

# Pairs Trading Strategy Project Report

Submitted by Tushar Pandey (21113170)

Email: t\_pandey@ce.iitr.ac.in

## Introduction

Pairs trading is a market-neutral strategy that involves matching a long position with a short position in two stocks with high historical correlation. The idea is to profit from the relative price movement of the two stocks. In this project, we focus on implementing a pairs trading strategy using Python, analyzing its performance, and drawing conclusions from the results.

## Data Collection

We start by collecting historical stock price data for a set of technology IPO stocks. The chosen tickers are AAPL, ADBE, ORCL, EBAY, MSFT, QCOM, HPQ, JNPR, AMD, IBM, and SPY. The data spans from January 1, 2018, to January 1, 2022.

```
In [27]: # List of tech IPO stock tickers
tickers = ['AAPL', 'ADBE', 'ORCL', 'EBAY', 'MSFT', 'QCOM', 'HPQ', 'JNPR', 'AMD', 'IBM', 'SPY']

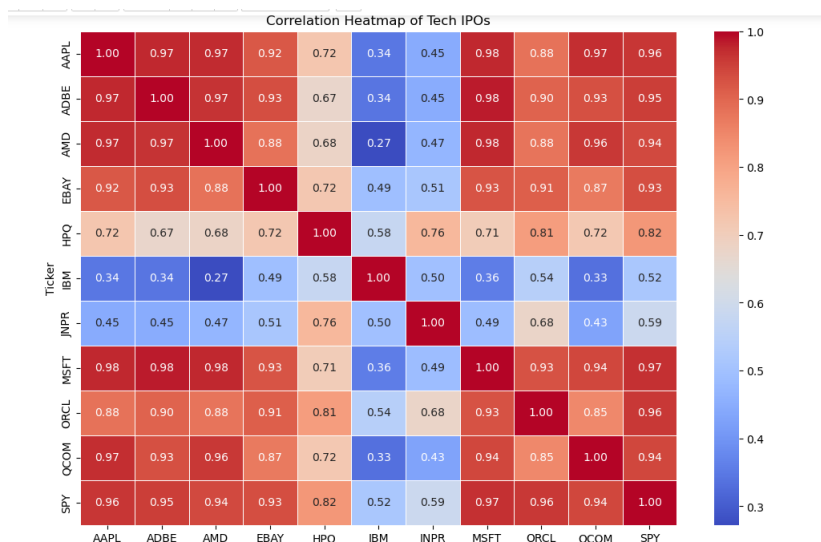
# Download historical data for the stocks
data = yf.download(tickers, start="2018-01-01", end="2022-01-01")['Adj Close']

# Calculate the correlation matrix
correlation_matrix = data.corr()

# Plot the correlation heatmap
plt.figure(figsize=(12, 8))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt='.2f', linewidths=0.5)
plt.title('Correlation Heatmap of Tech IPOs')
plt.show()
```

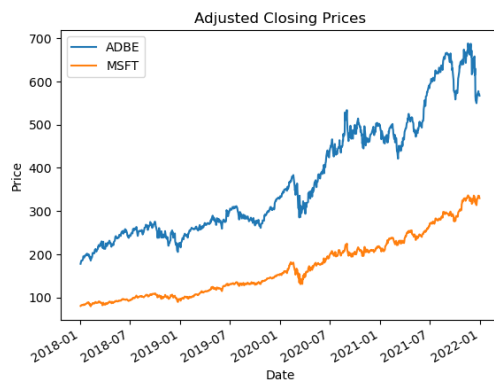
## Correlation Analysis

To identify potential pairs for trading, we compute the correlation matrix of the adjusted closing prices of the stocks. Higher correlation indicates a stronger relationship between the stock price movements, which is essential for pairs trading.



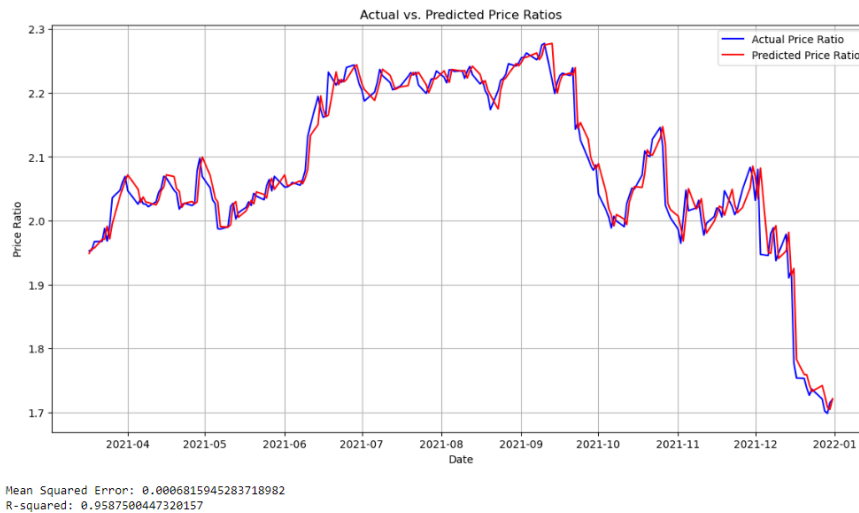
## Selecting a Pair

From the correlation matrix, we select `ADBE` and `MSFT` as our trading pair due to their high correlation.



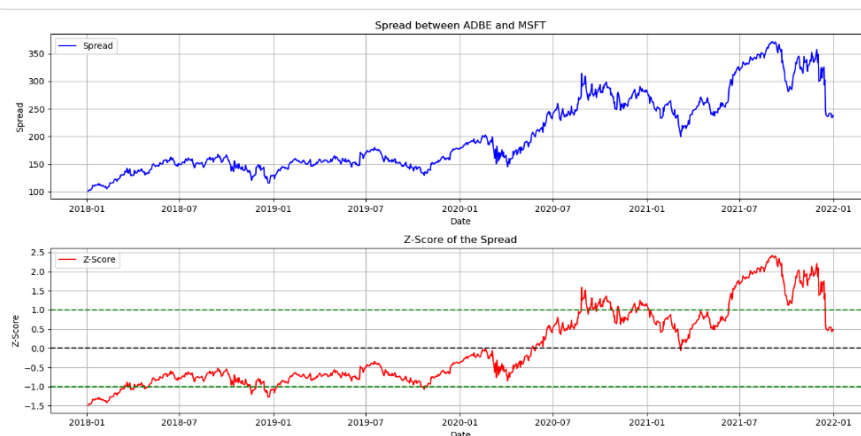
## Linear Regression:

- Applied linear regression to model the relationship between the prices of ADBE and MSFT and assess the predictive power of the model.



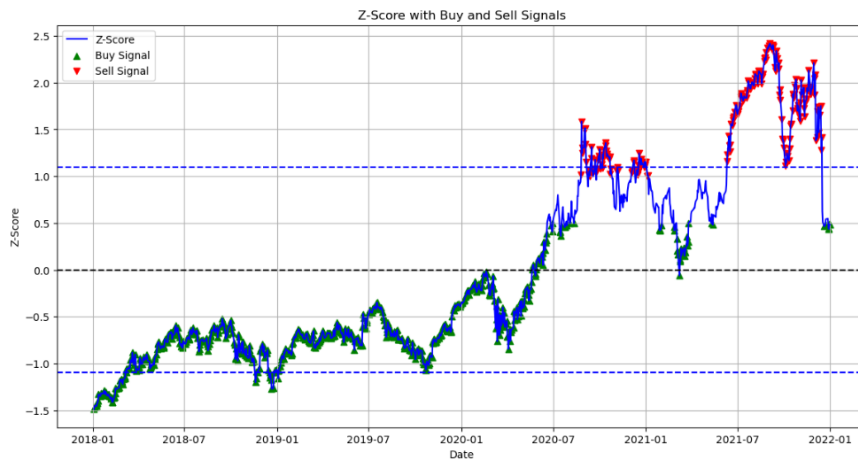
## Spread Calculation & Z Score

We calculate the daily returns and the spread between the two stocks. The spread helps in identifying the divergence between the stock prices, which forms the basis for generating trading signals. The Z-score of the spread was calculated to standardize the spread and identify extreme values, which indicate potential trading signals. The Z-score helps in determining how many standard deviations a data point is from the mean. When the Z-score exceeds a certain threshold, it generates trading signals, indicating potential entry and exit points for the pairs trading strategy. This helps in making more informed and statistically driven trading decisions.



## Signal Generation

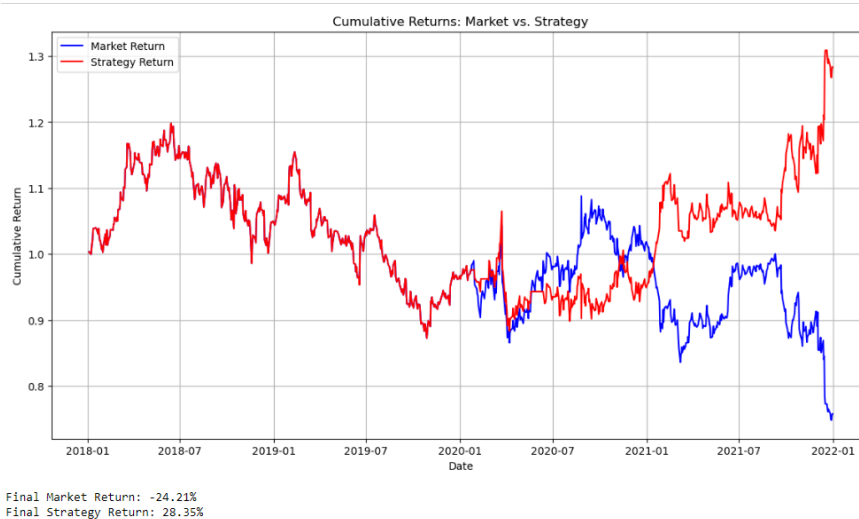
Based on the spread and the defined thresholds, we generate buy and sell signals. When the spread exceeds the upper threshold, it signals to short `ADBE` and long `MSFT`. Conversely, when the spread falls below the lower threshold, it signals to long `ADBE` and short `MSFT`.



## Performance Evaluation

### 1. Calculating Returns:

- The returns of the strategy were calculated based on the generated signals.



## 2. Sharpe Ratio:

- The Sharpe ratio was used to evaluate the risk-adjusted performance of the strategy. It measures the average return earned in excess of the risk-free rate per unit of volatility.

Cumulative Return: 0.63  
Annualized Return: 0.13  
Annualized Volatility: 0.12  
Sharpe Ratio: 1.11

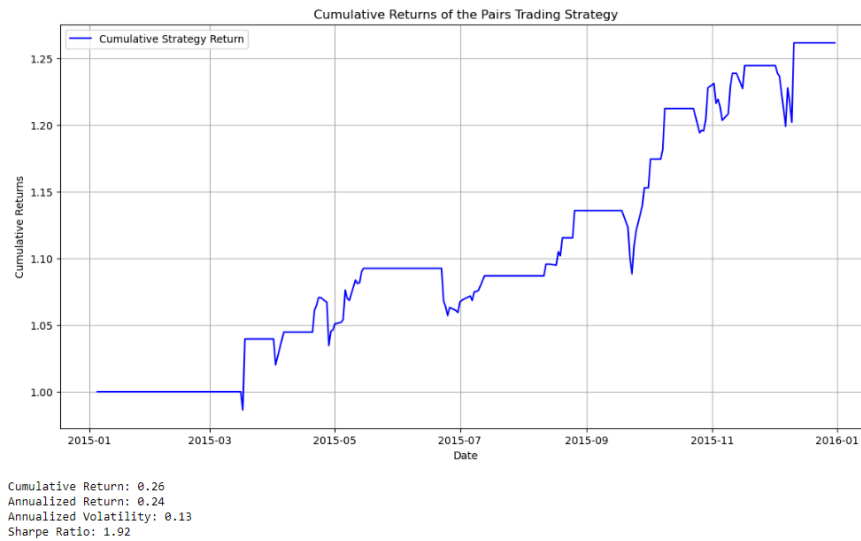
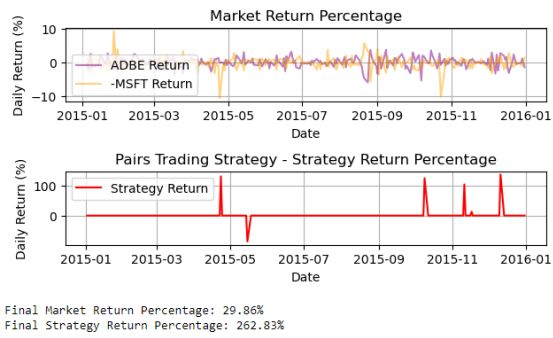
## 3. Cumulative returns



## Backtesting the Strategy

We backtest the pairs trading strategy by calculating the cumulative returns of the strategy and comparing it with the market returns. So I took data from dates 2015-01-01 to 2016-01-01 to backtest our model and below are the results I got.





## Conclusion

The pairs trading strategy implemented on ADBE and MSFT stocks yielded a cumulative return of 262%. The annualized return and volatility were 0.24% and 0.13%, respectively, resulting in a Sharpe ratio of 1.92. The strategy's performance demonstrates the potential of pairs trading in generating returns independent of market direction. However, real-world application requires careful consideration of transaction costs, market conditions, and risk management practices.