



QUANTITATIVE ANALYSIS - PAIRS TRADING STRATEGY

Presented by: Addysen Stewart, Molly Xie,
Zhengyu Zhu, Linxi Fan

Mentor: James Dufresne

WIM DRP Research Program

OUR RESEARCH OBJECTIVE:

Develop & Evaluate a Market-Neutral Pairs Trading Strategy

Using historical stock market data

PROFITABILITY

Returns?

ROBUSTNESS

Different regimes?

RISK

Risk-adjusted?



OVERVIEW OF OUR PIPELINE:

1. Data Collection & Pair Selection
 - Identify a suitable pair of stocks that exhibit a stable historical relationship.
 2. Z-Score Analysis
 - Detect when the spread deviates significantly from its historical mean.
 3. Backtesting
 - Evaluate whether the trading strategy would have been profitable using historical data.
 4. Benchmark Comparison
 - Compare the strategy against passive investment alternatives.
 6. Risk Evaluation
 - Measure the strategy's return and risk characteristics.
 7. Market Regime Analysis
 - Evaluate whether the strategy remains effective under different market environments.
-

1. DATA COLLECTION & PAIR SELECTION

- Pair we selected:

```
# KO-PEP    Coca-Cola vs PepsiCo (beverages)
# MA-V      Mastercard vs Visa (payments)
# XOM-CVX    ExxonMobil vs Chevron (energy/oil)
# TD-RY      TD Bank vs Royal Bank of Canada (Canadian banking)
# SPY-QQQ    S&P 500 ETF vs Nasdaq-100 ETF (broad market indices)
```

If $p < 0.05$:
cointegrated
Else:
False

- P-value & Cointegration:

Run: Engle–Granger Cointegration Test

How do we test for cointegration?

$$P_{1,t} = \alpha + \beta P_{2,t} + \varepsilon_t$$

- Is ε_t stationary?

	pair	p_value	is_cointegrated
0	KO-PEP	0.9939	False
1	MA-V	0.0348	True
2	XOM-CVX	0.0708	False
3	TD-RY	0.6530	False
4	SPY-QQQ	0.1678	False

2. Z-SCORE COMPUTATION

How many standard deviations away is today's spread from its historical mean?

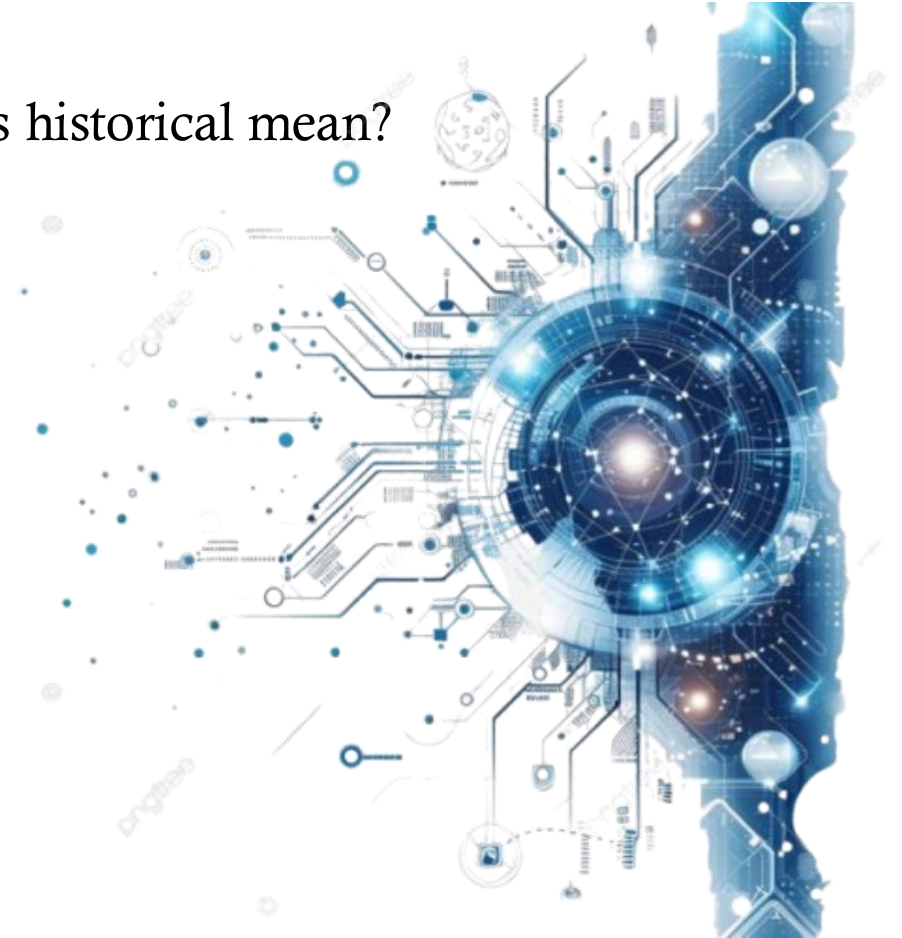
- z score answers the above question.

Formula:

1. Regression: $P_1 = \alpha + \beta P_2 + \varepsilon$

2. Spread: $\varepsilon_t = P_{1,t} - (\alpha + \beta P_{2,t})$

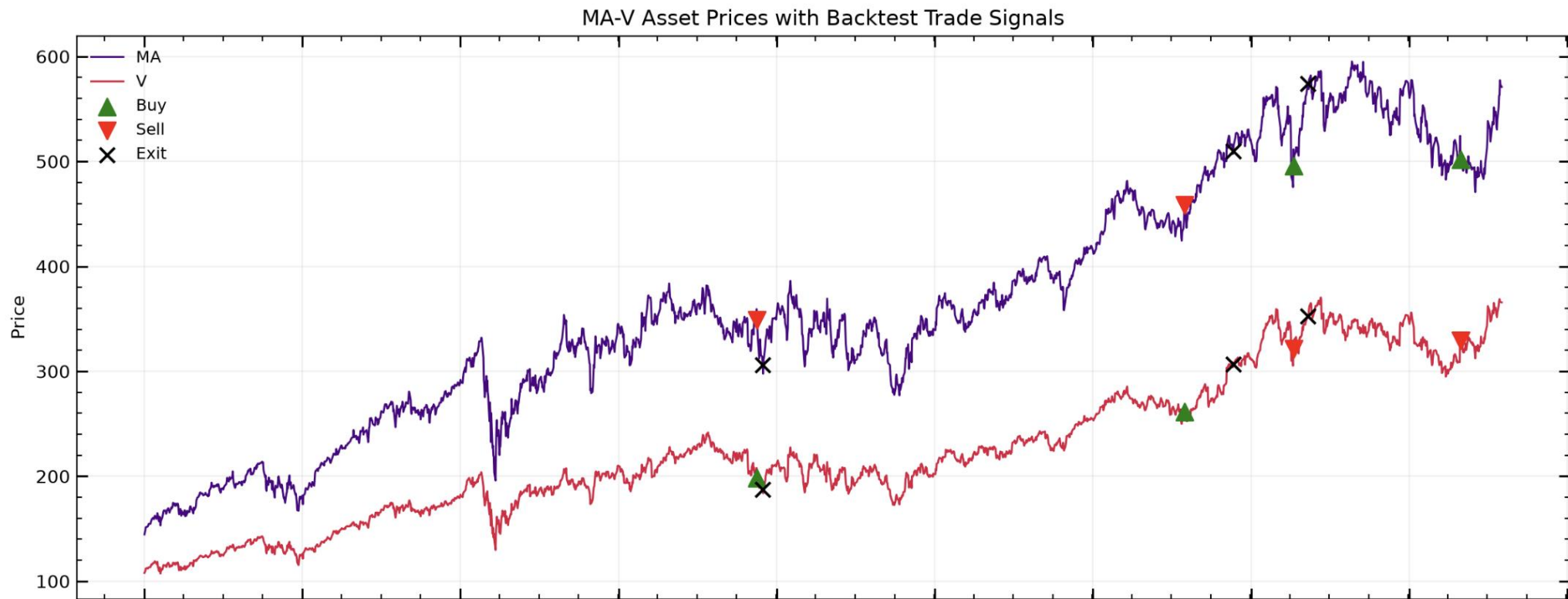
3. Z- score: $z_t = \frac{\varepsilon_t - \mu_\varepsilon}{\sigma_\varepsilon}$



Z-SCORE SIGNALS

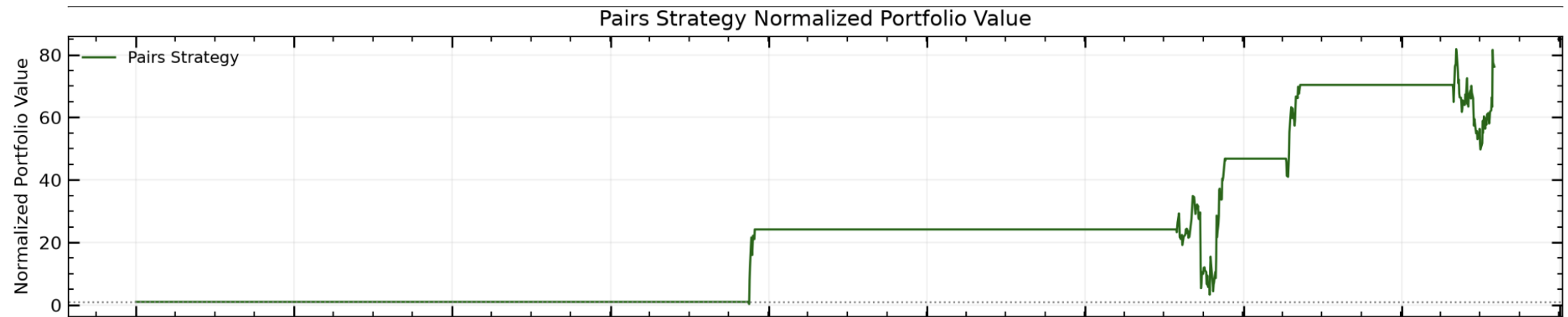


PLOT OF PAIR PRICE OVER TIME (WITH TRADE SIGNALS)

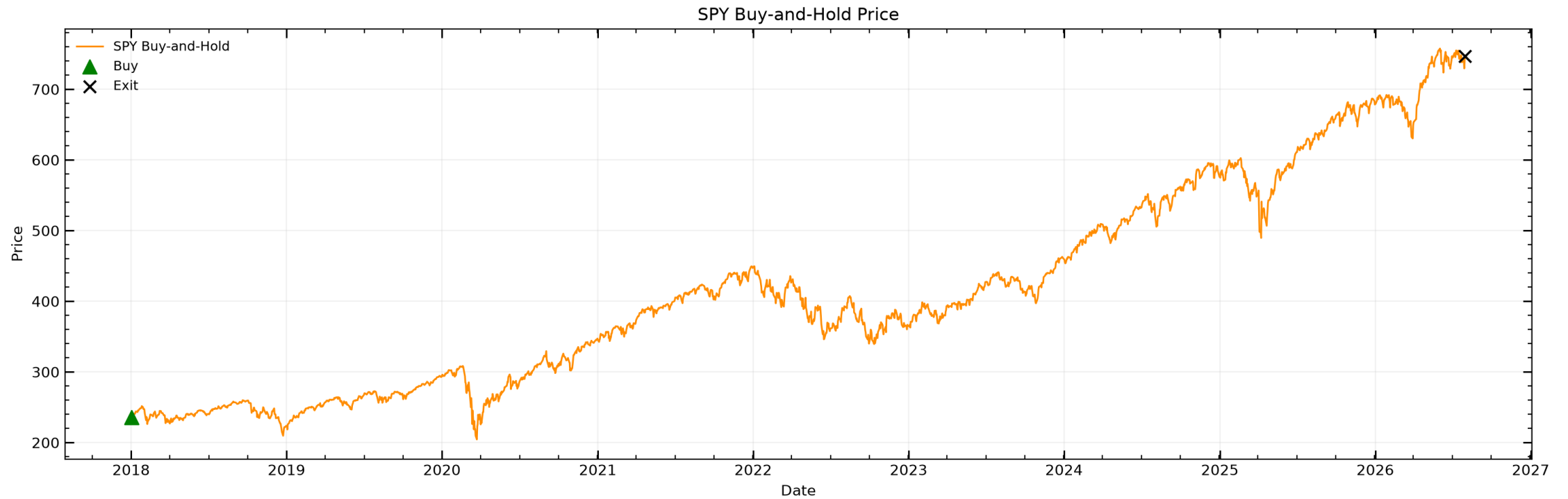


STRATEGY VALUE

- $\text{portfolio_value_t} = \text{initial_capital} + \text{cumulative_PnL_t}$



BUY-AND-HOLD BENCHMARK PRICE



COMPUTING RISK METRICS

$$\text{Annualized Volatility} = \sigma \cdot \sqrt{252}$$

$$\text{Sharpe} = \frac{R_p - R_f}{\sigma_e} \cdot \sqrt{252}$$

$$\text{Sortino} = \frac{R_p - R_f}{\sigma_d} \cdot \sqrt{252}$$

Where:

- σ is the standard deviation of the price
 - R_p is the return of the investment
 - R_f is the risk-free rate of the investment
 - σ_e is the standard deviation of the returns excess the risk free rate
 - σ_d is the standard deviation of the negative returns excess the risk free rate
-

COMPUTING RISK METRICS

$$\text{MDD} = \frac{\text{trough value} - \text{peak value}}{\text{peak value}}$$

$$\text{VaR} = v_m \cdot \left(\frac{v_i}{v_{i-1}} \right)$$

$$\text{CVaR} = \frac{1}{1-c} \int_{-1}^{\text{VaR}} x \cdot p(x) dx$$

Where:

- v_m is the current return (today)
 - v_i is the historical value on day i
 - v_{i-1} is the historical value on day $i-1$
 - i is from 1 to M , where M is the total number of historical days given
 - c is the cut off point in the distribution for the VaR
 - $p(x)$ is the probability density of having a return value = x
-

RISK METRICS COMPARED TO BUY AND HOLD

Strategy	Total Return	Volatility	Sharpe	Sortino	MDD	Daily 95% VaR	Daily 95% CVaR	Worst Daily Return
SPY Buy & Hold	217 %	19 %	.63	.89	-34 %	-1.8 %	-2.9 %	-11 %
MA & V Pair	7654 %	856 %	.41	6.77	-90 %	0 %	-7.3 %	-69 %

MARKET REGIME DEFINITIONS

Market Regime	Definition
High Volatility	20-day rolling volatility of SPY $\geq$ sample median
Low Volatility	20-day rolling volatility of SPY $<$ sample median
Bull Market	SPY closing price $\geq$ 200-day moving average (MA200)
Bear Market	SPY closing price $<$ 200-day moving average (MA200)

MARKET REGIME ANALYSIS

OVERALL PERFORMANCE (ALL TRADING DAYS)

Market Condition	Mean Daily Return (%)	Median Daily Return (%)	Number of Days	Active Trading Days
High Volatility	0.02%	0.00%	995	70
Low Volatility	0.04%	0.00%	1016	42
Bear Market	0.04%	0.00%	388	21
Bull Market	0.03%	0.00%	1424	91

MARKET REGIME ANALYSIS

PERFORMANCE ON ACTIVE TRADING DAYS ONLY

Market Condition	Mean Daily Return (%)	Median Daily Return (%)	Active Trading Days
High Volatility	0.22%	0.44%	70
Low Volatility	1.02%	-0.02%	42
Bear Market	0.66%	-0.22%	21
Bull Market	0.49%	0.39%	91

SOURCES

- Risk Metrics: <https://communities.sas.com/t5/SAS-Communities-Library/Risk-Simulation-in-SAS-Risk-Engine/ta-p/976747>, <https://www.investopedia.com/terms/t/totalreturn.asp>, <https://www.fool.com/investing/how-to-calculate/annualized-volatility/>, <https://www.investopedia.com/terms/s/sharperatio.asp>, <https://www.investopedia.com/terms/s/sortinoratio.asp>, <https://www.investopedia.com/terms/m/maximum-drawdown-mdd.asp>, <https://www.investopedia.com/articles/04/092904.asp>, https://www.investopedia.com/terms/c/conditional_value_at_risk.asp
-

Thank you
for listening!

