agnostiq Challenge
Pair Trading Stocks in the S&P 500

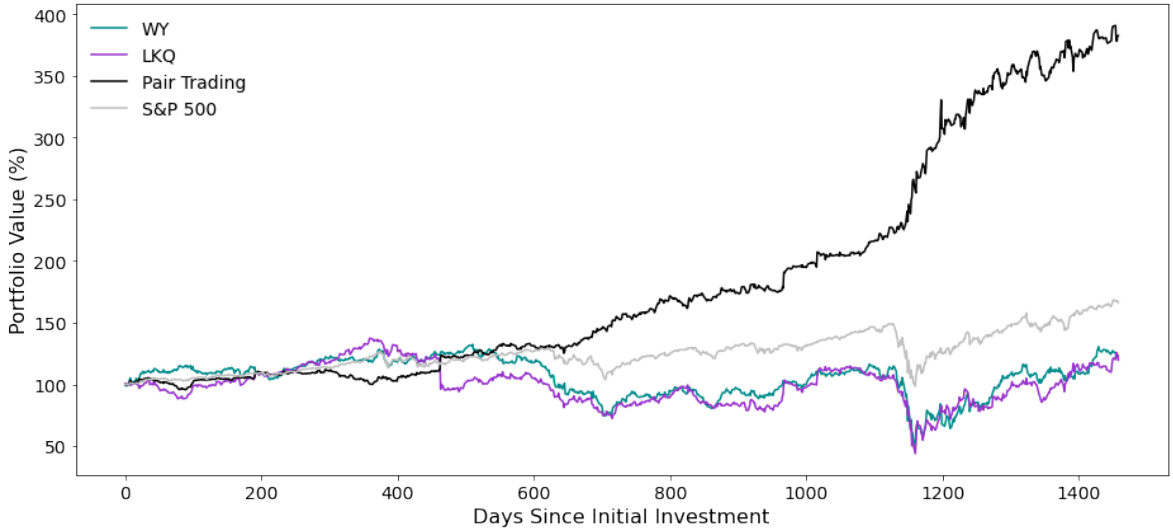
The Problem: For this challenge, I have been tasked with analyzing the S&P 500 market to find those stocks that pair trade. As stated in the challenge outline: "Pair trading is the method of modeling pairs of stocks that are co-integrated where the price difference follows a time series process that always returns to a constant value, given enough time. Using this model, one can find arbitrage opportunities that can then be used as a trading strategy."

Gathering Data: Using the python tool yfinance to access market data from Yahoo! Finance's API, I collected 5 years of historic trade data for the stocks in the S&P 500 index. I also collected analogous data for ^GSPC, which tracks the performance of the index writ large, to use as a baseline when comparing the effectiveness of varying investment strategies. I then stored this data in a SQL database file for ease of subsequent access and analysis.

Idealized Pair Trading Model: Effective pair trading relies on identifying stocks whose trading values tend to retain a fixed price differential. Let's call this differential Δ and assume we have two stocks with values P_1 and P_2 that vary throughout the trading period. The simplest pair trading case would be that where $\Delta = P_1 - P_2$ is approximately constant over time, but this seems unlikely to be a common arrangement in practice. Instead I will utilize a scaled Δ that takes the ratio of the two stock prices and their mean values, $\Delta_{scaled} = P_1/\bar{P}_1 - P_2/\bar{P}_2$.

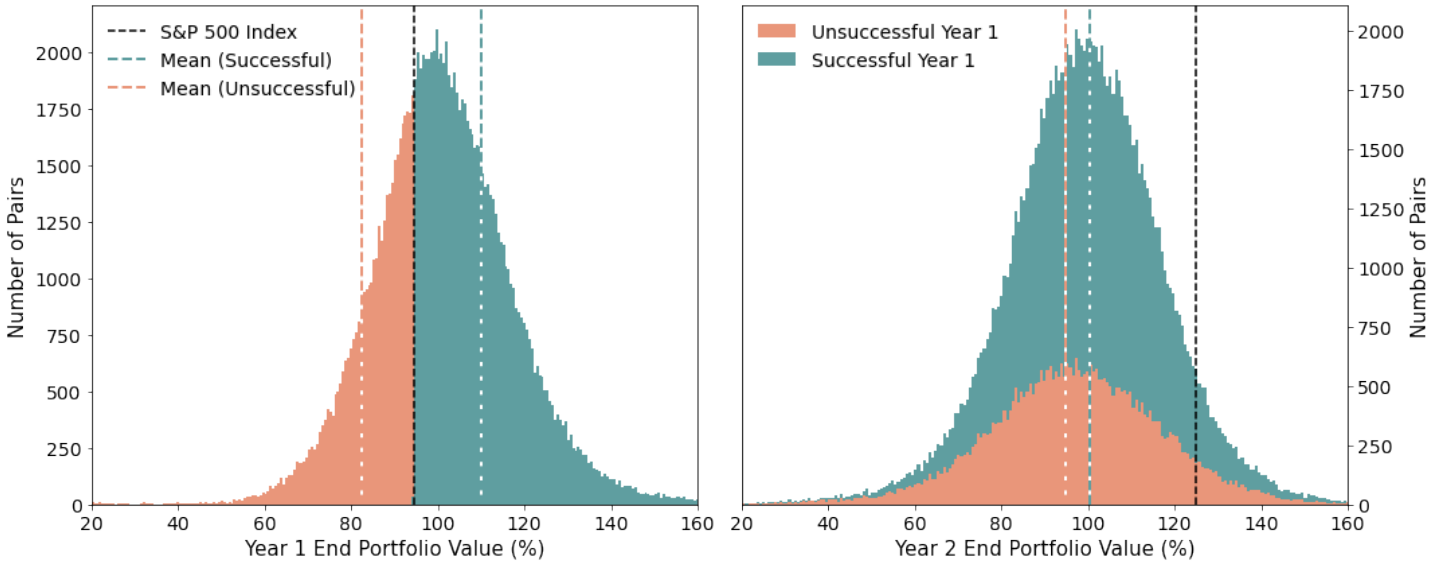
The basic assumption of the model is that, when $P_1/\bar{P}_1 - P_2/\bar{P}_2 > \Delta_{scaled}$, the value of Stock 1 will fall relative to that of Stock 2 in the relatively near future, and similarly, when $P_1/\bar{P}_1 - P_2/\bar{P}_2 < \Delta_{scaled}$, the value of Stock 1 will rise relative to that of Stock 2. Operating under these assumptions, a basic pair trading strategy can be constructed wherein we attempt to short those stocks expected to drop in value in and go long on – which is to say, buy shares of – those expected to rise. To explore this strategy in a somewhat practical setting, I coded up a two-stock portfolio manager that evaluates at market close whether or not their share values still satisfy the current positions and, if warranted, adjusts them at the next market open.

The figure below illustrates the potential benefits of pair trading. While neither WY nor LKQ come close to matching the S&P 500 as individual investments, a pair trading portfolio built around them does remarkably well, nearly quadrupling in value over a four year trial period and outperforming the S&P 500 by more than a factor of two. Notably, the pair trading portfolio consistently gains value even during significant downturns in the market writ large, as occurred in early 2020 (around Day 1150 below).



Real-World Limitations: It is not especially difficult, with the benefit of hindsight, to identify cases where pair trading would have been beneficial. In order to be a viable foundation for practical investing, however, one must be able to identify cases where pair trading is likely to be effective in the future. To that end, I produced three years of training data by examining the effectiveness of pair trading for each combination of S&P 500 stocks based on their Δ_{scaled} metric from the year prior, setting aside the most recent year of data for later use as a final test set. The idea is simple enough – if two stocks have consistently been effective pair trade targets in the past, they will likely continue to do so in the future. In practice, though, this does not seem to hold, at least not to the extent that it can easily be used to outperform the S&P 500.

The stacked histograms below illustrate the problem well. The distributions show the value of idealized pair trading portfolios for pairs of stocks in the S&P 500 over two years of trial trading (from January 2018 to January 2020). The pairs which outperformed the S&P 500 in Year 1 are shown in teal, and those which underperformed it are shown in salmon. While there is certainly some useful information encoded in that distinction – the successful pair trades from Year 1 clearly outperform the unsuccessful ones in Year 2 – to beat the market, it is not sufficient to simply identify previously successful pairs.



Applying Machine Learning: After updating the training data to include monthly information on stock returns and relative market performance, this situation essentially reduces down to a standard classification problem. Given the time constraints and my limited computational resources, I was not able to do an exhaustive search of potential classifiers or implement a full-fledged cross-validation approach to optimize those which I did examine. Nevertheless, I was able to obtain some promising early results.

An obvious concern with a problem like this is overfitting the training data. This is true in general for classification, of course, but is especially relevant here, where the three years of training data are almost certainly to fundamentally differ from future years to some extent. This is difficult to quantify when training/optimizing the classifiers because these fundamental changes in the market's future trading are not going to be evident in the validation data. That is, while we could generally combat overfitting directly by looking at accuracy metrics from the validation data, this approach is only partially effective here.

Because of this, if metrics are otherwise comparable, I lean toward using classifiers that have a higher intrinsic degree of bias. The idea here is that a biased model is more likely to be sensitive to broad features of the stock market, features that will hopefully be reflected in future trading periods. Thus, while the random forest and extra trees classifiers have nearly the same validation accuracy, as an analyst I put

more weight in future predictions made with the extra trees classifier. This has the added bonus of extra trees classifiers being a good bit faster to train than their random forest cousins.

I examined seven classifiers in total, the random forest and extra tree classifiers mentioned previously, a logistic regression classifier, LDA and QDA classifiers, a Naïve Bayes classifier, and a voting model. Ultimately the extra trees classifier performed the best, accurately classifying more than 82% of the pairs in the validation set as either overperforming or underperforming the S&P 500. I then used this classifier to make predictions for which pairs should be targeted in the future. Limiting my test portfolios to 100 pairs, I worked my way down the list of pair stocks with the highest average pair trading returns over the three years of testing. If they had outperformed the S&P 500 during all three previous years, then I added them to what I called the conventionally selected portfolio. If they were predicted by the extra trees classifier to beat the market in the next trading year, I added them to the machine learning portfolio.

Gaining 6.2% during the final year of test data, the machine learning portfolio significantly beat out the conventionally selected portfolio, which only gained 2.9%. Ultimately though, while both of them gained in value, neither test portfolio beat the market, which gained 17.8% during the trading year. It seems like an awful lot of work to set up and manage a hundred active pair trades only to make less profit than you would have made having simply invested in a good index fund. Thus, given the current state of these classifiers, I do not think a generalized pair trading strategy is a good idea.

Practical Recommendations: While this strategy is not likely to be worthwhile in the general case, when combined with additional modeling it may still prove very profitable. A major reason pair trading can fail is that when the market is consistently making significant gains, the short component of the investment will often negate gains made by the long positions. This problem is mitigated in times when the market is gaining only marginally, or when it is losing value. Thus, if other predictive models available to an investing firm – or obvious external factors like major world events – give cause to believe that a market retraction is imminent, it would be prudent to activate a pair trading scheme. In such a case choosing your portfolio using something like the extra trees classifier I identified would make sense and likely outperform a conventionally selected portfolio significantly. Additionally, if a firm has sufficient archival data and modeling capability to accurately estimate Δ_{scaled} for the next trading period, pair trading becomes a viable option even during strong market growth. If, for example, I had been able to accurately predict Δ_{scaled} during the final trading year, my machine learning-selected portfolio would have gained 25.7%, comfortably outperforming the market.