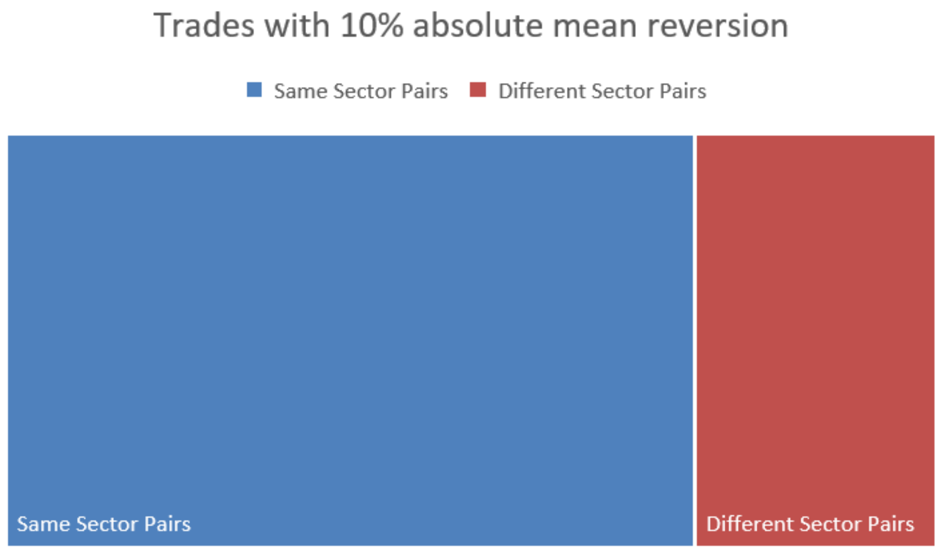

Figure 12: Sector Trade Categorization (10% absolute mean reversion): Author's own calculations

There were a total of 4844 trades with the 10% stop loss with absolute mean reversion strategy. Out of these, 74.07% or 3588 of these trades occurred between stocks of the same sectors as shown in Figure 12.

The categorization of same and different sector trades for the 1%,5% and 10% relative mean reversion strategy are given in the appendix section under A5, A6 and A7 respectively. The results for almost all the relative and absolute mean reversion strategies almost have a similar division of same sector and different sector trades, which highlights the consistency of the trading strategies.

Now, we will discuss in detail the top pairs of these strategies by two criteria.

ii. Top Pair by Win Rate

First, we will see the top pairs by win percentage, i.e., the stocks that have the highest win% of all the trades.

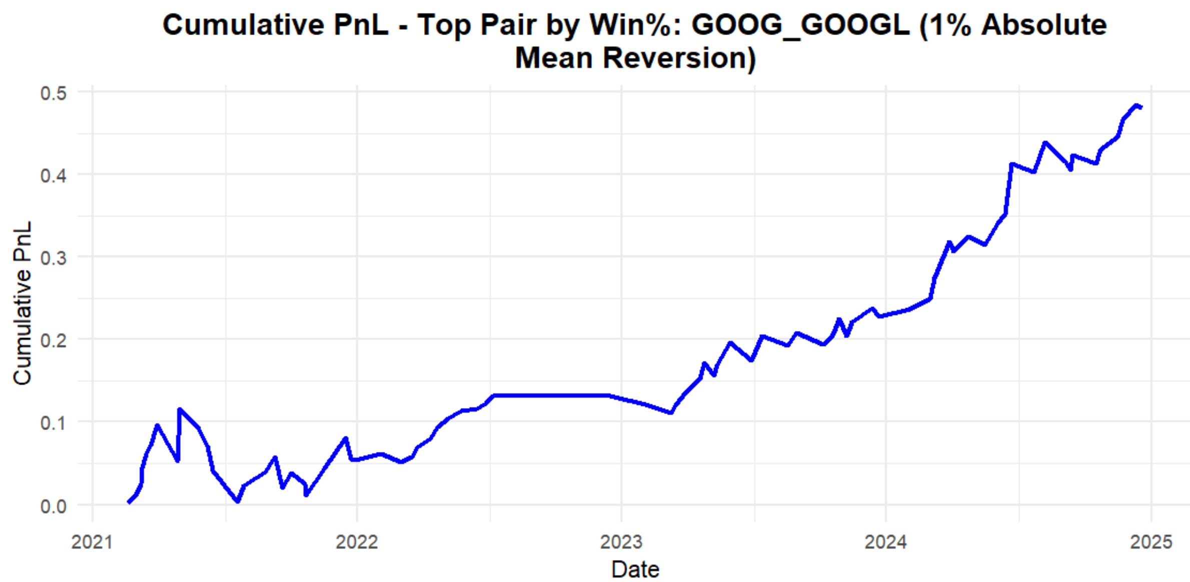

Figure 13: Top Pair by Win Rate (1% absolute mean reversion): Author's own calculations

Figure 13 shows us the pathway of the top trading pair with the 1% absolute mean reversion strategy. The top pair with respect to the win percentage for this strategy are GOOG (Alphabet Inc Class C) and GOOGL (Alphabet Inc Class A). Both the stocks not only represent the same sector but the same ownership in Alphabet. This pair was traded 78 times during the course of the four years with a success rate of about 64.1%. The cumulative PnL of this pair in terms of standardized units was about 0.5 standardized units.

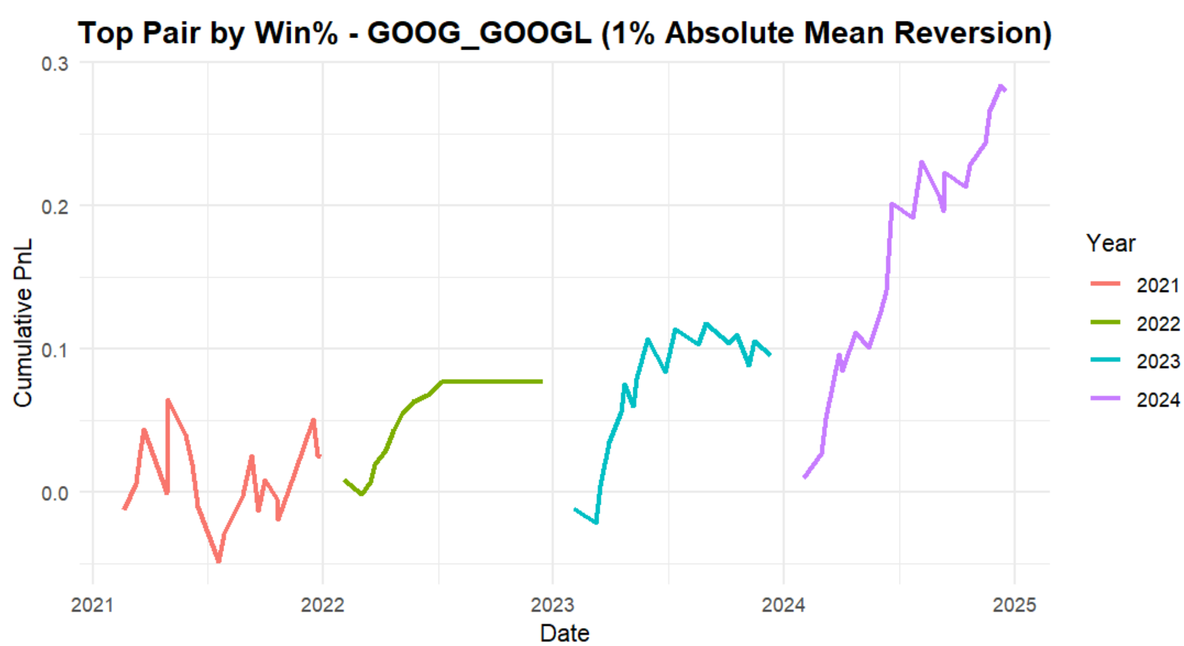

Figure 14: Yearly Breakdown of 1% Absolute Mean Reversion Top Pair (Author's own calculations)

Figure 12 breaks down the yearly performance of the top pair for the same 1% absolute mean reversion strategy. We can see from the figure that even though the success rate is very high, there is still variation in the yearly Cumulative PnL. 2024 was by far the best year in term of cumulative PnL for this top pair.

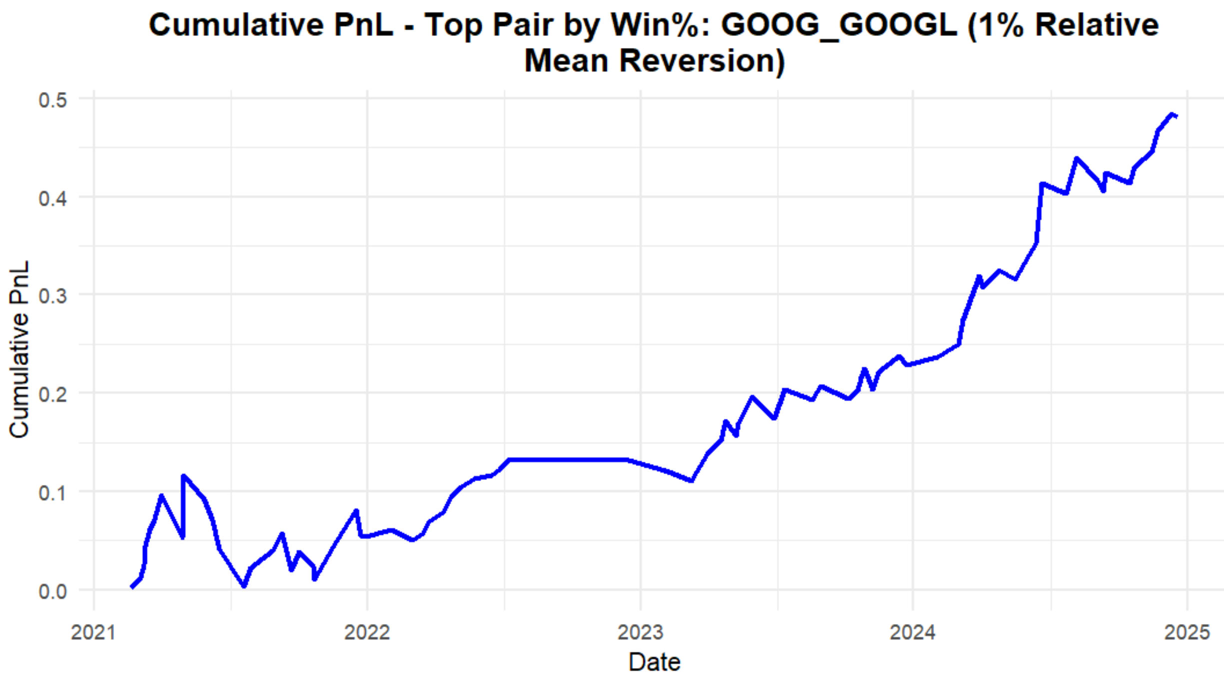

Figure 15: Top Pair by Win Rate (1% relative mean reversion): Author's own calculations

Figure 15 shows the top pair by win rate for the 1% relative strategy. The top pair for the 1% stop loss with relative mean reversion strategy was again GOOG and GOOGL. The pair was traded 85 times over the four years with a success rate of 67.06%. The cumulative PnL for this pair was less than 0.5 standardized units.

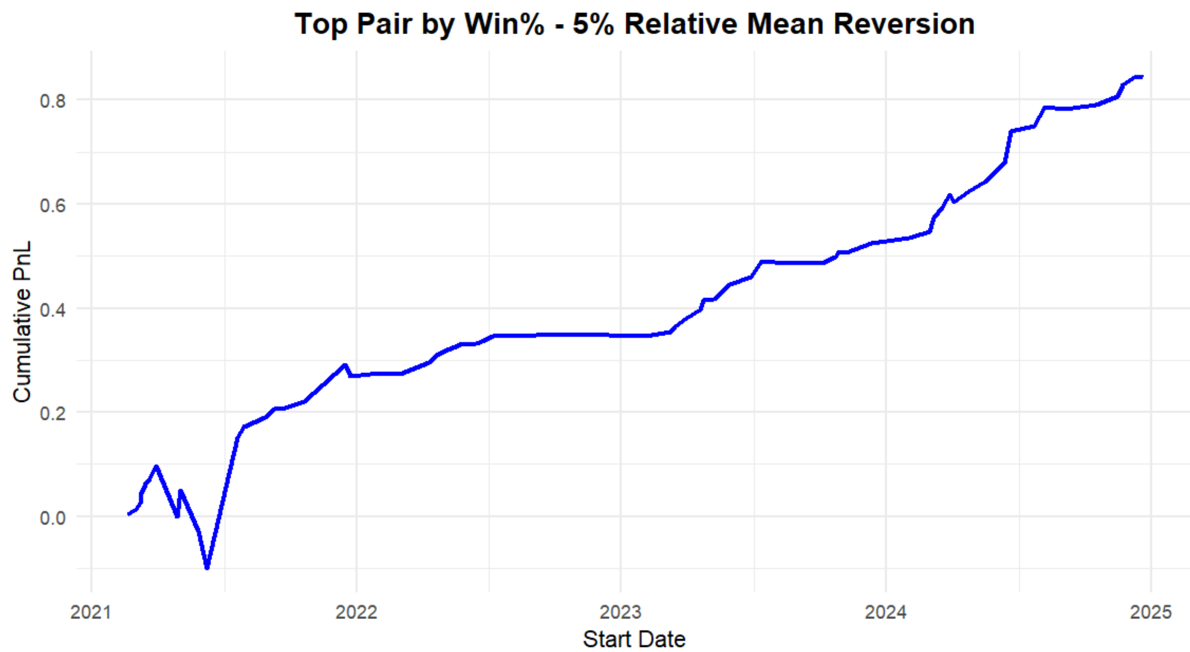

Figure 16: Top Pair by Win rate (5% relative mean reversion): Author's own calculations

Figure 16 shows the top pair by win rate for 5% relative mean reversion strategy. The top pair is again GOOG and GOOGL. The pair was traded a total of 73 times with a success rate of 79.45%. The cumulative PnL was about 0.8 standardized units for the strategy.

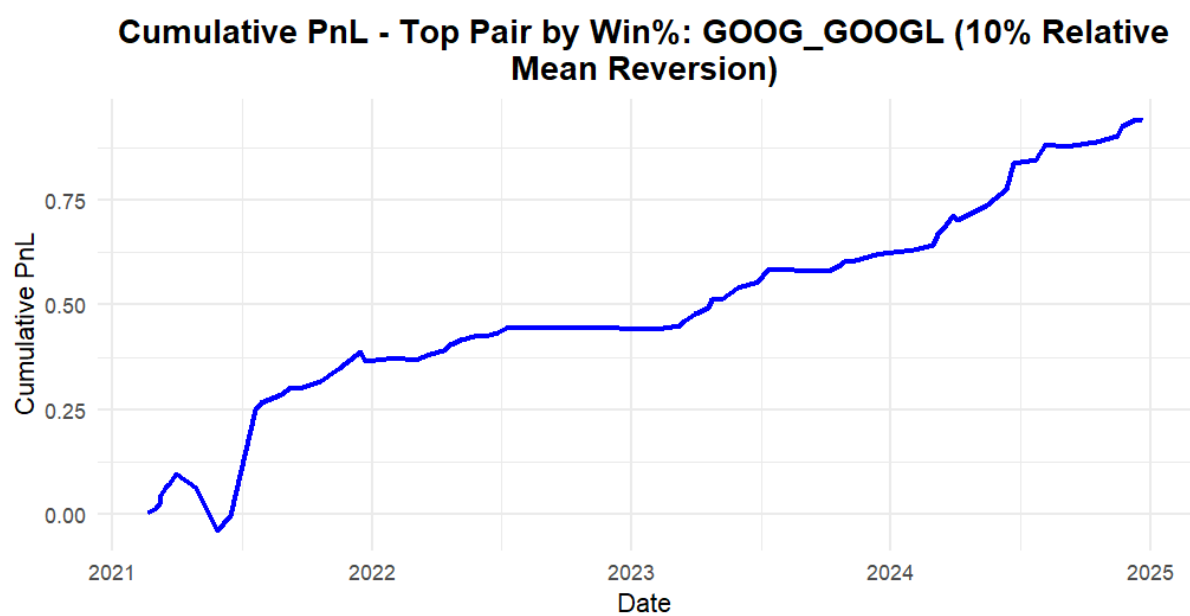

Figure 17: Top Pair by Win rate (10% relative mean reversion): Author's own calculations

Figure 17 shows the top pair by win rate for the 10% relative mean strategy. The top pair was once again GOOG and GOOGL and the pair was traded for a total of 72 times with a success rate of 80.56%. The cumulative PnL for the pair was just shy of the 0.8 standardized units mark.

The top pair for the 5% absolute and 10% absolute mean reversion strategy are provided in the appendix section in figures A8 and A9 respectively. The top pair for both these strategies were the same pair of GOOG and GOOGL with success rates of 76.92% and 78.12% respectively.

Alphabet Inc.'s Class A (GOOGL) and Class C (GOOG) stocks are such a good pair for the distance-based approach because they represent the same underlying company. The only difference between them is that the Class A (GOOGL) shareholders have 1 vote per share while Class C (GOOG) do not have any voting rights. This means that generally we find that the Class A's prices are relatively higher than Class C. However, we still expect them to have a strong direct relationship and to move in the same direction in regard to any good or bad news.

The divergence between these two stocks could occur due to short-term imbalances in supply and demand. For example, large institutions are expected to buy large chunks of Class A because they want the luxury of voting rights. Hence, It might take some time for the price to balance itself again. The economic value of both the companies are the same, the difference between them occurs in how the shareholders view them.

iii. Top 10 Pairs by Cumulative PnL

Now, we move on to the second criteria of categorizing the top pair, i.e., in order of the highest cumulative PnL that they generate from every trade over the four years.

Number	Pair	Cumulative_PnL
1	EBAY_TROW	6.31
2	MCK_OXY	6.31
3	MCK_XOM	6.27
4	IDXX_MTCH	6.05
5	DHR_WST	4.46
6	OXY_AMR	4.40
7	ETN_IR	3.78
8	KEY_PNC	3.65
9	IPG_TXT	3.64
10	AVGO_NVDA	3.11

Table 5: Top Pair by Cumulative PnL (1% relative mean reversion): Author's own calculations

Table 3 shows the top ten pairs of stocks, based on a 1% relative mean reversion strategy, ranked by their Cumulative PnL. This value represents the total profit generated by each pair over the analysis period. The table highlights that EBAY and TROW, along with MCK and OXY, were the most profitable pairs, both generating a cumulative PnL of $6.31. Of the top 10 pairs listed, four are composed of stocks from the same sector.

Rank	Pair	Cumulative_PnL
1	EBAY_TROW	7.01
2	AVB_UDR	5.09
3	KEY_PNC	4.66
4	IPG_TXT	4.57
5	DHR_WST	4.23
6	ANSS_MSFT	4.02
7	MCK_XOM	3.88
8	DVN_SLB	3.73
9	FITB_KEY	3.53
10	PHM_TDG	3.47

Table 6: Top Pair by Cumulative PnL (1% absolute mean reversion): Author's own calcualtions

Table 6 presents the top ten pairs of stocks identified by a 1% absolute mean reversion strategy, ranked by their Cumulative PnL. The table shows that EBAY and TROW were the most profitable pair, generating a cumulative PnL of $7.01. Of the top 10 pairs listed, six are composed of stocks from the same sector, highlighting the effectiveness of the strategy when applied to companies that operate in similar market environments.

Rank	Pair	Cumulative_PnL
1	IPG_TXT	4.85
2	OXY_AMR	4.76
3	KEY_PNC	4.73
4	AVB_UDR	4.45
5	CFG_KEY	4.36
6	EBAY_TROW	4.09
7	AVB_EQR	3.58
8	DHR_WST	3.53
9	ANSS_MSFT	3.32
10	CCL_SLB	3.29

Table 7: Top Pair by Cumulative PnL (5% absolute mean reversion): Author's own calculations

Table 7 shows the top ten pairs of stocks identified by a 5% absolute mean reversion strategy, ranked by their Cumulative PnL. The table highlights that IPG and TXT were the most profitable pair, generating a cumulative PnL of $4.85. Of the top 10 pairs listed, six are composed of stocks from the same sector.

Rank	Pair	Cumulative_PnL
1	MCK_OXY	7.09
2	EBAY_TROW	7.01
3	MCK_XOM	5.90
4	IPG_TXT	4.69
5	KEY_PNC	4.22
6	CFG_RF	4.19
7	DHR_WST	3.94
8	AVB_EQR	3.89
9	KEY_TFC	3.73
10	IDXX_MTCH	3.29

Table 8: Top Pair by Cumulative PnL (10% absolute mean reversion): Author's own calculations

Table 8 presents the top ten pairs of stocks identified by a 10% absolute mean reversion strategy, ranked by their Cumulative PnL. The table shows that MCK and OXY were the most profitable pair, generating a cumulative PnL of $7.09. Of the top 10 pairs listed, five are composed of stocks from the same sector.

The top 10 pairs by cumulative PnL for 5% and 10% relative mean reversion are provided in the appendix section in tables A1 and A2. There is no significant contribution but an even one of same and different stocks pair combinations contributing to the top 10 pairs for the six strategies.

Now, that we have gone through the top pairs of the strategies, we will now move to a discussion of the sectors in which these trades took place. We will have this discussion in two parts. In first part, we will discuss the top 10 combination pairs of trades that took place.

iv. Top 10 Trade Pair Sector Combinations

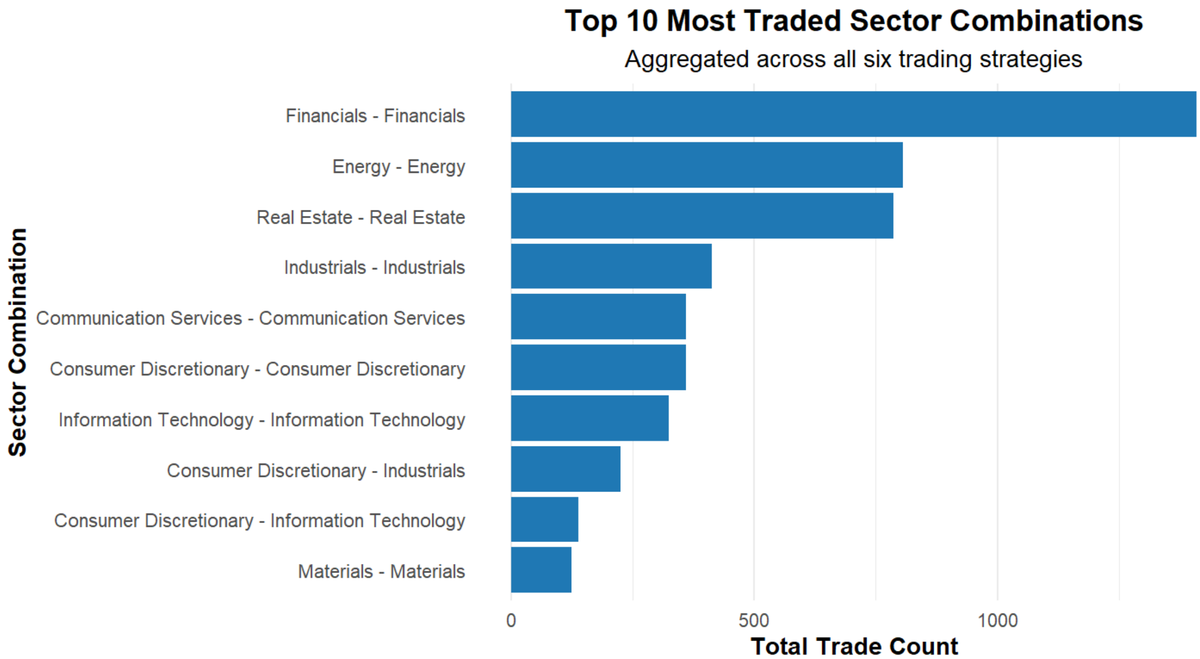

Figure 18: Top 10 Aggregated Traded Sector Combinations: Author's own calculations

Figure 18 shows the top 10 aggregated traded sector combinations. The top pair is both stocks from the financial sector. All the top 7 top sector combinations are of stocks from the same sectors emphasizing the strong results of the distance-based approach.

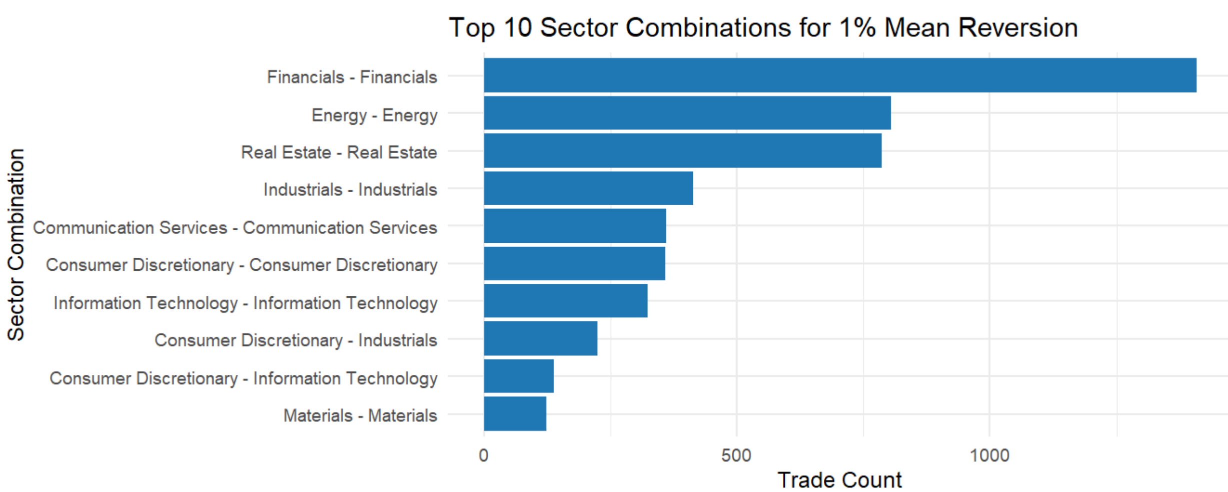

Figure 19: Top 10 Sector Combinations for 1% Mean Reversion: Author's own calculations

The top sector combination was financial and financial for the 1% absolute mean reversion strategy. The top combination was traded 1409 times while the second-best combination of energy and energy contributed 805 trades.

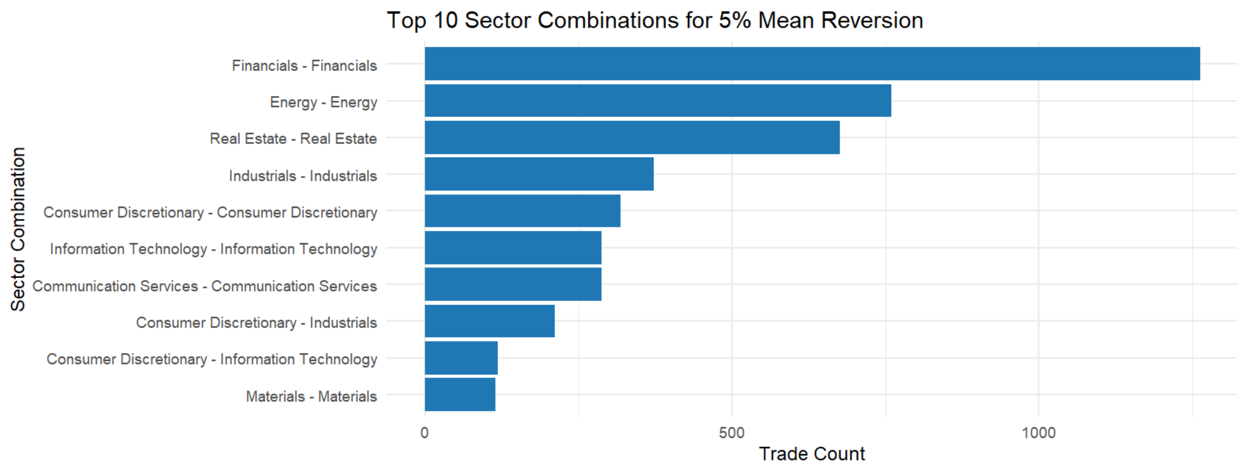

Figure 20: Top 10 Sector Combinations for 5% Mean Reversion: Author's own calculations

The top sector combination again was financial and financial for the 5% absolute mean reversion strategy. The top combination was traded 1228 times while the second-best combination of energy and energy contributed 742 trades.

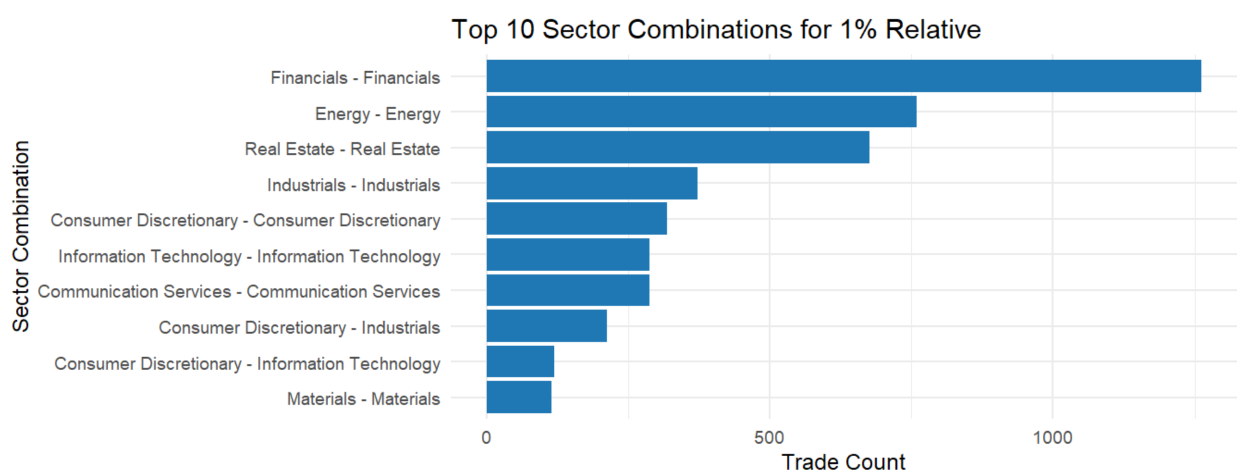

Figure 21: Top 10 Sector Combinations for 1% relative mean reversion: Author's own calculations

The top 10 sector combinations for the 1% relative mean reversion strategy were the same sector combinations as the 1% and 5% absolute mean reversion strategy. The top pair financial and financial was traded 1262 times while the second-best pair was traded 759 times.

The top 10 sector combinations for 5% relative, 10% relative and absolute mean reversion strategies are given in the appendix section under

v. Individual Sector Trades

Now, we will have a discussion of trades using the categorization of trades in individual sectors. This will classify how many times a sector has been a part of either of the pair or both when traded.

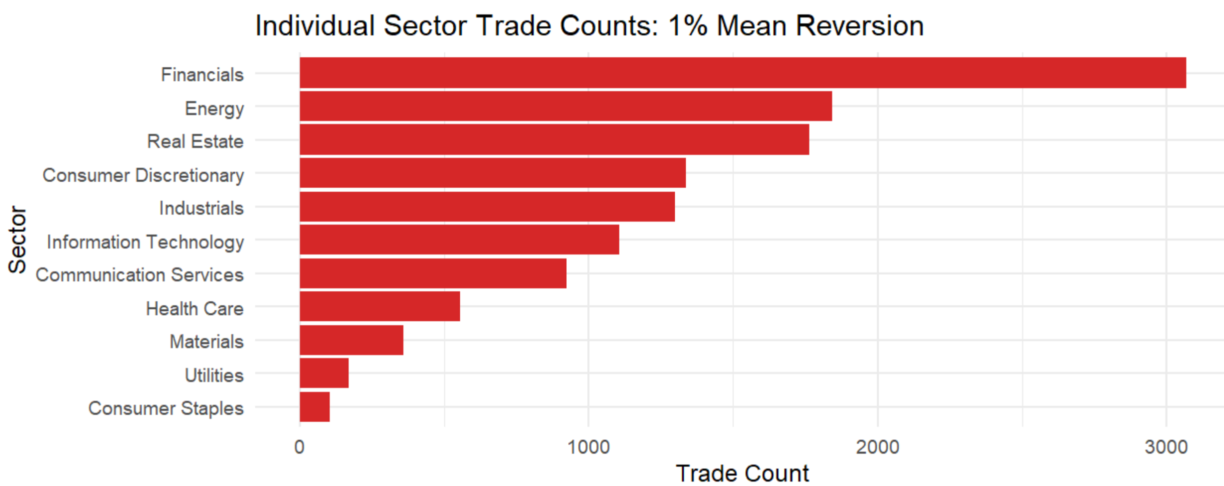

Figure 22: Individual Sector Trades (1% Absolute Mean Reversion): Author's own calculations

Figures 22 shows us that the financial sector's stocks have been traded the most for the 1% absolute mean reversion , with 3068 trades across the four years. Energy stocks from the S&P 500 have been traded 1844 times, Real Estate stocks about 1765 times.

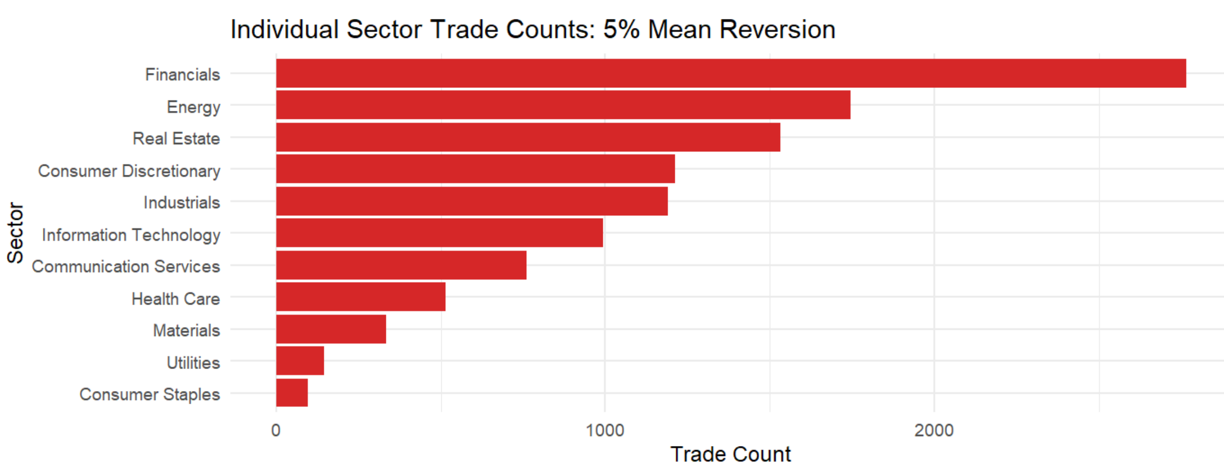

Figure 23: Individual Sector Trades (5% Absolute mean reversion): Author's own calculations

Figure 23 shows the individual sector trades for 5% absolute mean reversion strategy. Stocks from the financial sectors were again traded the most number of times, about 2765 times. The top 10 sector stocks were almost the same as the top 10 for 1% absolute mean reversion strategy.

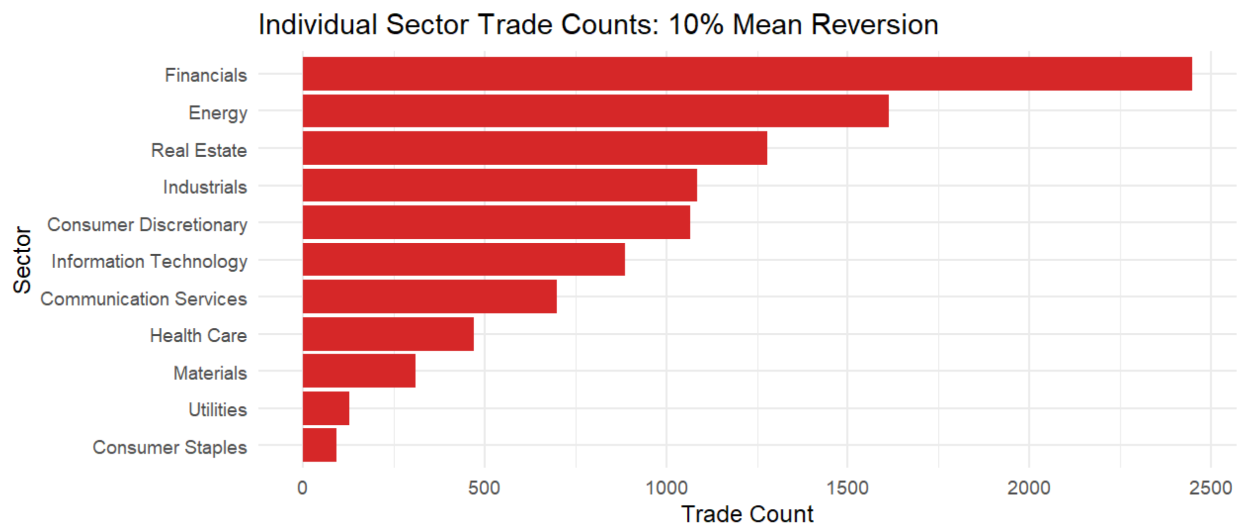

Figure 24: Individual Sector Trades (10% Absolute mean reversion): Author's own calculations

Figure 24 shows the individual sector trade count for the 10% absolute mean reversion strategy. Financial, energy and real estate sectors were again the top 3 with 2447, 1612 and 1279 trades respectively.

The individual sector trade count for the 1%, 5% and 10% relative mean reversion strategies are provided in the appendix section under figures A13, A14 and A15. The top 10 individual sectors were almost in the same order for all the strategies except for a couple of changes in ranks here and there. This suggests strong consistency in the strategies and their respective results.

vi. Win Rate of Different Sectors

The final discussion point would be the win rate of different sectors. This could help identify sectors that could be targeted further to increase the potential returns. We shall see the top 10 sector combinations with respect to their win rate for the six trading strategies.

The top 10 sector combinations are provided in the appendix section. The top sector combinations for all the six different trading strategies are provided through tables A3 to A8. Here, we shall discuss the common sector combinations found throughout the six trading strategies by using a heatmap.

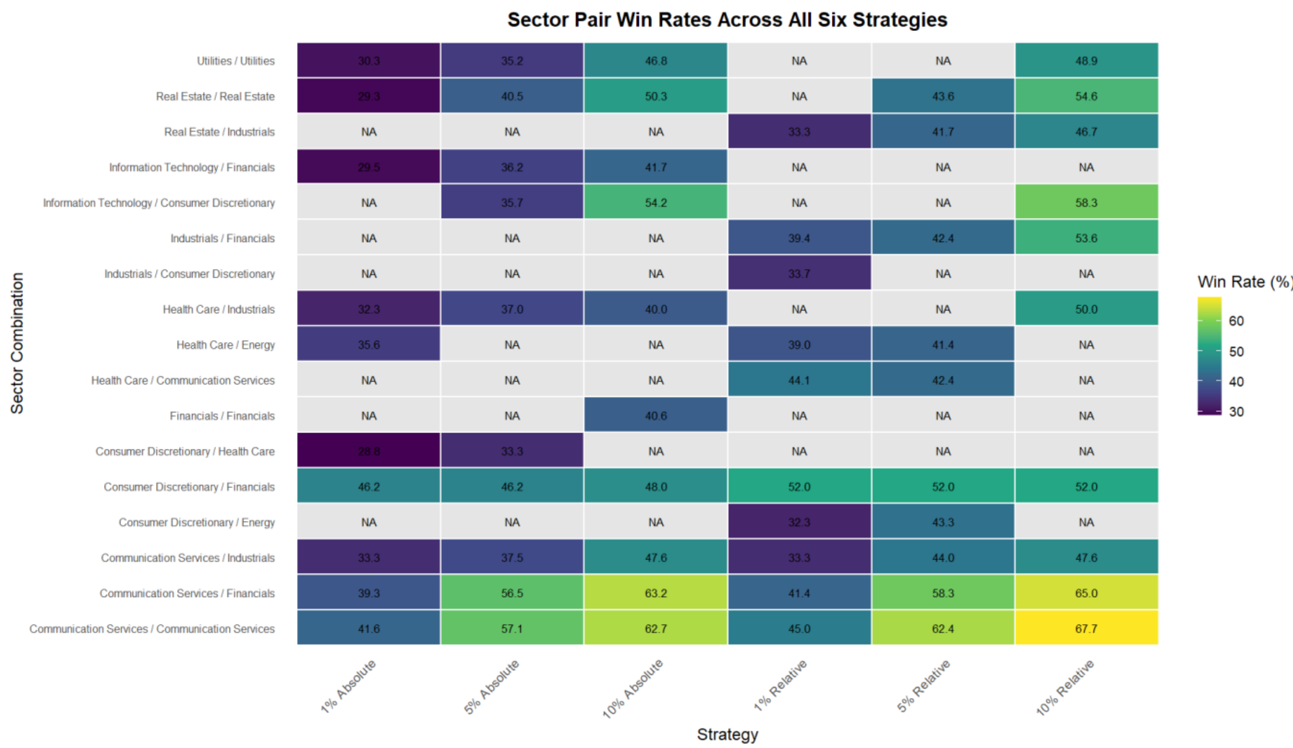

Figure 25: Heatmap of Top Sector Pairs across the six trading strategies: Author's own calculations

From the figure 25, we can see that higher success rates sector pairs were found in the 1%,5% or 10% stop loss with relative mean reversion strategies. The success rate for the 1% absolute mean reversion was already close to 30%, so it makes sense that even top pair combinations have success rate closer to 40%. With the use of these figures, we can identify further the sectors in which we can trade with relatively less risk and can afford to provide a relatively larger capital in comparison to the other sector combinations.

While the results demonstrate the efficacy of the distance-based approach, the study has certain limitations. Daily closing prices were used, excluding intraday and after-hours price movements. Transaction costs and slippage were not incorporated, which may impact practical performance. The analysis was restricted to a six-year period and the top 50 pairs. Future research applying these parameters; potentially over longer timeframes, with dynamic pair selection, or using higher-frequency data could provide guidance for achieving even more robust and insightful results

Conclusion

This research confirms the critical role of a tightly controlled stop-loss in maximizing returns and minimizing risk within a distance-based pairs trading framework. Our analysis, which focused on a direct distance-based approach for identifying pairs with minimal price divergence; a methodology that differs from the more traditional cointegration or correlation-based models prevalent in the recent existing literature, yielded several important findings. Most notably, the 1% Absolute mean reversion strategy consistently emerged as the top performer, achieving a remarkable cumulative PnL of 23.19996 by the end of 2024. This success can be attributed to the compounding effect of consistent gains over time. The analysis also revealed a high degree of consistency in both the win rates and Sharpe ratios, which reinforces the robustness of this particular strategy.

While the use of daily price data means the stop-loss mechanism is not as precise as an intraday approach, our results clearly demonstrate its effectiveness as a crucial risk management tool. The tight stop-loss on the 1% Absolute strategy allowed for the efficient closing of losing trades, protecting capital and enabling the compounding of winning trades. It is important to note that these impressive results do not incorporate transaction costs, which would impact the final profitability.

A deeper analysis of the pairs traded also addressed a key theoretical issue: the effectiveness of a distance-based model in handling pairs of stocks that are fundamentally identical but trade under different tickers, such as GOOG and GOOGL. Our model successfully identified and capitalized on the temporary divergences that occurred between these very closely related stocks, demonstrating its ability to exploit market inefficiencies. While an analysis of individual sectors showed frequent trades within the Financials sector, the core success of our strategies lies in the model's ability to find and profit from these temporary dislocations regardless of industry classification. This paper, therefore, sets the stage for further work on the distance-based approach, re-evaluating its potential in a modern context and challenging the notion that it has been made redundant by the more widely used cointegration models.

Appendix

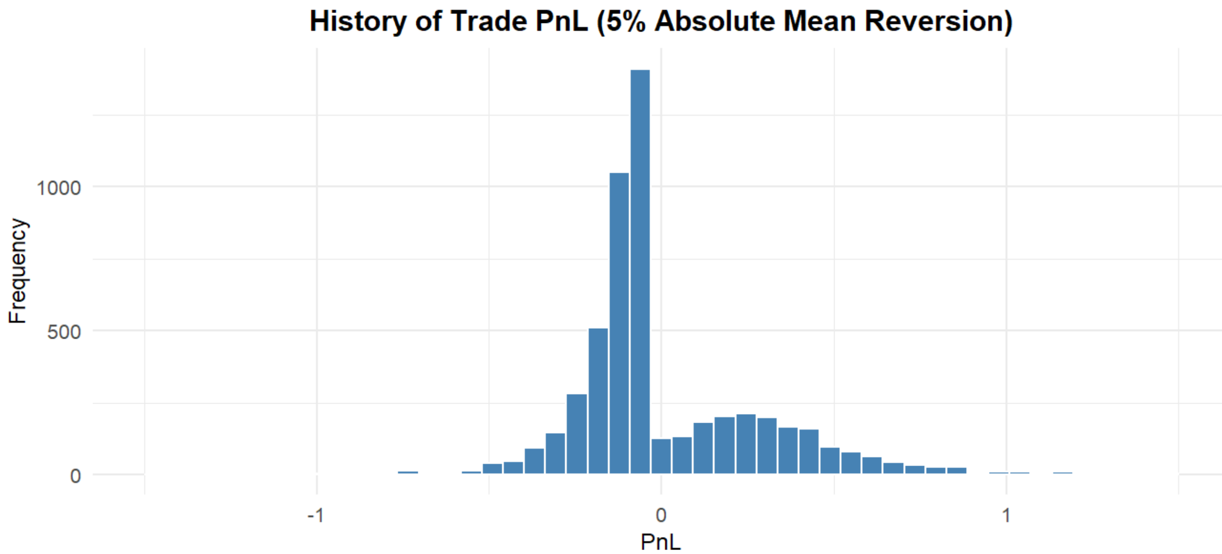

Figure A 1: PnL Frequency (5% absolute mean reversion): Author's own calculations

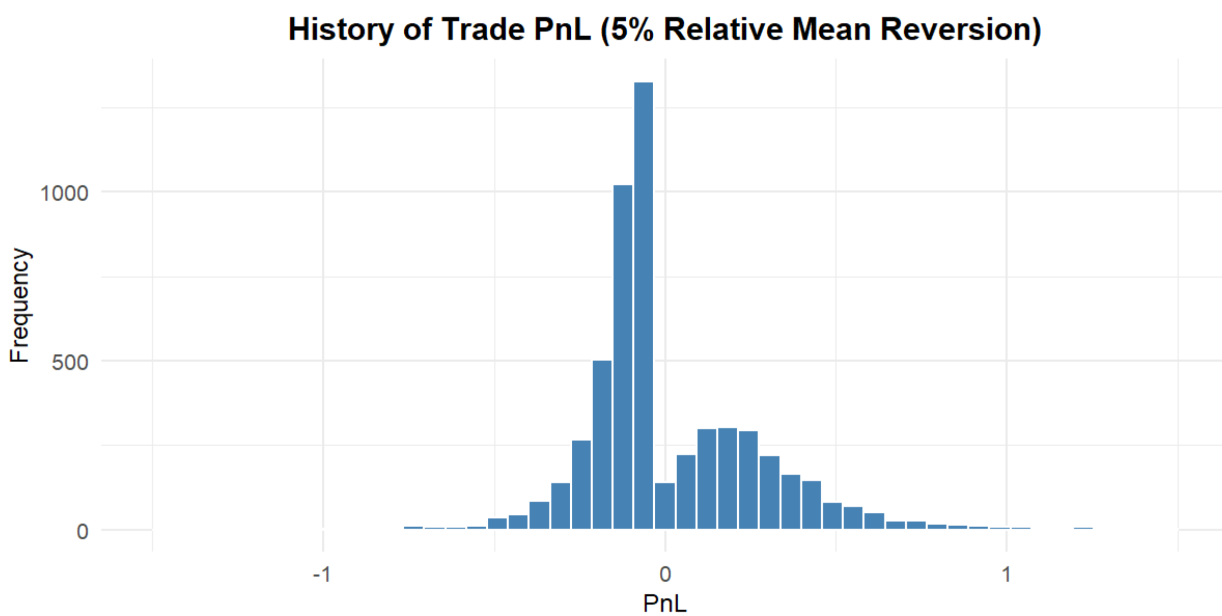

Figure A 2: PnL Frequency (5% relative mean reversion): Author's own calculations

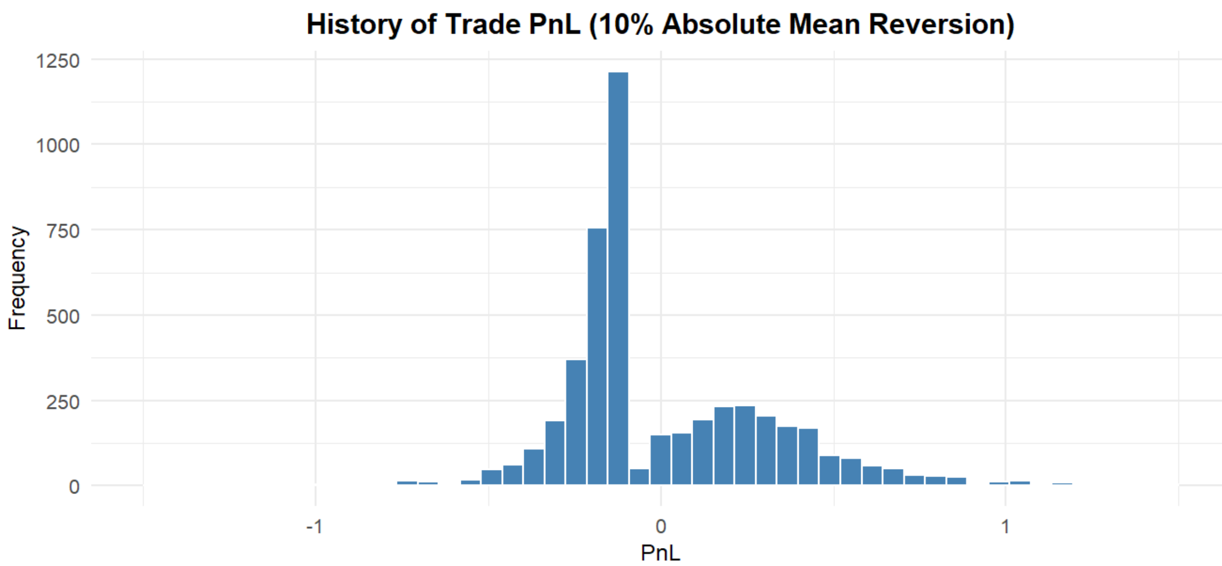

Figure A 3: PnL Frequency (10% absolute mean reversion): Author's own calculations

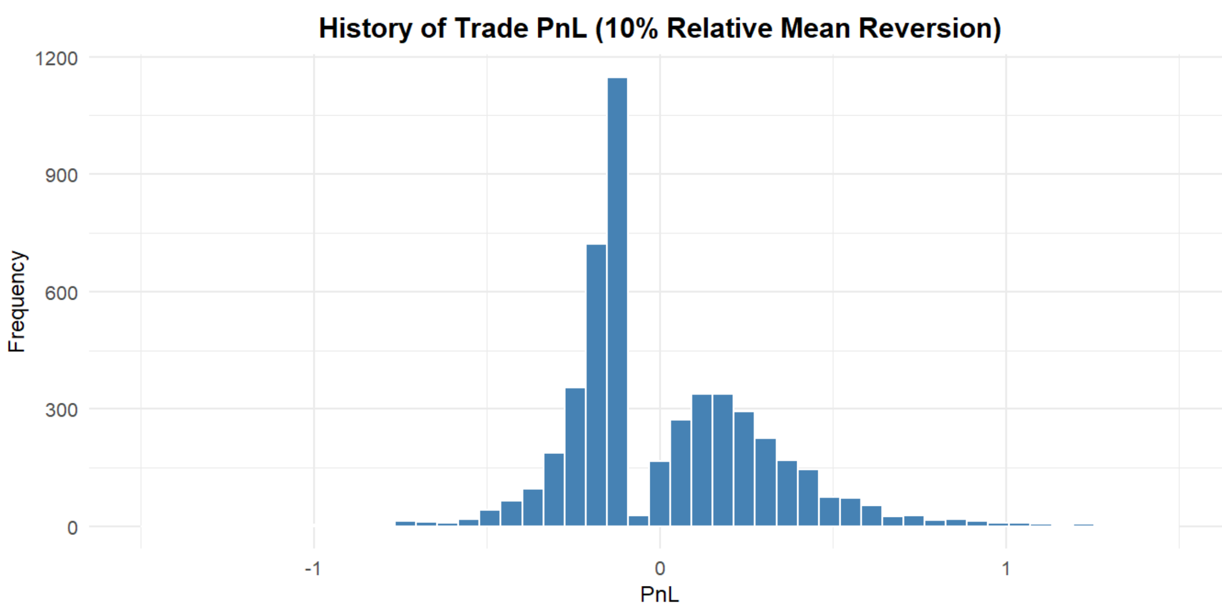

Figure A 4: PnL Frequency (10% relative mean reversion): Author's own calculations

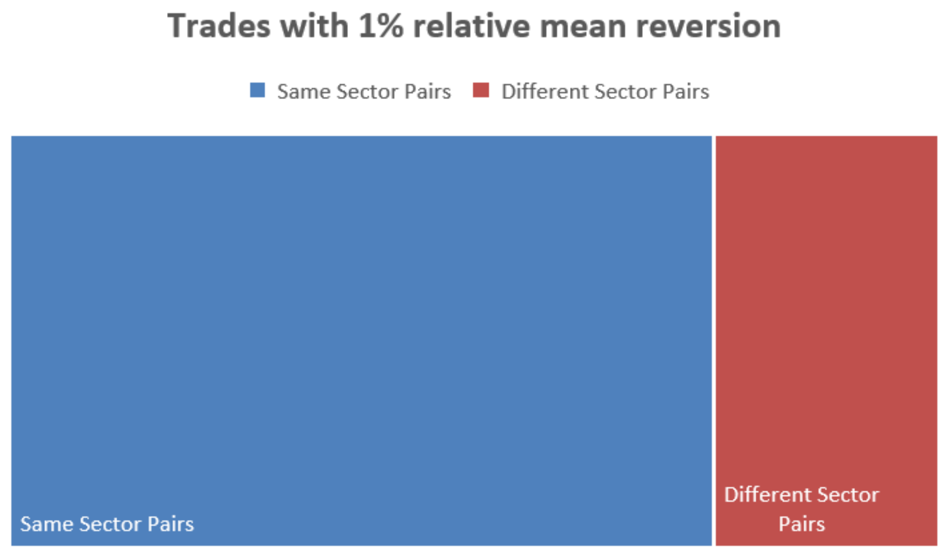

Figure A 5: Sector Trade Categorization (1% relative mean reversion): Author's own calculations

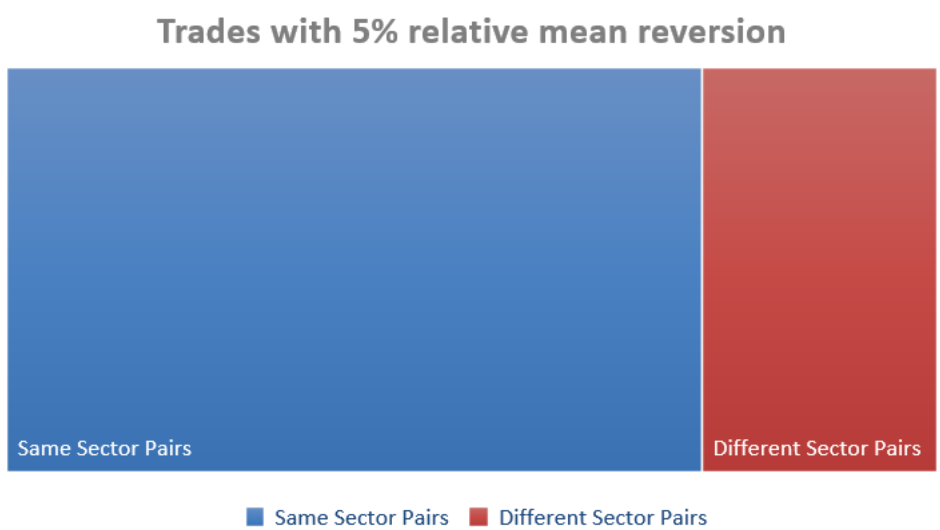

Figure A 6: Sector Trade Categorization (5% relative mean reversion): Author's own calculations

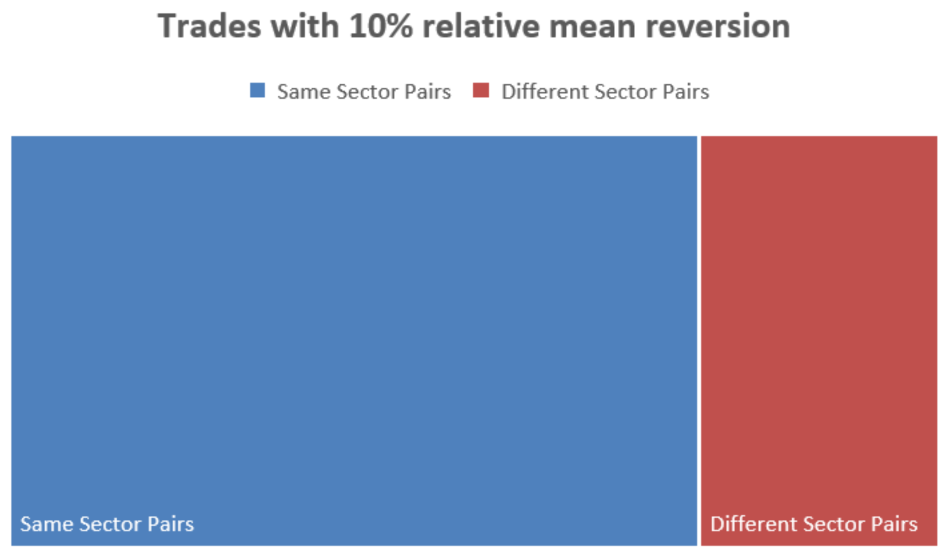

Figure A 7: Sector Trade Categorization (10% relative mean reversion): Author's own calculations

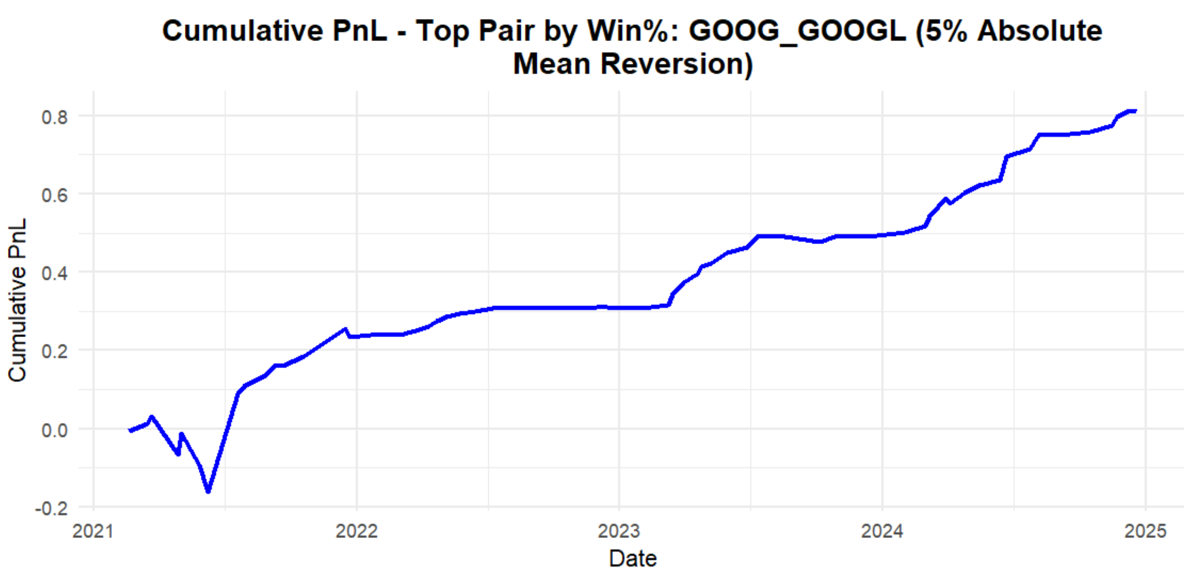

Figure A 8: Top Pair by Win Rate (5% absolute mean reversion): Author's own calculations

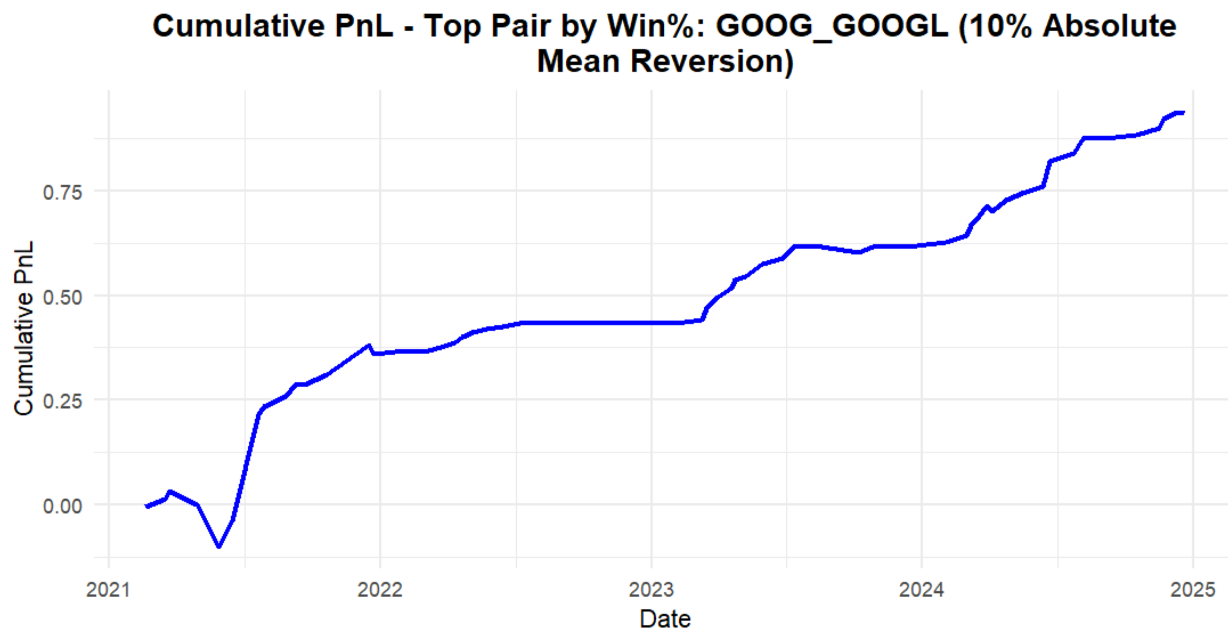

Figure A 9: Top Pair by Win Rate (10% absolute mean reversion): Author's own calculations

Number	Pair	Cumulative_PnL
1	MCK_XOM	5.06
2	DHR_WST	5.01
3	OXY_AMR	4.48
4	KEY_PNC	4.38
5	DVN_CHK	4.37
6	MCK_OXY	4.12
7	CFG_KEY	3.95
8	IPG_TXT	3.89
9	AVB_UDR	3.84
10	EBAY_TROW	3.49

Table A 1: Top Pair by Cumulative PnL (5% relative mean reversion)

Number	Pair	Cumulative_PnL
1	MCK_XOM	7.26
2	EBAY_TROW	6.31
3	DHR_WST	4.98
4	OXY_AMR	4.93
5	IPG_TXT	4.54
6	AVB_UDR	4.27
7	MCK_OXY	4.12
8	KEY_PNC	3.99
9	IDXX_MTCH	3.84
10	FITB_KEY	3.03

Table A 2: Top Pair by Cumulative PnL (10% relative mean reversion): Author's own calculations

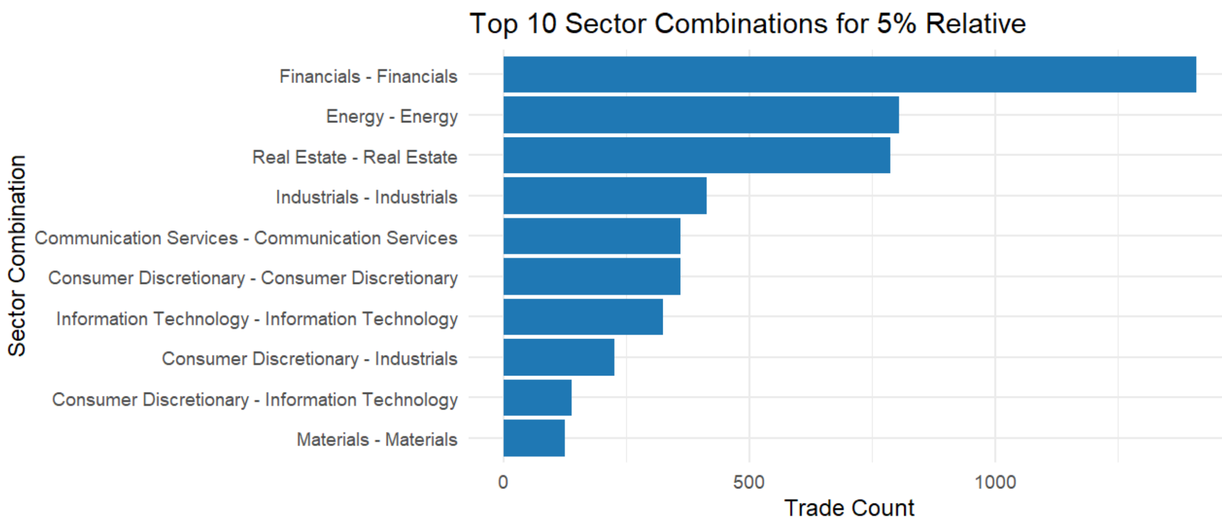

Figure A 10: Top 10 Sector Combinations for 5% relative mean reversion: Author's own calculations

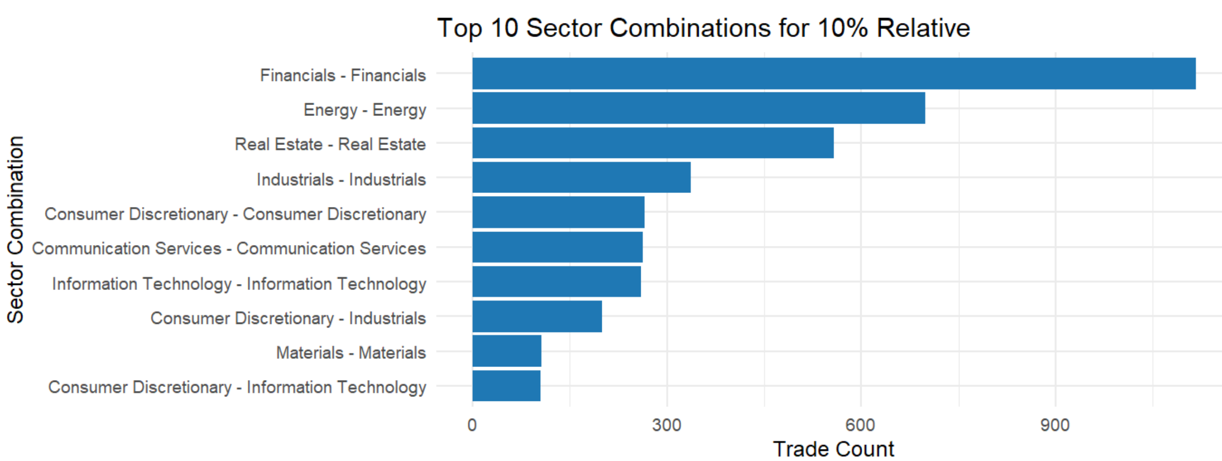

Figure A 11: Sector Combination for 10% relative mean reversion: Author's own calculations

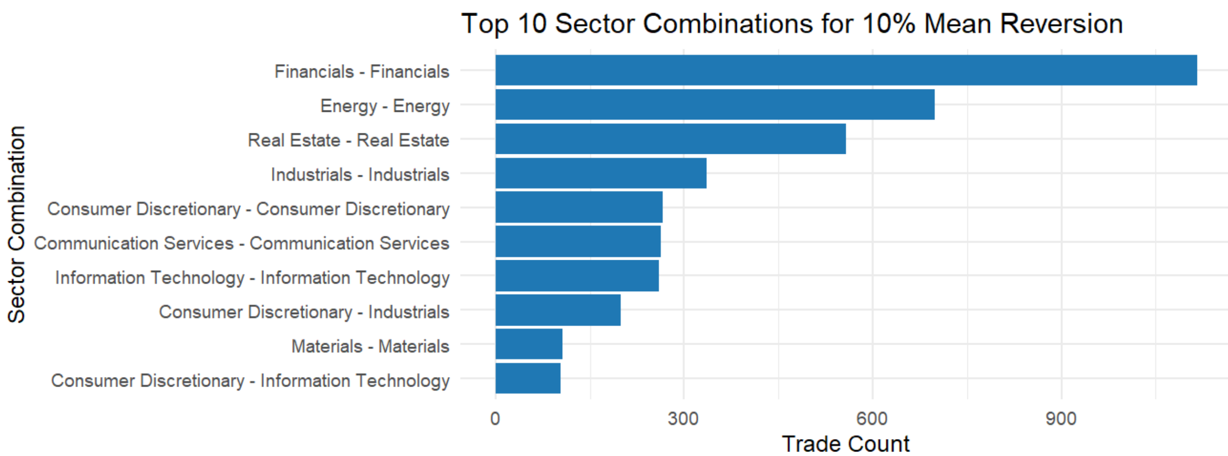

Figure A 12: Sector Combinations for 10% Mean reversion strategy: Author's own calculations

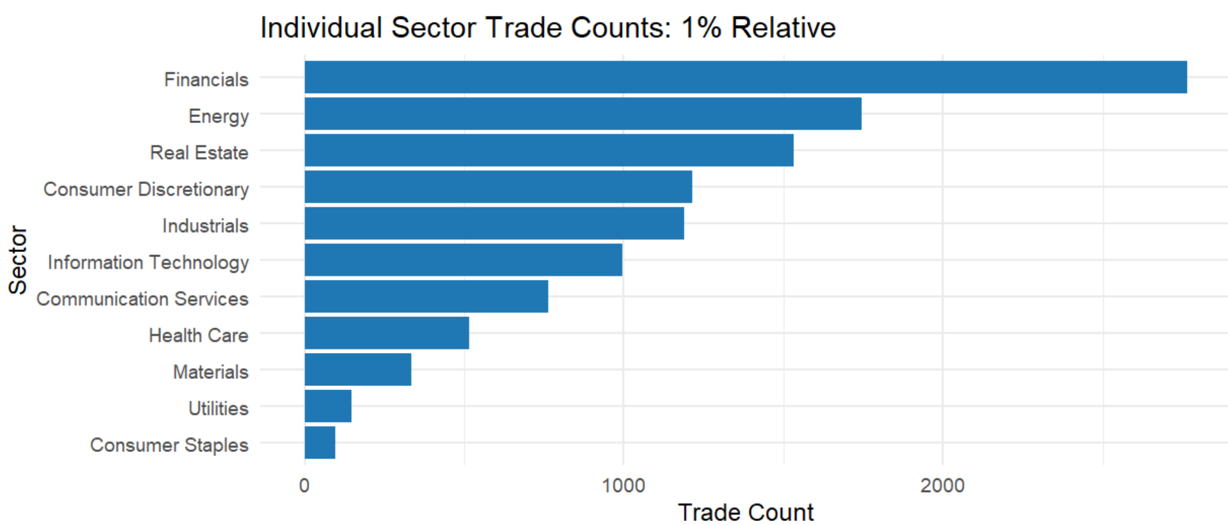

Figure A 13: Individual Sector Trades (1% relative mean reversion): Author's own calculations

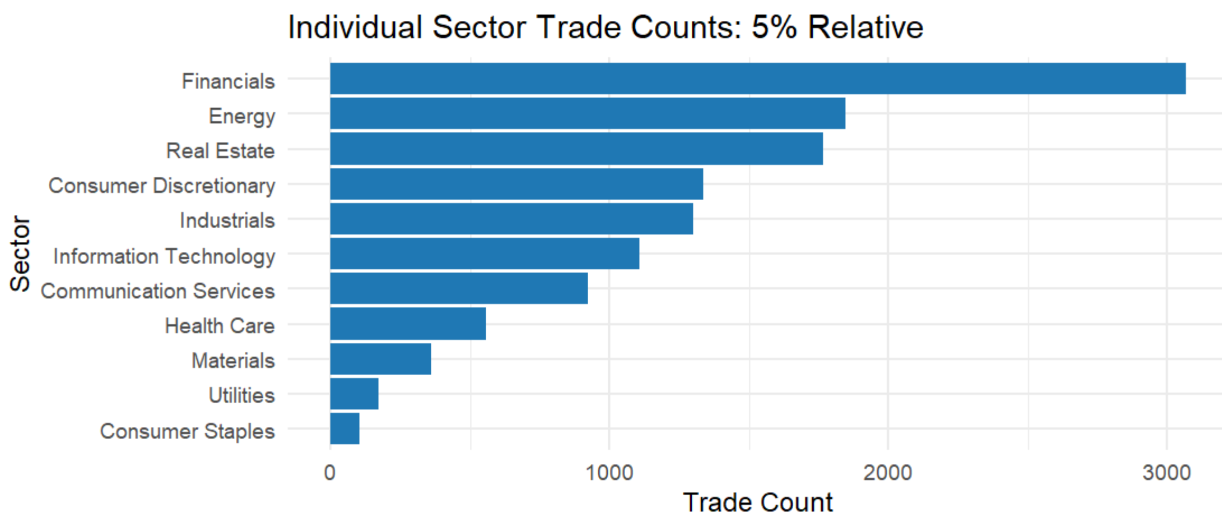

Figure A 14: Individual Sector Trades (5% relative mean reversion): Author's own calculations

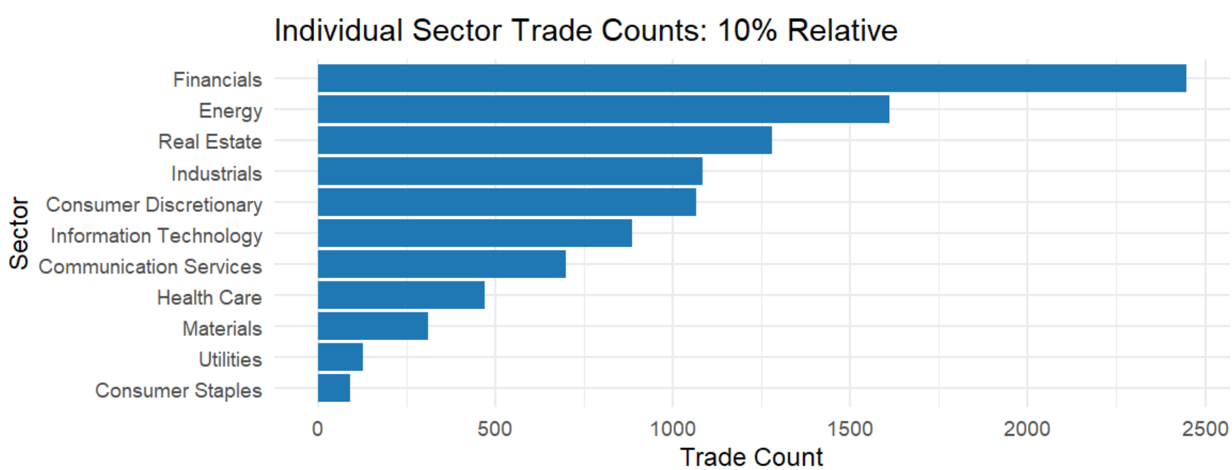

Figure A 15: Individual Sector Trade for 10% relative mean reversion strategy: Author's own calculations

Rank	Sector Combination	Win Rate (%)
1	Consumer Discretionary / Financials	52.0
2	Communication Services / Communication Services	45.0
3	Health Care / Communication Services	44.1
4	Communication Services / Financials	41.4
5	Industrials / Financials	39.4
6	Health Care / Energy	39.0
7	Industrials / Consumer Discretionary	33.7
8	Communication Services / Industrials	33.3
9	Real Estate / Industrials	33.3
10	Consumer Discretionary / Energy	32.3

Table A 3: Top 10 Sector Combinations by Win Rate (1% relative mean reversion strategy): Author's own calculations

Rank	Sector Combination	Win Rate (%)
1	Communication Services / Communication Services	62.4
2	Communication Services / Financials	58.3
3	Consumer Discretionary / Financials	52.0
4	Communication Services / Industrials	44.0
5	Real Estate / Real Estate	43.6
6	Consumer Discretionary / Energy	43.3
7	Health Care / Communication Services	42.4
8	Industrials / Financials	42.4
9	Real Estate / Industrials	41.7
10	Health Care / Energy	41.4

Table A 4: Top 10 Sector Combinations by Win Rate (5% relative mean reversion strategy): Author's own calculations

Rank	Sector Combination	Win Rate (%)
1	Communication Services / Communication Services	67.7
2	Communication Services / Financials	65.0
3	Information Technology / Consumer Discretionary	58.3
4	Real Estate / Real Estate	54.6
5	Industrials / Financials	53.6
6	Consumer Discretionary / Financials	52.0
7	Health Care / Industrials	50.0
8	Utilities / Utilities	48.9
9	Communication Services / Industrials	47.6
10	Real Estate / Industrials	46.7

Table A 5: Top 10 sector combinations by win rate (10% relative mean reversion): Author's own calculations

Rank	Sector Combination	Win Rate (%)
1	Consumer Discretionary / Financials	46.2
2	Communication Services / Communication Services	41.6
3	Communication Services / Financials	39.3
4	Health Care / Energy	35.6
5	Communication Services / Industrials	33.3
6	Health Care / Industrials	32.3
7	Utilities / Utilities	30.3
8	Information Technology / Financials	29.5
9	Real Estate / Real Estate	29.3
10	Consumer Discretionary / Health Care	28.8

Table A 6: Top 10 sector combinations by win rate (1% absolute mean reversion): Author's own calculations

Rank	Sector Combination	Win Rate (%)
1	Communication Services / Communication Services	57.1
2	Communication Services / Financials	56.5
3	Consumer Discretionary / Financials	46.2
4	Real Estate / Real Estate	40.5
5	Communication Services / Industrials	37.5
6	Health Care / Industrials	37.0
7	Information Technology / Financials	36.2
8	Information Technology / Consumer Discretionary	35.7
9	Utilities / Utilities	35.2
10	Consumer Discretionary / Health Care	33.3

Table A 7: Top 10 sector combinations by win rate (5% absolute mean reversion): Author's own calculations

Rank	Sector Combination	Win Rate (%)
1	Communication Services / Financials	63.2
2	Communication Services / Communication Services	62.7
3	Information Technology / Consumer Discretionary	54.2
4	Real Estate / Real Estate	50.3
5	Consumer Discretionary / Financials	48.0
6	Communication Services / Industrials	47.6
7	Utilities / Utilities	46.8
8	Information Technology / Financials	41.7
9	Financials / Financials	40.6
10	Health Care / Industrials	40.0

Table A 8: Top 10 sector combinations by win rate (10% absolute mean reversion): Author's own calculations

Bibliography

1. Bowen, D., Hutchinson, M. C., & O'Sullivan, N. (2010). High Frequency Equity Pairs Trading: Transaction Costs, Speed of Execution and Patterns in Returns. *Journal of Trading*.

2. Braudel, F. (1982). *The wheels of commerce.*

3. Cboe. (2025). *North American Equities Year in Review.*

4. Chen, H. (., Chen, S. J., Chen, Z., & Li, F. (2017). Empirical Investigation of an Equity Pairs Trading Strategy.

5. Chuen, D. L. (2015). *Handbook of Digital Currency: Bitcoin, Innovation, Financial Instruments, and Big Data.* London: Academic Press.

6. Companiesmarketcap.com. (2025). *Countries Ranked by Market Cap.*

7. Cronon, W. (1991). *Nature's Metropolis: Chicago and the Great West.* New York: W.W.Norton.

8. CRSP. (2024). *WHARTON RESEARCH DATA SERVICES.*

9. Do, B. H., & Faff, R. (2010). Does Simple Pairs Trading Still Work? *Financial Analysts Journal*.

10. Domowitz, I. (2002). Liquidity, Transaction Costs, and Reintermediation in Electronic Markets. *Journal of Financial Services Research*.

11. FINRA. (2024). *2024 Snapshot Data.*

12. Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (2006). Pairs Trading: Performance of a Relative Value Arbitrage Rule. *NATIONAL BUREAU OF ECONOMIC RESEARCH*.

13. Goetzmann, W. N., & Rouwenhorst, K. G. (2005). *The Origins of Value: The Financial Innovations that Created Modern Capital Markets.* Oxford University Press.

14. Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *Review of Financial Studies*.

15. Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does Algorithmic Trading Improve Liquidity? *Journal of Finance*.

16. Holý, V., & Černý, M. (2021, June). Bertram's Pairs Trading Strategy with Bounded Risk. *arxiv*.

17. Hsia, M. (2021, October 12). Exploring distance-based pairs trading strategies using QuantConnect. *Medium*.

18. Huck, N., & Afawubo, K. (2015). Pairs trading and selection methods: is cointegration superior? *Applied Economics* .

19. Jevons, W. S. (1876). *Money and the Mechanism of Exchange.* New York: D. Appleton and Co.

20. Kindleberger, C. P. (1974). "The Great Transformation" by Karl Polanyi. *Daedalus*, 45-52.

21. Kindleberger, C. P. (1993). *A Financial History of Western Europe.* Oxford University Press.

22. Kirilenko, A. A., Kyle, A. S., Samadi, M., & Tuzun, T. (2017). The Flash Crash: High-Frequency Trading in an Electronic Market. *Journal of Finance*.

23. Ma, B., & Ślepaczuk, R. (2022). The profitability of pairs trading strategies on Hong-Kong stock market: distance, cointegration, and correlation methods. *Faculty of Economic Sciences, University of Warsaw*.

24. Macey, J. R., & O'Hara, M. (1999). Golbalization, exchange governance, and the future of exchanges. *Brookings-Wharton papers on financial services*.

25. Maverick, J. (2025). *S&P 500 Average Returns and Historical Performance.*

26. Michie, R. (1999). *The London Stock Exchange: A history.* Oxford University Press.

27. Neal, L. (1990). *The Rise of Financial Capitalism: International capital markets in the Age of Reason.* Cambridge University Press.

28. Qureshi, K., & Zaman, T. (2024, March 12). Pairs Trading Using a Novel Graphical Matching Approach. *arxiv*.

29. Rad, H., Low, R. K., & Faff, R. W. (2016). The Profitability of Pairs Trading Strategies: Distance, Cointegration, and Copula Methods. *Quantitative Finance*.

30. Roover, R. D. (1948). *Money, banking and credit in mediaeval Bruges: Italian merchant-bankers Lombards and money changers, a study in the origins of banking.* Medieval Academy of America.

31. Sarmento, S. M., & Horta, N. (2020). Enhancing a Pairs Trading strategy with the application of Machine Learning. *Expert Systems with Applications*.

32. Sen, J. (2022). Designing Efficient Pair-Trading Strategies Using Cointegration for the Indian Stock Market. *Quantitative Finance*.

33. Sylla, R. (1998). *U.S. Securities Markets and the Banking System, 1790-1840.* Federal Reserve Bank of St. Louis Review.

34. Tadi, M., & Witzany, J. (2023). Copula-Based Trading of Cointegrated Cryptocurrency Pairs.

35. Tapscott, D., & Tapscott, A. (2016). *Blockchain Revolution: How the Technology behind Bitcoin Is Changing Money, Business, and the World.* Penguin.

36. Vidyamurthy, G. (2004). *Pairs Trading: Quantitative Methods and Analysis*. New York: John Wiley & Sons.