PAIRS TRADING AND COINTEGRATION STRATEGY

Neha Sellappan Devadas (230123040)

Pairs trading is a market-neutral trading strategy that involves simultaneously buying one asset (or going long) and selling another related asset (or going short). The objective of pairs trading is to profit from the relative performance of these two assets while minimizing exposure to overall market movements. It's based on the idea that, over time, certain related assets will move in tandem, and when they diverge from their historical relationship, there may be opportunities for profit.

PAIR IDENTIFICATION AND SELECTION

We have used 12 top performing stocks from the energy industry for this project, and the stock data to find the pair used is from April 2019 to March 2022, which is then back tested on a period from April 2022 to March 2025

After collecting stocks closing prices, we have performed the following test to find a pair:

Cointegration Test

- The primary purpose of a cointegration test is to determine whether there exists a stable, long-term relationship between two or more non-stationary time series.
- Cointegrated time series move together in the long run, even if they might exhibit short-term fluctuations and divergences.

Two-step Engle & Granger test for cointegration

Step 1: Estimate the Long-Run Relationship

The idea is to regress one non-stationary variable on another:

$$y_t = c + \gamma x_t + z_t$$

Here:

- y_t and x_t are both assumed to be non-stationary, typically I(1) (integrated of order 1),
- c is an intercept (constant term),
- γ captures the long-run relationship (like slope in regression),
- z_t is the residual, also called the error correction term.

The code snippet performs an OLS regression of y_t on x_t, including a constant.

Step 2: Test for stationarity

You test whether the residuals zt are stationary, i.e., whether:

$$z_t = \phi z_{t-1} + \eta_t$$

with $|\phi| < 1$ implying stationarity.

This is what the ADF test checks:

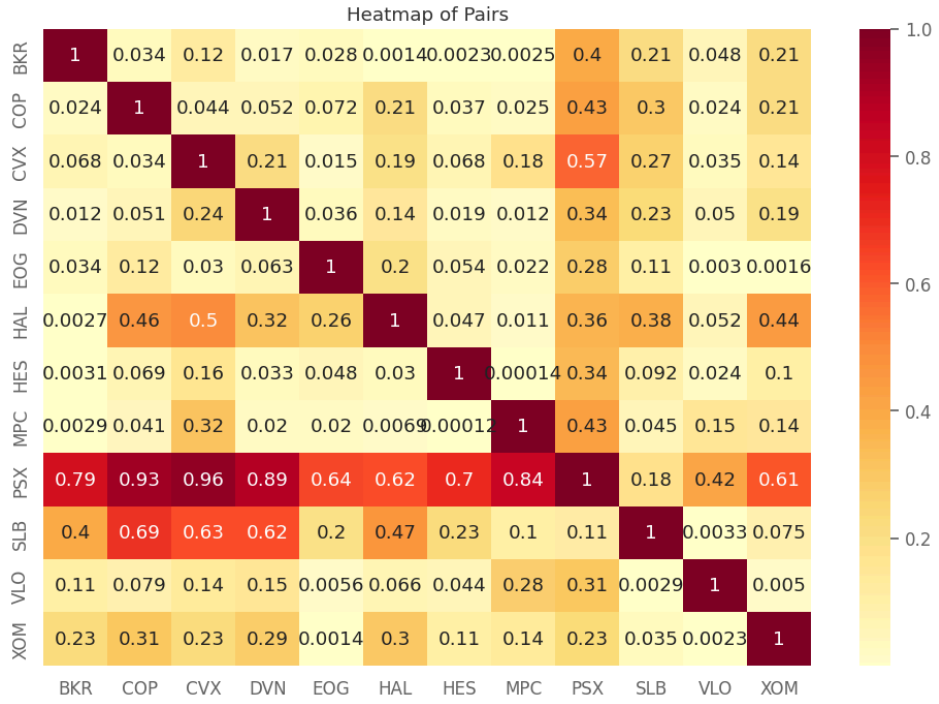
- The null hypothesis is: $\phi = 1$ (i.e., unit root — non-stationary).

- ❖ The alternative is: $\phi < 1$ (i.e., stationary).
- ❖ So if `p-value < 0.05` (typically), you reject the null ⇒ zt is stationary ⇒ y and x are cointegrated.

We perform the above test to all pairs whose data we extracted.

P Values

p value can be considered as a measure of strength of cointegration between the two series.Lower p value indicates a strong cointegration.



Based on p value, we select the pair with the least p value and in our case the pair turns out to be "HEC" and "MPC".

Log prices of both the stocks (over the period for choosing stocks)

STRATEGY DEVELOPMENT

The trading strategy we formed will be back tested in the period from April 2022 to March 2025 . We implement our trading strategy in following steps:

1.Calculating Spread:

Spread calculates the difference between the stock prices of selected pairs.

$$S_t = Y_t - \beta * X_t$$

Where,
S_t = Spread
Y_t= Log of daily closing price of second stock
X_t = Log of daily closing price of first stock
β = Hedge Ratio from previous period

This spread represents the deviation between the two stock prices, adjusted for scale.

2. Calculating z-score:

After spread is calculated, z-score of spread is calculated with a rolling window with a lookback period of 20 days.

$$z = (X - \mu)/\sigma$$

Where,
z = z-score
X = Spread of past 20 days
μ = mean of X
σ = standard deviation of X

It measures how far the spread is from its average in terms of standard deviations.
A high or low z-score implies the spread has diverged significantly from the mean.

3. Entry and Exit Positions:

We could perform a profitable trade when z-score deviates from its mean value.

Enter positions:

- **If z-score > threshold**:
  - The spread is unusually high.
  - → **Sell S1, Buy S2** (expecting spread to decrease).
  - Signal = `+1`
- **If z-score < -threshold**:
  - The spread is unusually low.
  - → **Buy S1, Sell S2** (expecting spread to increase).
  - Signal = `-1`

No trade / hold position:

- If z-score is within the ±threshold range, the signal is `0`

3. Risk Management:

- **Stop-loss condition**:

- ➢ If the z-score exceeds the threshold by an additional margin (`threshold + stop_loss`):
  - The trade is considered too risky or has moved against expectations.
  - → **Signal is set to 0** to exit the trade.

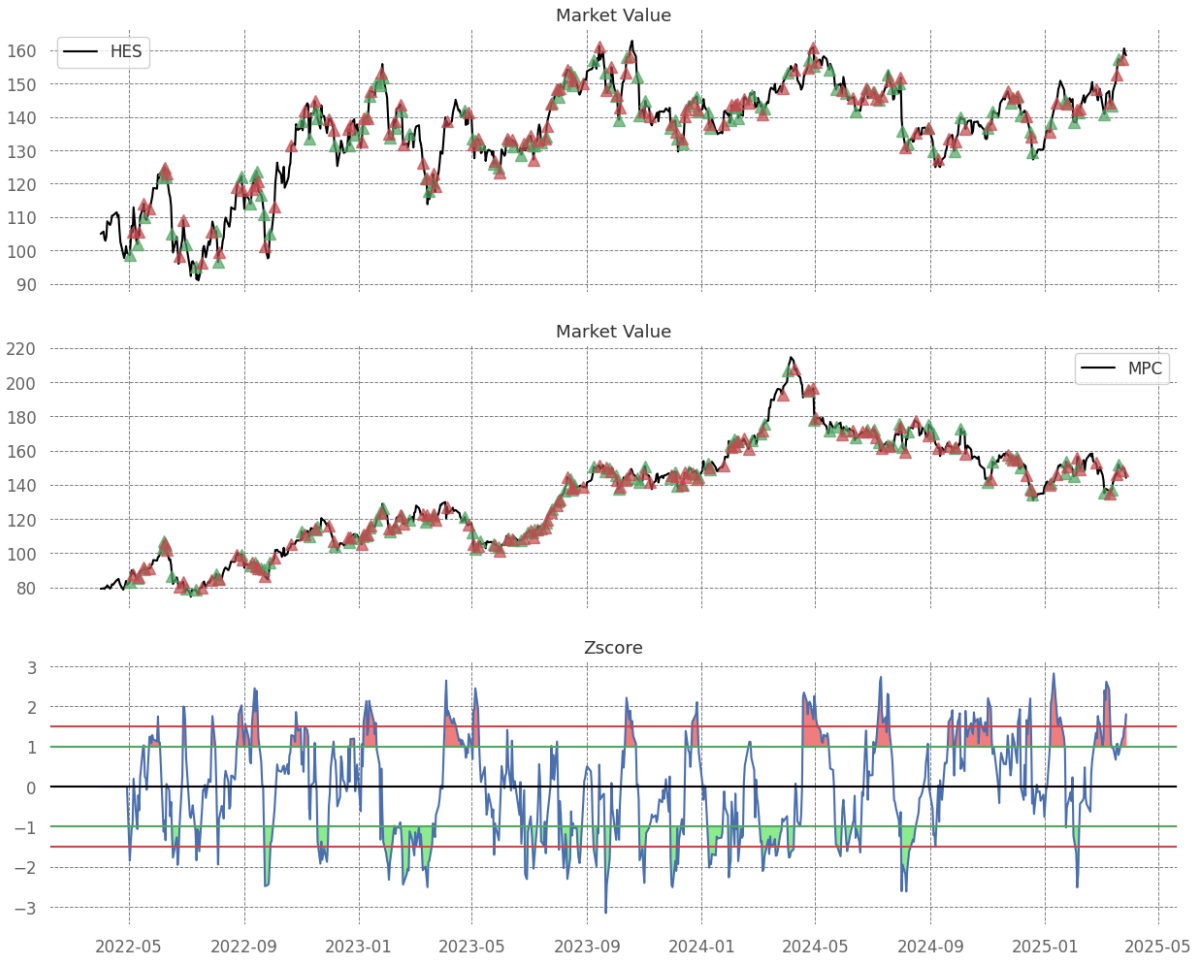
❖ **Take-profit / mean-reversion exit**:

- ➢ If the z-score returns to `0` (i.e., spread has normalized to the mean):
  - → Exit the trade by setting the signal to `0`

3. Position Generation:

Calculate position changes:

- ❖ `position1 = -signal.diff()` → Captures when to enter or exit trades for stock S1.
- ❖ `position2 = -position1` → Takes the opposite side for S2.
- ❖ This ensures the trades are always paired (one long, one short).

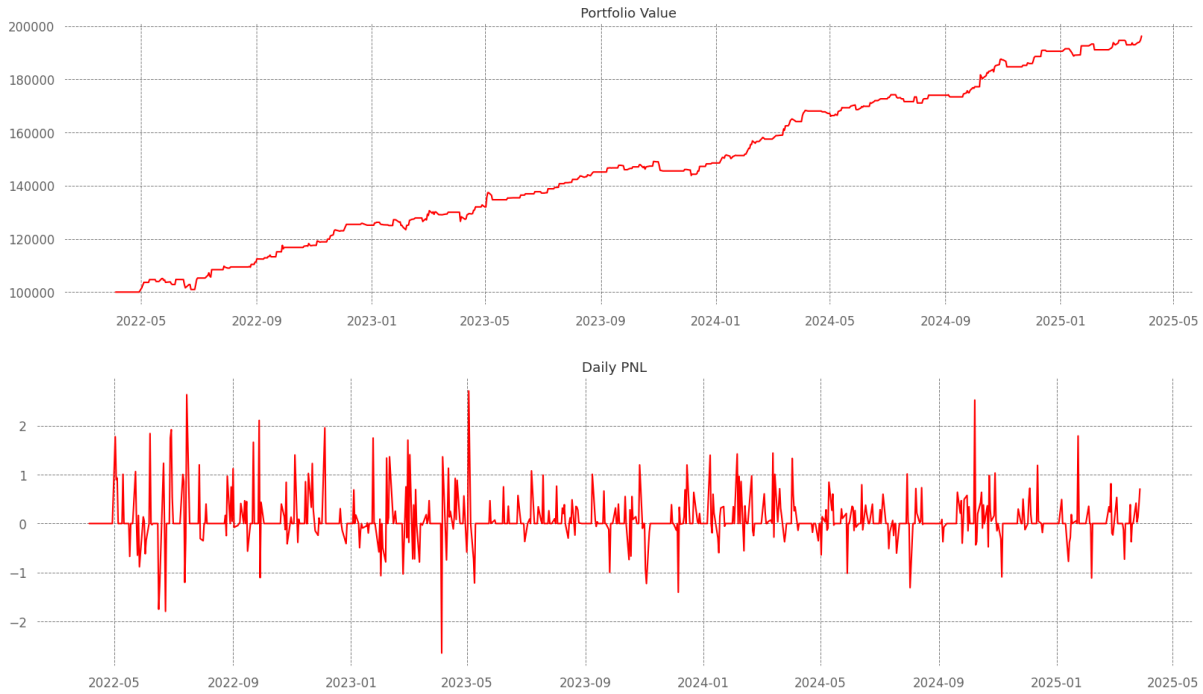


The threshold of 1 is chosen to enter trades when the spread deviates by one standard deviation, signaling a meaningful divergence likely to revert. A stop loss of 0.5 adds a buffer,

exiting trades if the z-score exceeds 1.5 in either direction, limiting potential losses if the spread continues to diverge. This setup balances early trade entry with controlled risk.

Capital is divided in the ratio `gamma:1` to maintain consistency with the spread calculation and ensure the trade is hedged correctly based on the historical relationship between the two stocks.

BACKTESTING RESULTS



Profit (Cumulative Return): 96.21%

CAGR: 25.46%

Sharpe Ratio: 2.71

Maximum Drawdown: 3.91%

Trades Taken: 237

This trading strategy yielded a strong cumulative return of **96.21%**, with a solid **Compound Annual Growth Rate (CAGR) of 25.46%**, indicating consistent performance over time.

The **Sharpe Ratio of 2.71** reflects excellent risk-adjusted returns, meaning the strategy generated high returns relative to its volatility.

With a **maximum drawdown of just 3.91%**, it maintained strong downside protection.

A total of **237 trades** were executed, suggesting frequent opportunities while keeping risk in check.Overall, a well-balanced and efficient strategy.