Finance Club
Open Project Summer 2024

USER GUIDE

Title: Pairs Trading Algorithm for Financial Markets (#FC24OPS2)

Strategy Description :

This strategy involves selecting two correlated stocks from the IT sector on the NSE (National Stock Exchange). The strategy uses statistical methods such as cointegration to determine the relationship between the selected pairs. By using the methods we can see TECHM(Tech Mahindra) & PERSISTENT(Persistent Systems Ltd) are correlated stocks to be used in the stratergy.

Calculating the spread = Y-b*X
where b is the linear regression constant for making the spread stationary which is independent of time

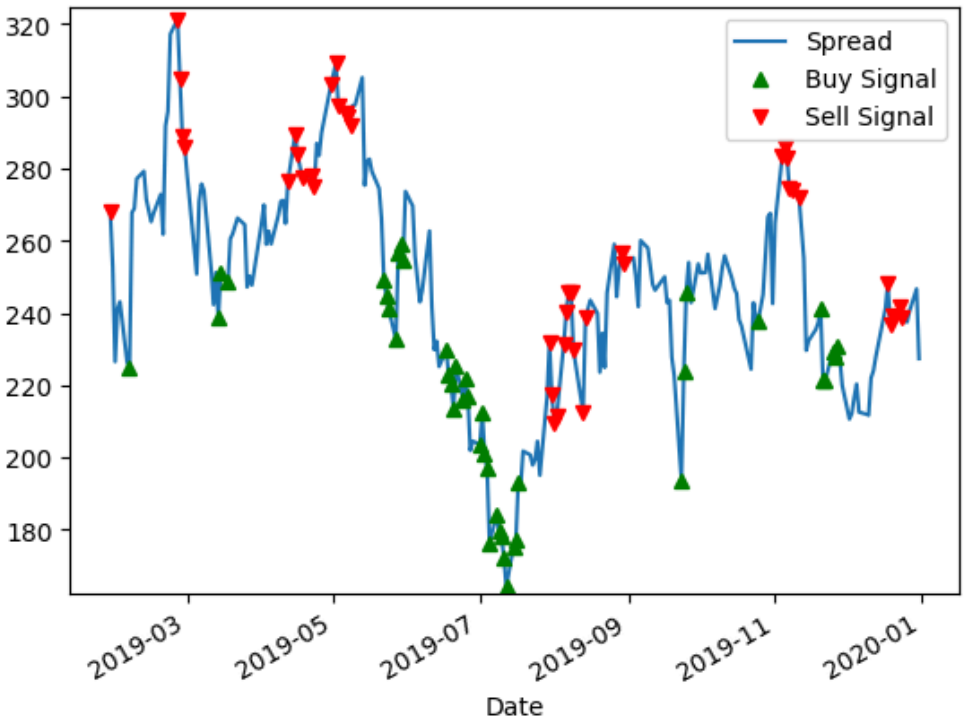
Computing the z score of the Moving average of the spread

Zscore = mean(returns)/stdev(returns)

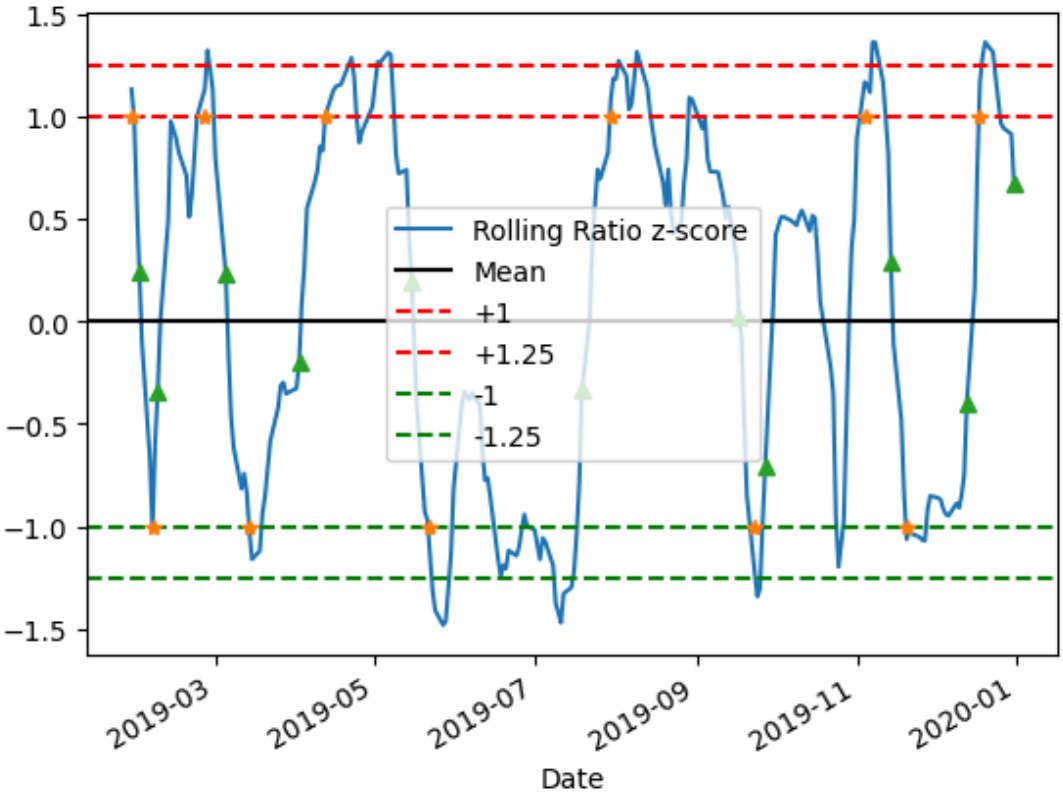
If the z score exceeds +1 giving a SELL signal to spread (SHORT Y LONG X)

And if the z score drops below -1 giving BUY signal to spread(LONG Y SHORT X)

Both the positions are squared off whenever the Z-score changes its sign or touches the 0.



INTERPRETING BUY AND SELL SIGNALS FOR SPREAD



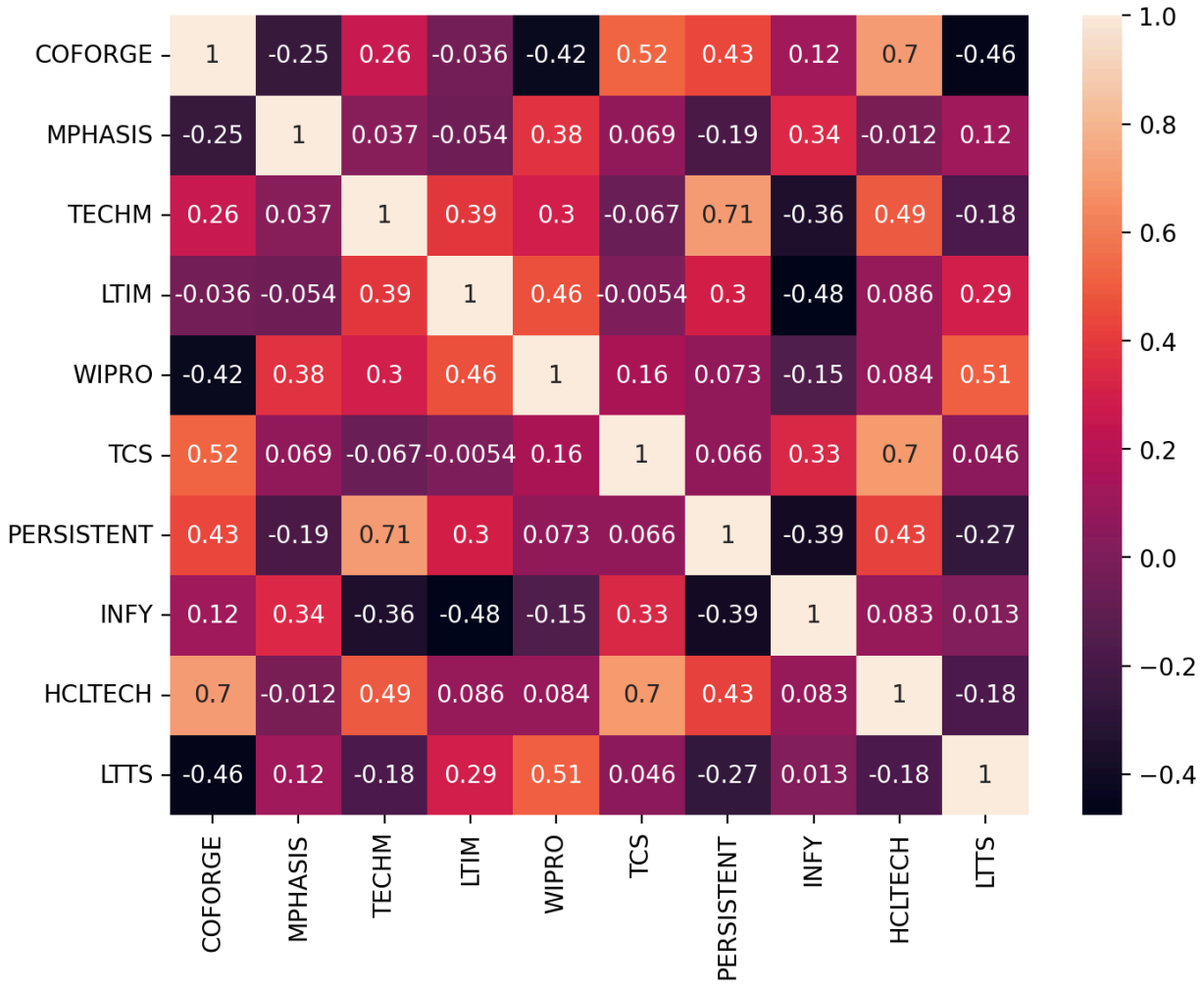
Z SCORE AND PLOTTING THE OPENING AND CLOSING POSITIONS

CHANGING THE PARAMETERS:

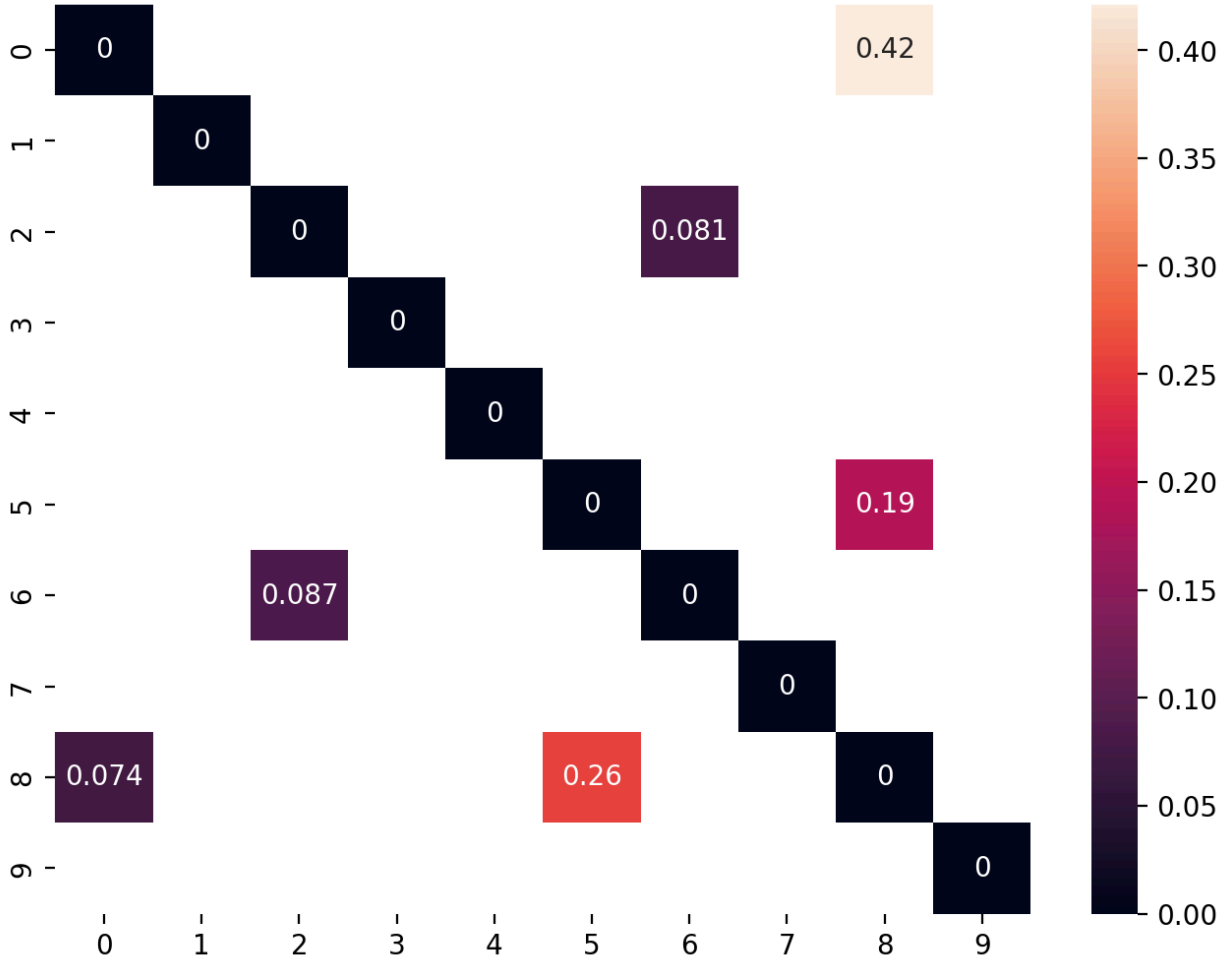
1. INCLUDING THE DATA SET FROM NIFTY-50

You have the flexibility to select any group of assets, such as IT stocks from NIFTY, ensuring they belong to the same sector for the algorithm's pair selection process. This approach allows for focused analysis within specific industries, leveraging sector-related trends and correlations to optimize pairings effectively. By homing in on a cohesive asset category, like IT within NIFTY, you enhance the algorithm's ability to identify and capitalize on sector-specific patterns, potentially improving trading outcomes through more precise and informed decision-making based on industry dynamics and performance metrics.

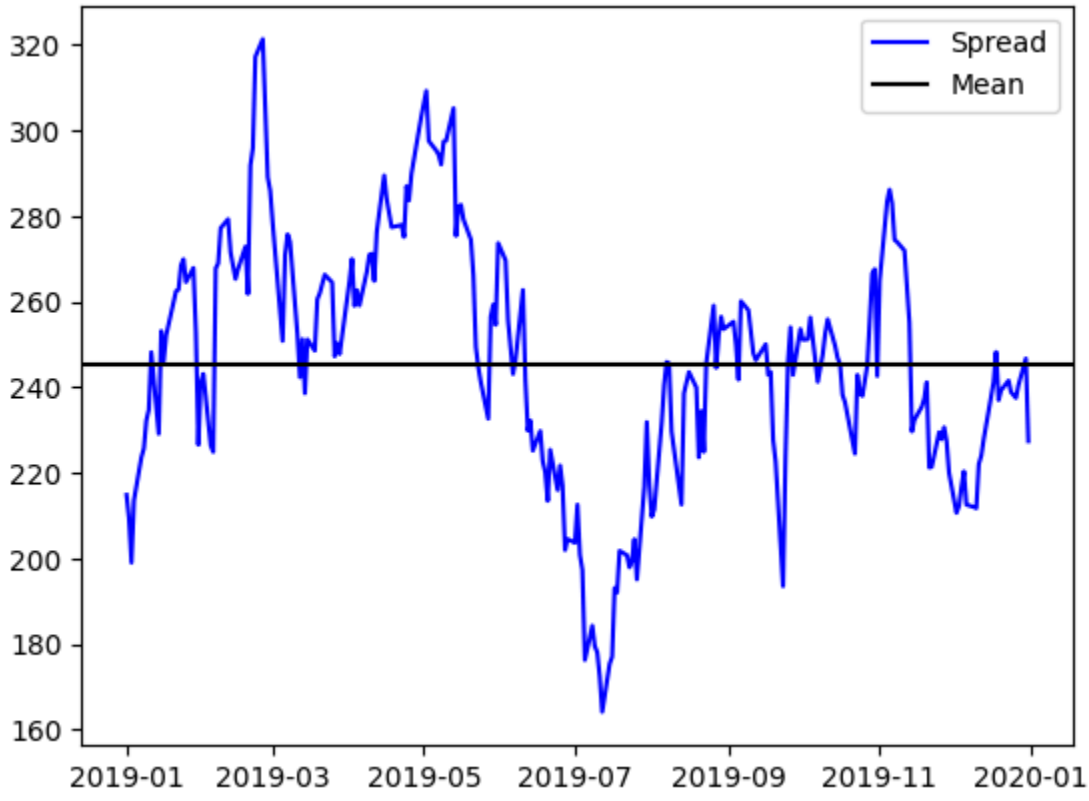
You can now plot correlation graph choosing the pair of assets with Correlation > 0.6/0.7 by choosing min_corr



After filtering stocks based on corrected correlations (> 0.6), pairs with the lowest cointegration are selected. For example, in this analysis, Tech Mahindra and Persistent emerge as a pair due to their minimal cointegration. This approach ensures robustness by focusing on stocks with significant correlated movements and low integration, optimizing the algorithm's ability to identify pairs likely to maintain distinctive price behaviors, crucial for informed trading strategies.



After selecting pairs with low cointegration and high (> 0.6) corrected correlations, calculate the spread Y - b * X, where b is determined optimally. Use linear regression to find the best-fit value for b. This step ensures the spread calculation is based on the most accurate relationship between the pair's prices. For instance, a chosen value like b = 1.3718448073406406 can be used to define the spread, optimizing trading strategies based on calculated price differentials.



You can adjust the value of (b) as necessary to ensure the spread remains stationary, consistent over time in Time Series analysis. This stability is crucial for reliable trading strategies, ensuring the spread calculation reflects a consistent relationship between asset prices, typically optimized through methods like linear regression to find the most suitable (b) value.

After calculating the spread between paired assets, the next step involves computing the Z-score to assess deviations from the historical moving average (MAVG). Typically, a moving average of 5 days and 20 days is chosen, ensuring these periods are shorter than the overall training period to capture recent price trends effectively. The Z-score is calculated using the difference between these moving averages, where the mean and standard deviation of the spread over the selected MAVG periods serve as the basis. This Z-score graph provides insights into whether the spread is above or below its historical average, helping to identify potential buy or sell signals based on deviations from the mean.

Once the Z-score graph is generated, adjusting thresholds becomes crucial for determining actionable trading signals. Commonly, thresholds such as Z-score values of +1 and -1 are used to indicate overbought and oversold conditions respectively. These thresholds can be fine-tuned based on historical data analysis and market conditions to optimize signal accuracy. Interpreting the results involves monitoring the Z-score graph: crossing above the upper threshold suggests a potential sell signal, while dropping below the lower threshold indicates a possible buy signal. This approach leverages statistical analysis to enhance trading decisions, aligning strategies with observed market dynamics and historical performance metrics.

Here's a summary of the parameters and steps for optimizing the trading strategy:

1. Pair Selection: Choose pairs with high (> 0.6) corrected correlations and low cointegration for stable spread calculations.

2. Coefficient Determination: Use linear regression to find the optimal coefficient \(b \) for the spread formula \(Y - b \times X \), ensuring the spread is stationary over time.

3. Z-Score Calculation: Compute the Z-score using moving averages (e.g., 5-day and 20-day), where the mean and standard deviation of the spread over these periods determine deviations.

4. Threshold Adjustment: Adjust buy and sell signal thresholds based on Z-score values (e.g., +1 for overbought, -1 for oversold), optimizing for historical performance and market conditions.

5. Iterative Optimization: Continuously refine parameters through iterative testing and analysis of historical data to enhance trading strategy effectiveness and adaptability.

By systematically adjusting these parameters, you can optimize the strategy to capitalize on market trends while minimizing risks associated with price volatility and other market dynamics.