# Pairs Trading Algorithm for Financial Markets: Logic, Inputs, and Outputs

The Pairs Trading Project is structured across four distinct phases, each implemented in a separate Jupyter notebook:

1. Identifying the Optimal Pair
2. Strategy Formulation
3. Strategy Back-testing
4. Risk Analysis

Each notebook corresponds to a crucial step in the project, facilitating a systematic approach to pairs trading strategy development. Below are the detailed descriptions of each phase:

## 1. Identifying the Optimal Pair: “findingPair.ipynb”

### Stock Selection

* **Selection Process**: The project commenced by selecting a comprehensive set of 200 stocks listed on the Bombay Stock Exchange (BSE).
* **Data Acquisition**: Data was acquired from Yahoo Finance, covering a period from 2013 to 2023, ensuring a robust dataset for analysis.

### Filtering

* **Loss-Making Stocks**: Stocks demonstrating consistent losses throughout the decade were filtered out.
* **Rationale**: This step was essential to align the strategy with potentially profitable outcomes even amid downturns, ensuring that only fundamentally strong stocks were considered.

### Correlation Analysis

* **Identification of Pairs**: Pairs of stocks with a correlation index exceeding 0.9 were identified. A high correlation indicates a strong relationship between the price movements of the two stocks.
* **Significance**: This step ensures that the selected pairs have a tendency to move together, which is critical for the pairs trading strategy.

### Cointegration Analysis

* **Statistical Tests**: Further analysis was conducted to identify correlated stocks exhibiting cointegration, using statistical tests named Engle-Granger test.
* **Criteria**: Cointegration was confirmed if the p-value was less than 0.05, ensuring the presence of mean reversion characteristics.
* **Importance**: Cointegration ensures that despite short-term deviations, the prices of the stocks tend to revert to a long-term equilibrium, which is a key premise of pairs trading.

### Optimal Pair Selection

* **Final Pair**: After rigorous analysis, the optimal stock pair was determined to be “ZEELEARN.BO” and “GICRE.BO”.
* **Reasoning**: These stocks exhibited strong correlation and cointegration, making them suitable for the pairs trading strategy.

## 2. Strategy Formulation: “pairTradingStrategy.ipynb”

### Mean Reversion

* **Observation**: Pronounced mean reversion trends were observed in the ratio graph of the stocks, indicating that the ratio of their prices frequently reverted to the mean.
* **Defining Characteristic**: The ratio of the stocks was chosen as the defining characteristic for the strategy due to its mean-reverting nature.

### Z-score Calculation

* **Purpose**: The z-score of the ratio was computed to normalize the ratio and establish trading boundaries.
* **Formula**: The z-score is calculated as (short-term moving average − long-term moving average) / standard deviation.
* **Interpretation**: This z-score indicates how far the current price ratio deviates from its historical mean.

### Strategy Development

* **Trading Rules**:
  + **Long Position**: If the z-score fell below -0.6, indicating a negative deviation from the mean, it signaled that ZEELEARN.BO was undervalued relative to GICRE.BO. The strategy involved taking a long position on ZEELEARN.BO and a short position on GICRE.BO.
  + **Short Position**: Conversely, if the z-score exceeded 0.6, suggesting a positive deviation from the mean, it implied that ZEELEARN.BO was overvalued relative to GICRE.BO. The strategy involved taking a short position on ZEELEARN.BO and a long position on GICRE.BO.
  + **Exit Signal**: An exit signal was triggered when the z-score approached 0.25, indicating that the ratio was close to its mean and it was time to close positions.

### User Parameters

* **Customization**:
  + **Window Sizes**: Users could choose the window sizes for the short-term (window1) and long-term (window2) moving averages.
  + **Trading Frequency**: Users could select the trading frequency (e.g., 1 day, 5 days, or 1 week), allowing them to tailor their trading approach according to their preferences.
  + **Parameter Details**:
    - **window1**: Determines the window size for the short-term moving average (ma1), making it more responsive to recent price changes.
    - **window2**: Determines the window size for the long-term moving average (ma2) and the standard deviation (std), providing a broader perspective of the price trend.

## 3. Strategy Back-testing: “backTesting.ipynb”

### Performance Metrics

* **Assessment Criteria**: The strategy was assessed based on profitability, Sharpe ratio, volatility, and a comparative analysis of returns against the BSE index and NIFTY 50 returns.
* **Key Metrics**:
  + **Profitability**: Measured the total returns generated by the strategy.
  + **Sharpe Ratio**: Evaluated the risk-adjusted returns.
  + **Volatility**: Assessed the risk associated with the strategy.

### Testing Periods

* **Time Frames**: The strategy underwent rigorous testing over periods of 1, 3, and 5 years.
* **Results**: Consistent profitability was observed, with respectable Sharpe ratios and volatility metrics.

### Stress Testing

* **Market Conditions**: The strategy was tested during both bullish (post-COVID recovery, 2022-23) and bearish (COVID pandemic, 2019-20) market conditions.
* **Outcomes**: Remarkably, the strategy remained profitable during these adversities, demonstrating its robustness.

### Trading Frequencies

* **Frequency Testing**: Various trading frequencies were tested, including daily, weekly, and monthly trades.
* **Findings**: Each frequency yielded profitability, though returns varied, highlighting the flexibility of the strategy.

## 4. Risk Analysis: “riskProfiling.ipynb”

### Risk Profiles

* **Flexibility**: Three distinct risk profiles (low, medium, and high risk) were devised to offer users flexibility according to their risk tolerance.
* **Profile Characteristics**:
  + **Low Risk**: Conservative parameters with lower potential returns.
  + **Medium Risk**: Balanced parameters with moderate potential returns.
  + **High Risk**: Aggressive parameters with higher potential returns.

### Profile Parameters

* **Parameters**: These profiles were formulated based on diverse parameters, including moving average windows, trading frequencies, and z-score thresholds.
* **Customization**: Users can select a profile that aligns with their risk appetite and investment goals.

### Profitability Analysis

* **Back-testing**: Back-testing across these profiles showed significant variability in profitability.
* **Results**:
  + **High Risk Profile**: Yielded the highest profit.
  + **Low Risk Profile**: Resulted in the lowest profit.
* **Conclusion**: This analysis allows users to choose a risk profile that suits their investment strategy.