Pairs Trading Strategy Exploration

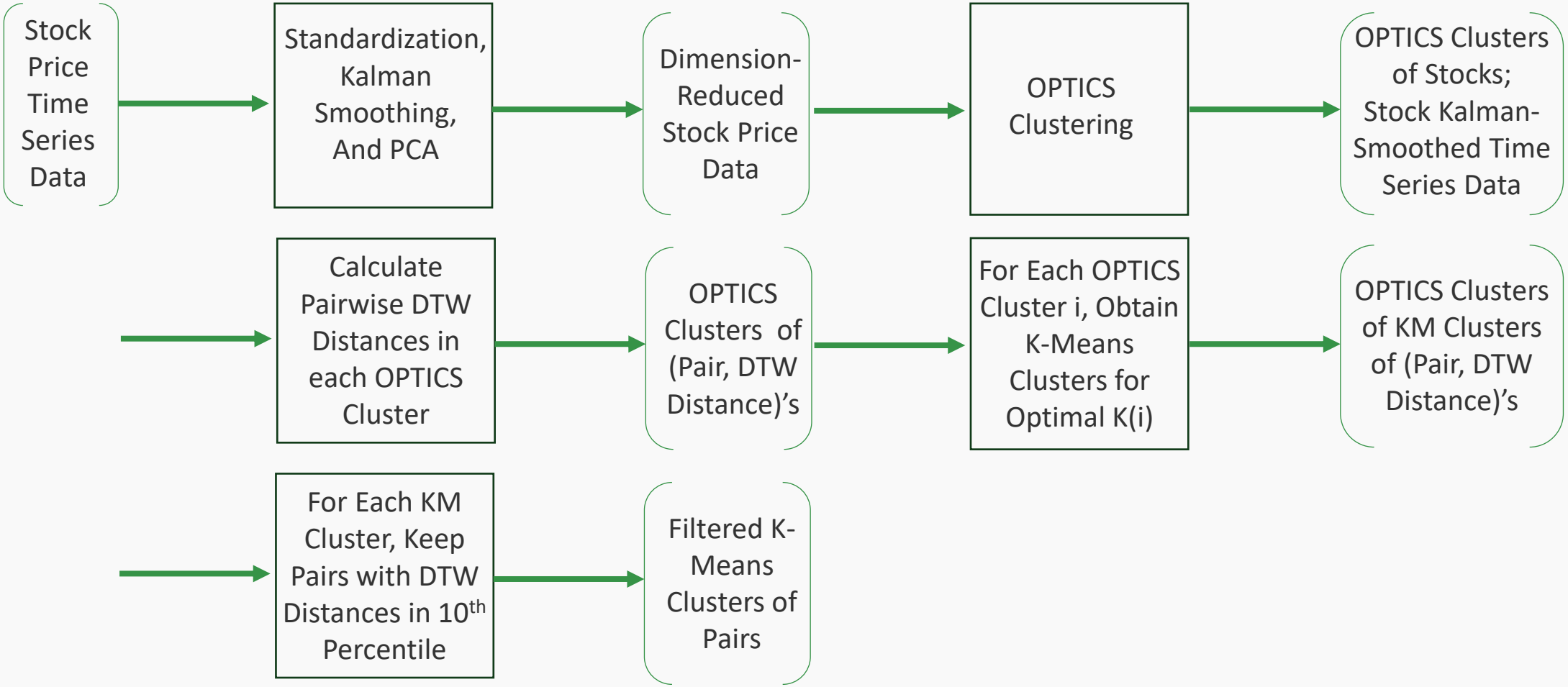
Harvey Hoang, Arthur Faynin, William Villacis

FQE
Financial Quants and Engineers Club

Optics DTW K-Means Filter

Spring 2024 Pairs Team & William Villacis

Pair Finding Pipeline



Back-test Formulation

- Time Horizon – Oct 2004 – Sep 2024:
  - Stock Universe For Year yyyy: S&P 500 Constituents on 9-30-yyyy
  - Pair-Finding Lookback Horizon: 2 years
  - Trading Time Horizon: 1 year

- Trading Rules:
  - Signal: Standardized Spread $P_1 - \beta P_2$
    - 63-day lag for beta and standardization calculations
    - Enter Trade if abs(z-score) >= 2
    - Exit Trade if abs(z-score) <= 1
  - Max Time in Position: 63 days

Back-test Results (1 of 3): Total Results

FQE
Financial Quants and Engineers Club

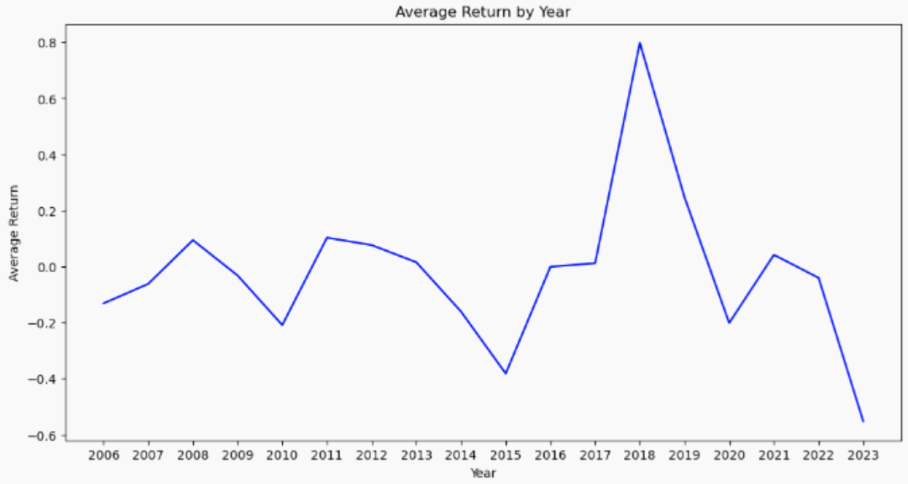
Metric	Filtered_2_1	Filtered_3_2	Total_2_1	Total_3_2
Avg Return	2.03%	0.95%	-8.34%	-13.55%
Sortino	0.0090	0.0075	-0.0073	-0.0138
Win Rate	0.5023	0.4826	0.4949	0.4859
Profit Factor	1.1025	1.0622	0.7353	0.5074
Avg Trade Duration	33.52 days	16.76 days	32.95 days	16.57 days
Pairs Found	605	605	3981	3981
Pairs Traded	604	604	3967	3967

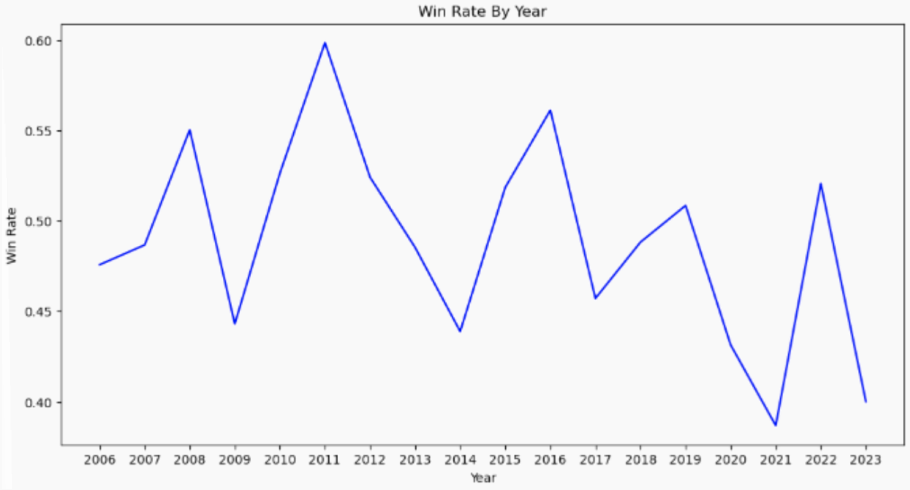
$$\text{Sortino} = \frac{avg\ return}{std\ dev\ negative\ returns}$$

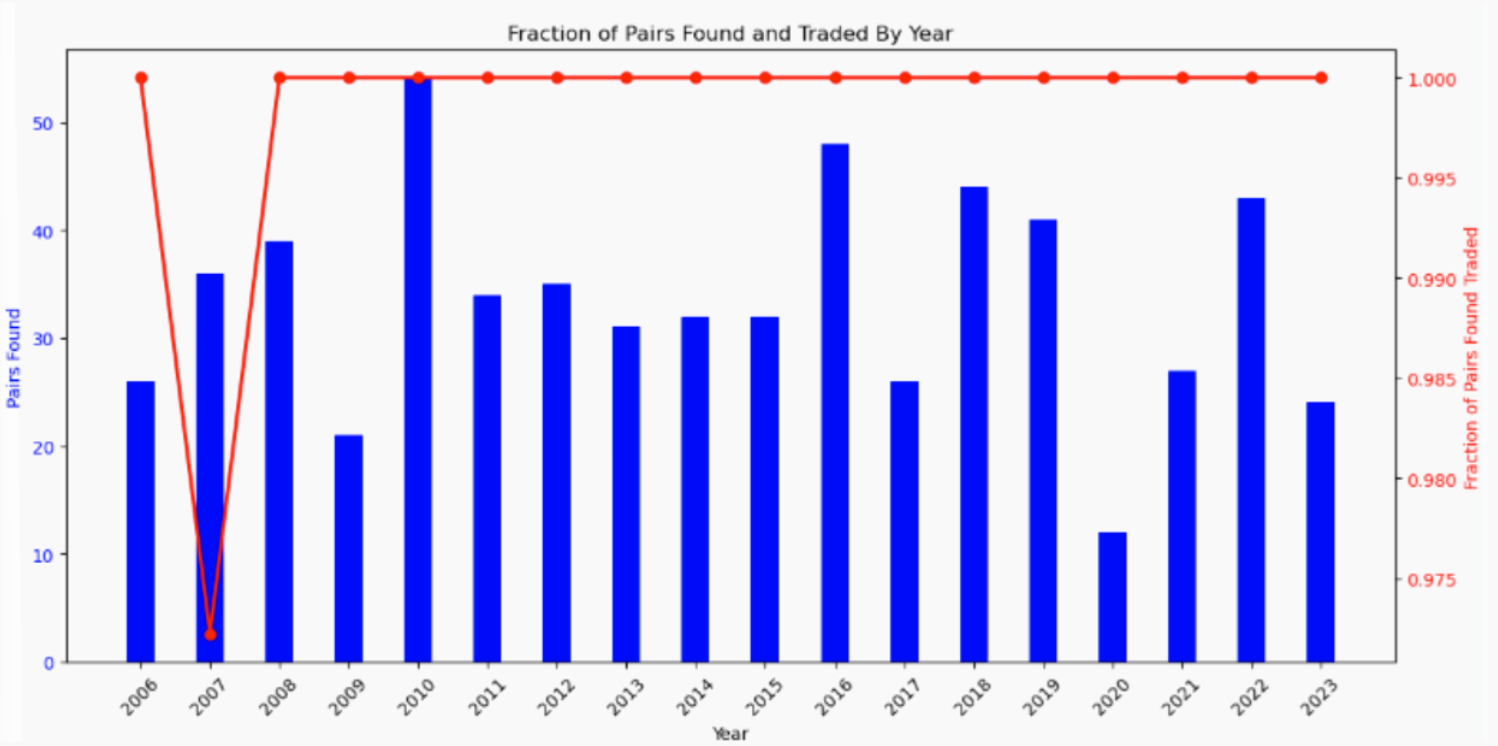
$$\text{Win Rate} = \frac{\#\ postive\ return\ trades}{\#\ trades}$$

$$\text{Profit Factor} = \frac{sum\ of\ positive\ returns}{abs(sum\ of\ negative\ returns)}$$

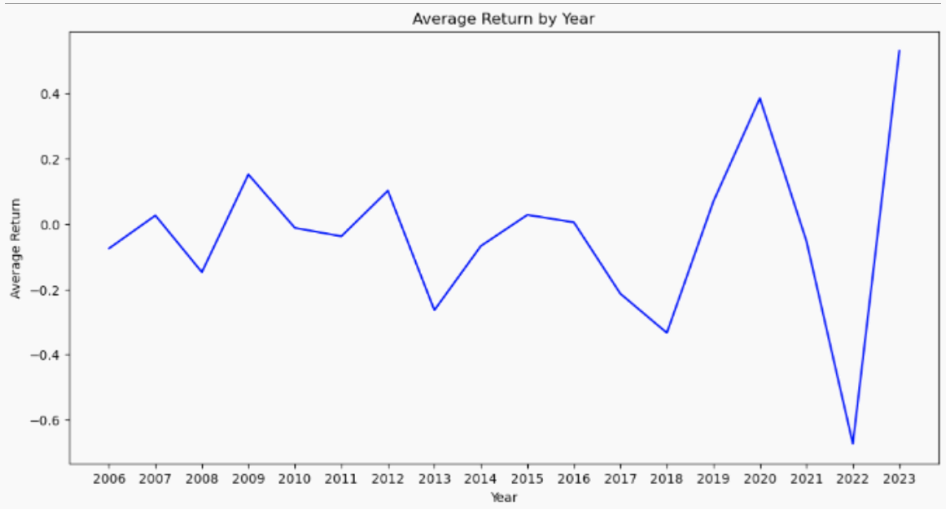
Back-test Results (2 of 3): Filtered 2-1

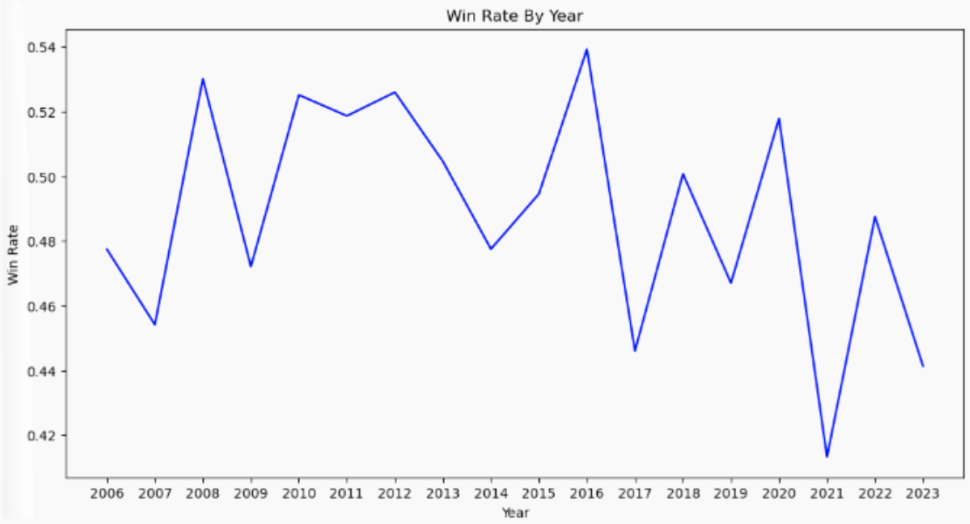


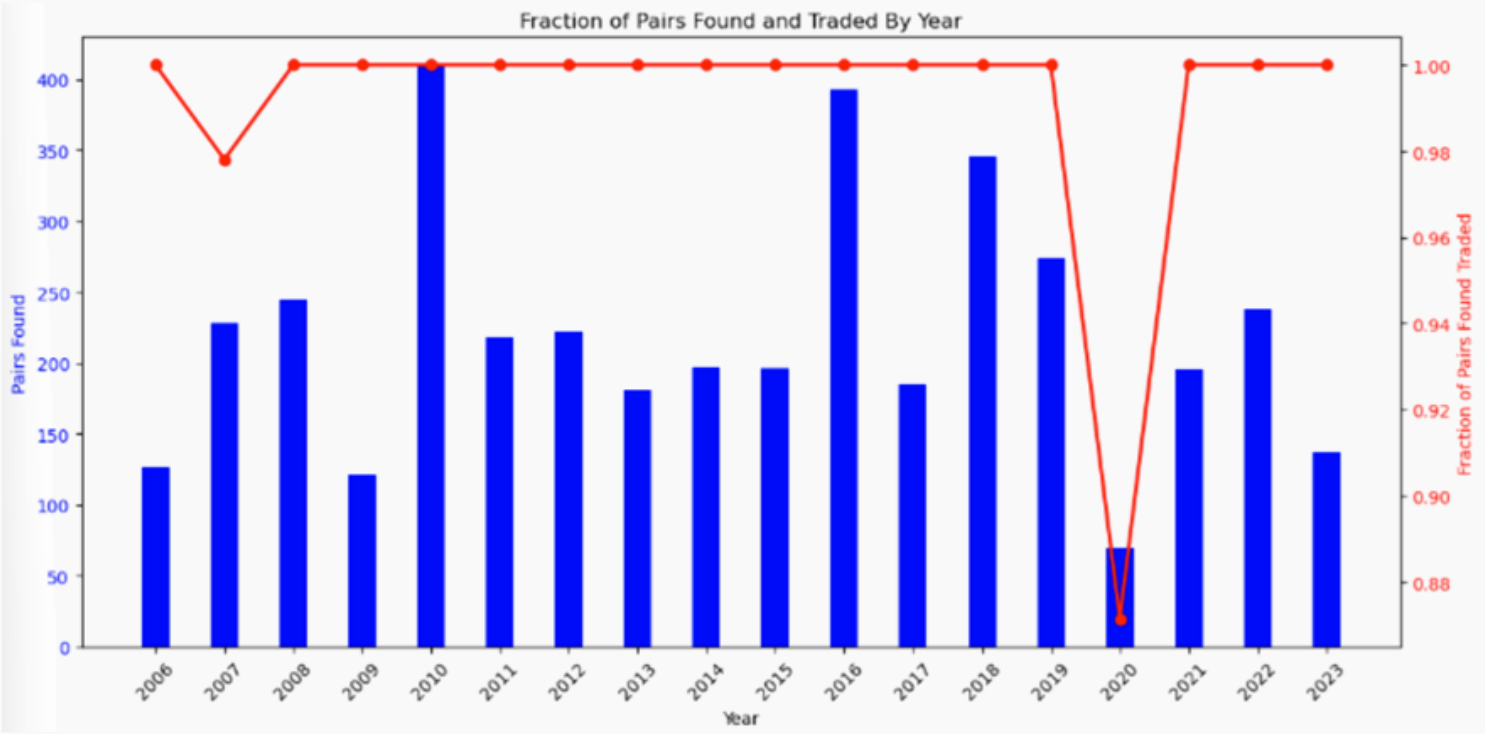




Back-test Results (3 of 3): Total 2-1







Considerations & Further Research

FQE
Financial Quants and Engineers Club

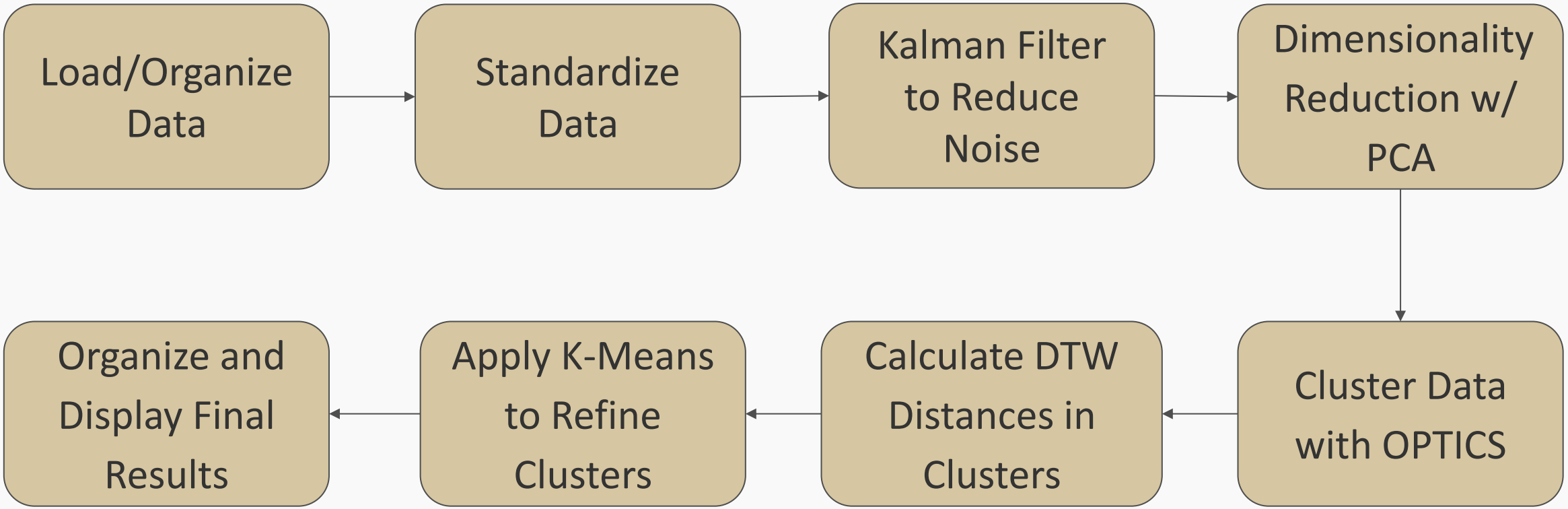
- Use signal related to pair-finding metric:
  - Ex: rolling DTW Distance of pair instead of z-score of spread
- Adjust look-back horizons:
  - standardization, pair finding
- Include more stock features into pair-finding filter:
  - Ex: firm-level features
- Improve return and portfolio metric calculations

FQE
Financial Quants and Engineers Club

Enhanced Optics DTW K-Means Filter

Arthur Faynin

Overview


Load/Organize Data
Standardize Data
Kalman Filter to Reduce Noise
Dimensionality Reduction w/ PCA
Cluster Data with OPTICS
Calculate DTW Distances in Clusters
Apply K-Means to Refine Clusters
Organize and Display Final Results

Add Firm Data

FQE
Financial Quants and Engineers Club

```
firm_data = pd.read_csv("stock_performance/firm_level.csv")
firm_data.rename(columns={'permno': 'PERMNO'}, inplace=True)
firm_data['public_date'] = pd.to_datetime(firm_data['public_date'])

firm_data['divyield'] = firm_data['divyield'].str.rstrip('%').astype('float') / 100.0
numeric_cols = ['pe_exi', 'dpr', 'roe', 'GProf', 'divyield']
```

Trading Strategies

FQE
Financial Quants and Engineers Club

Kalman Filter State Transition Equation:

$$\mathbf{x}k = \mathbf{F}\mathbf{x}k - 1 + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k$$

xk = State vector at time k .

F = State transition matrix.

B = Control input matrix.

Uk = Control vector.

Wk = Process noise (assumed to be Gaussian).

Silhouette Score:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

a(i) : Mean intra-cluster distance (average distance between i and all other points in the same cluster).

b(i) : Mean nearest-cluster distance (average distance between i and all points in the nearest neighboring cluster).

Backtesting Ideas

FQE
Financial Quants and Engineers Club

Rolling Mean/STD:

```
# Calculate rolling mean and std of the spread
mean_spread = spread.rolling(window=30).mean()
std_spread = spread.rolling(window=30).std()
z_score = (spread - mean_spread) / std_spread
```

Limit Date:

```
pivot_data = pivot_data.loc['2023-01-01':'2023-12-29']
```

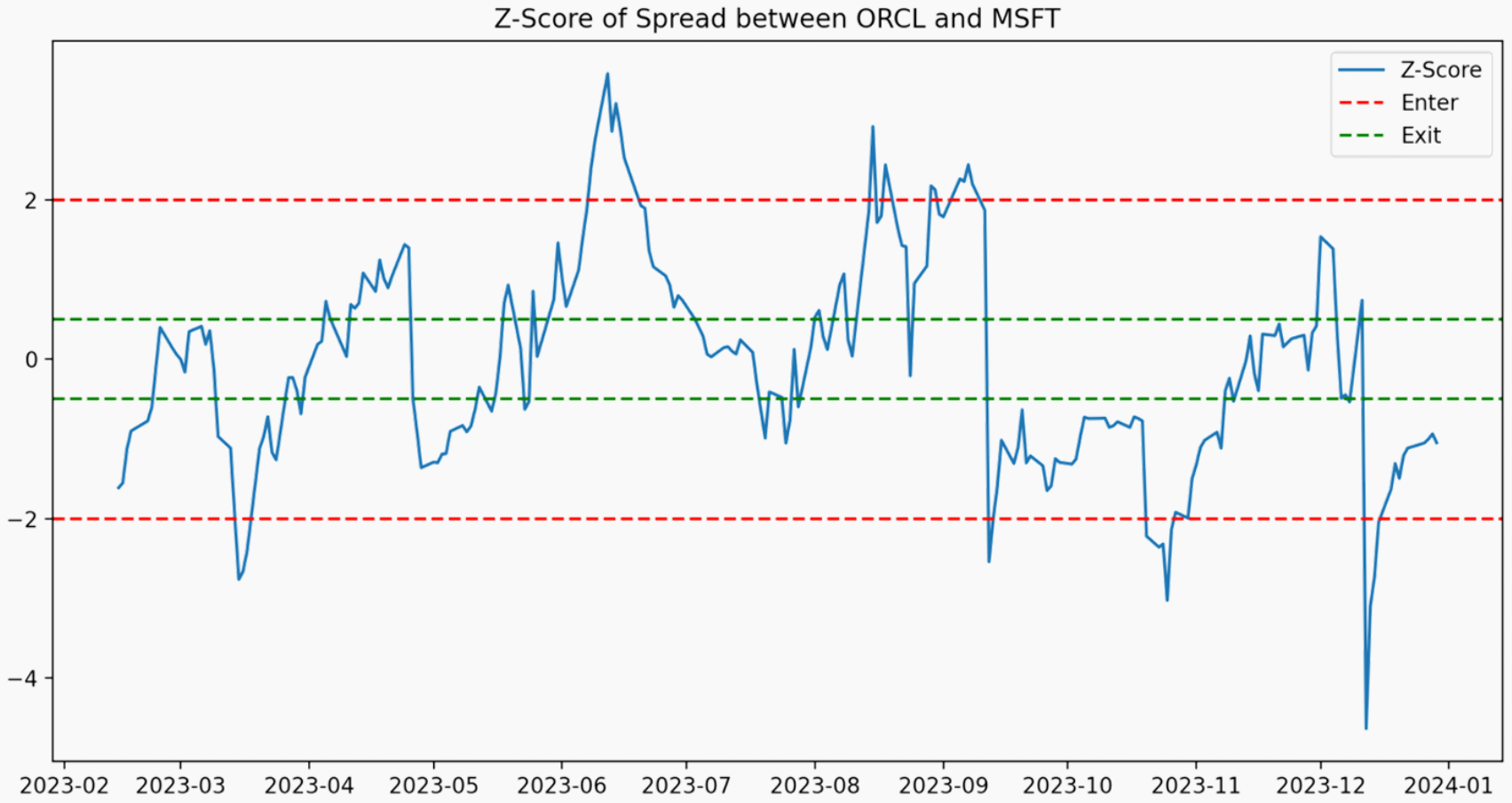
Exit Signal when both stocks are within .5 STDs Z-score

```
# Exit signal
elif position != 0 and abs(z_score.iloc[i]) < z_exit:
    exit_price = (price1, price2)
    pnl_stock1 = (exit_price[0] - entry_price[0]) * entry_shares_stock1
    pnl_stock2 = (exit_price[1] - entry_price[1]) * entry_shares_stock2
    cost_stock1 = abs(entry_shares_stock1) * (entry_price[0] + exit_price[0]) * transaction_cost
    cost_stock2 = abs(entry_shares_stock2) * (entry_price[1] + exit_price[1]) * transaction_cost
    total_cost = cost_stock1 + cost_stock2
    pnl = pnl_stock1 + pnl_stock2 - total_cost
    portfolio.append(pnl)
```

Entry Signal when both stocks exceed two STDs Z-score

```
# Entry signal
if position == 0:
    if z_score.iloc[i] > z_entry:
        # Short spread: Sell stock1, buy stock2 * beta
        position = -1
        entry_price = (price1, price2)
        trade_capital = capital * trade_size_fraction
        shares_stock1 = trade_capital / (price1 + beta * price2)
        shares_stock2 = beta * shares_stock1
        entry_shares_stock1 = -shares_stock1
        entry_shares_stock2 = shares_stock2
        trades.append({
            'date': date,
            'position': position,
            'entry_price': entry_price,
            'entry_shares_stock1': entry_shares_stock1,
            'entry_shares_stock2': entry_shares_stock2
        })
```

One of the closest pairs with Firm level Data



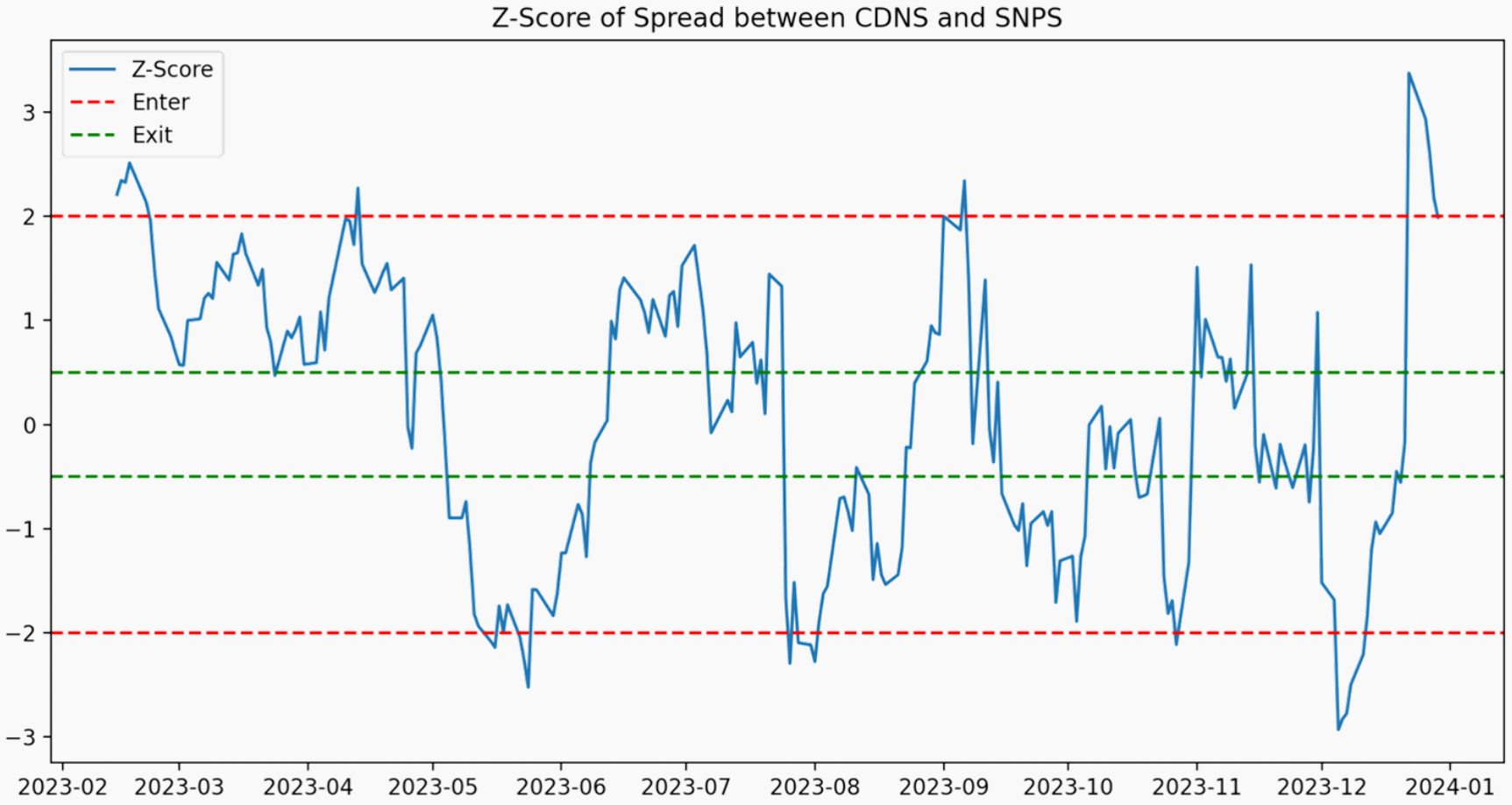
DTW Distance between ORCL and MSFT: 1.81194429983846

```
Backtesting pair: ORCL and MSFT
Annualized Return: 1.28%
```

Severely Underperforms compared to S and P

```
Backtesting pair: MSFT and MPWR
Annualized Return: -0.81%

Backtesting pair: MSFT and AVGO
Annualized Return: 1.46%

Backtesting pair: CDW and AMAT
Annualized Return: 0.62%

Backtesting pair: CDW and KLAC
Annualized Return: -0.34%

Backtesting pair: CDW and LRCX
Annualized Return: 2.36%

Backtesting pair: CDW and TER
Annualized Return: 0.64%
```

One of the closest pairs with just stocks



DTW Distance between ORCL and MSFT: 1.81194429983846

```
Backtesting pair: CDNS and SNPS
Annualized Return: 0.36%
```

Severely Underperforms compared to S and P

FQE
Financial Quants and Engineers Club

FX Two-Stage Correlation & Cointegration

Harvey Hoang

FX Market Microstructure

Market Dynamics, Technical Terms, Currency Quote, and Market Details

- Pip Value: The difference in fourth decimal place in most currency pairs (1 pip)
  - For example: 1.2345 -> 1.2346
- Currency Quote: The price of one currency in terms of another currency – Base/Quote
  - For example: USD/EUR = 1.2000
- Some exchanges might have a different price quote, but the trading logic is the same thing
  - Long USD/EUR is equivalent to short EUR/USD
- Independent Order Book between currency pairs
  - Long USD/EUR then short USD/GBP does not offset your position
- Brokers control price quote and order filling/ Risks of unfilled order/ Low Spread
  - Long USD/EUR in Broker A for 1.2000 -> Short USD/EUR in Broker B for 1.2010

This sounds like easy money but there are way more beyond that!

Trading Strategies: Cointegration Arbitrage

George J. Miao – "Dynamics Pair Trading and Two-Stage Correlation"

- Cointegration vs Correlation: Cointegration tests whether two stocks price move along together in the long run, despite the short-term deviation, while correlation measure their relationship in the short term
- Correlation Formula:

$$\rho = \frac{\sum_i^N (A_i - \bar{A})(B_i - \bar{B})}{\left[\sum_i^N (A_i - \bar{A})^2 \sum_i^N (B_i - \bar{B})^2\right]^{1/2}} \quad \text{for} \quad \bar{A} = \frac{1}{N}\sum_i^N A_i \quad \& \quad \bar{B} = \frac{1}{N}\sum_i^N B_i$$

- Cointegration Approach and Cointegration Verification:

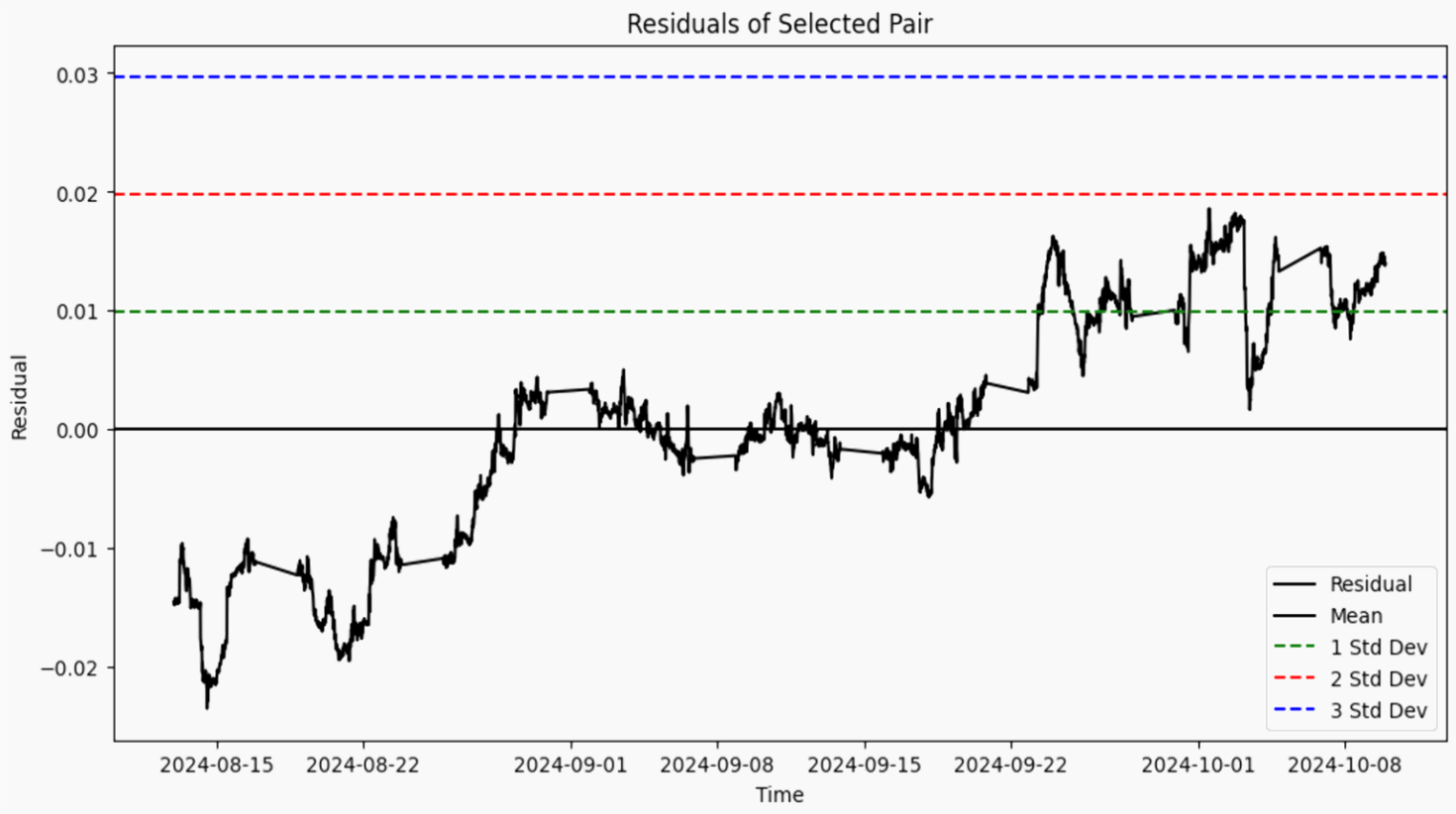
$$P_t^A - \gamma P_t^B = \mu + \varepsilon_t, \quad \& \quad \Delta Z_t = \alpha + \beta t + \gamma Z_{t-1} + \sum_{i=1}^{p-1} \delta_i \Delta Z_{t-i} + u_t$$

- Statistical test for P value (lag-order):

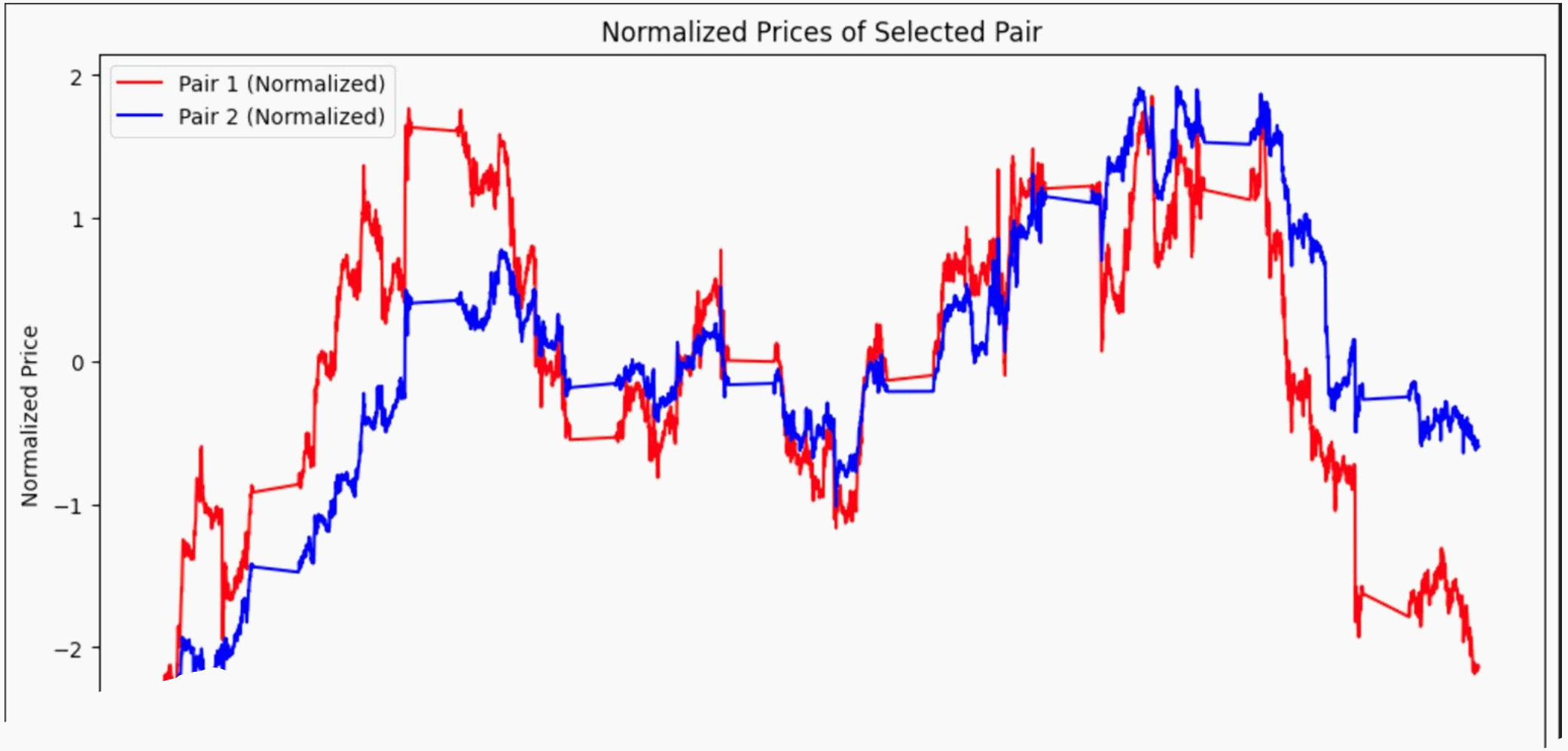
$$AIC = \ln\left(\hat{\sigma}_p^2\right) + \frac{2p}{T} \qquad SIC = \ln\left(\hat{\sigma}_p^2\right) + \frac{p\ln(T)}{T} \qquad HQC = \ln\left(\hat{\sigma}_p^2\right) + \frac{2p\ln[\ln(T)]}{T}$$

Back-testing Ideas and Techniques

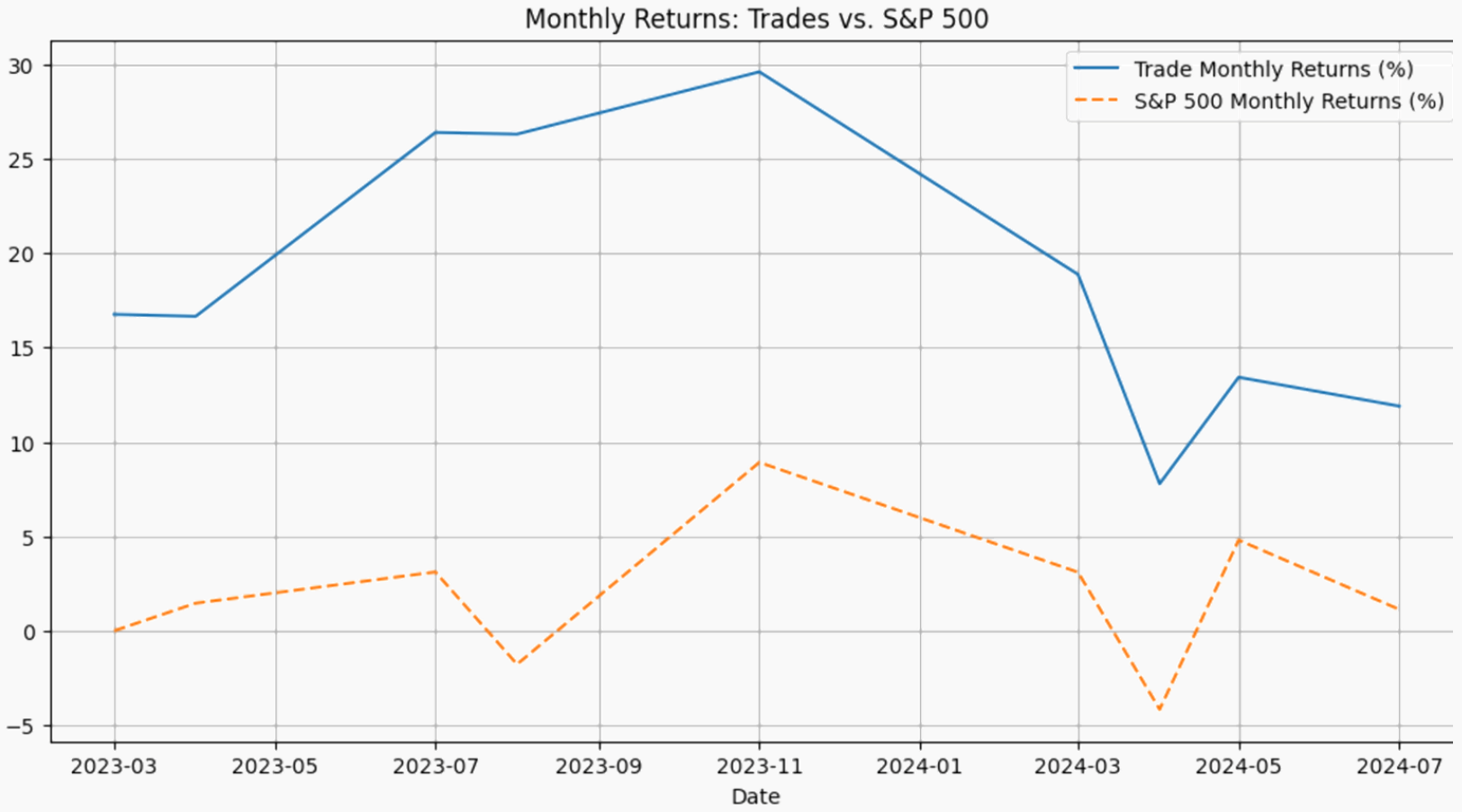
- 15 minutes bar candle with close price on EUR/USD and GBP/USD over the past 4 years
- Rolling window for 66 days and Testing window for 30 days
- Entering orders whenever Z_i= +/- 2 standard deviations from the mean
- Stop loss at 0.5 standard deviations



Cointegration Visualization



Cointegration Visualization



Summary of Trades

FQE
Financial Quants and Engineers Club

Pairs Trading Portfolio (All Trading Results)	Value	Pairs Trading Portfolio (Monthly Return Results)	Value
Number of Winning Trades	12	Mean	13.97
Number of Loss Trades	0	Median	12.89
Total Number of Trades	12	Minimum	4.29
Winning/Loss Ratio	N/A	Maximum	29.59
Percentage of Winning Trades	100.0	Standard Deviation	7.64
Percentage of Loss Trades	0.0	Skewness	0.98
Average Percentage of Winning Trades	13.97	Kurtosis	0.49
Average Percentage of Loss Trades	nan	Monthly Sharpe Ratio	1.83
Reward/Risk Ratio	N/A	Annual Sharpe Ratio	6.34
Profit (Annually, $)	16765.5	Profit (Monthly, $)	1397.1
Profit (Annually, % of Capital)	167.655	Profit (Monthly, % of Capital)	13.971

Q&A