Louis Kabelka
3/24/2022
Statistical Computing

Pairs Trading Report

Pairs Trading is a stock trading strategy that was employed intensively during the 1980s. We find a pair of stocks that are positively correlated. Then we use the ratio to determine the best time to sell/buy each stock. When one stock is overpriced, we sell it and buy the other. When one stock is underpriced, we buy it and sell the other.

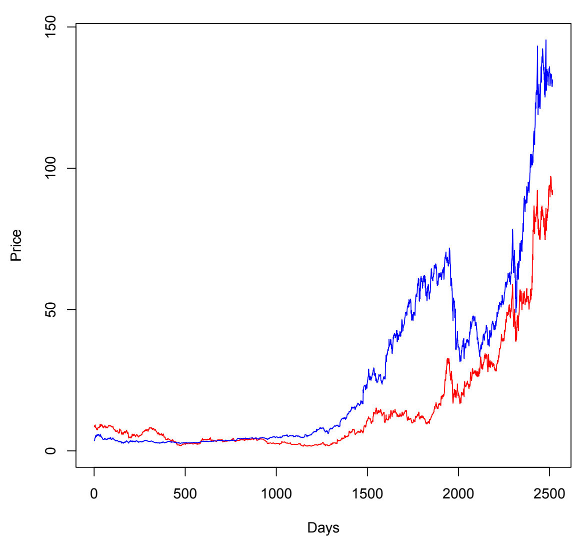
This project seeks to emulate the strategy on stock data from Yahoo! Finance. Next, we utilize a time series to simulate pairs of stocks with various different properties. Finally, we extend the project by finding the hypothetical optimal training fraction.

1. Two Positively Correlated Stocks

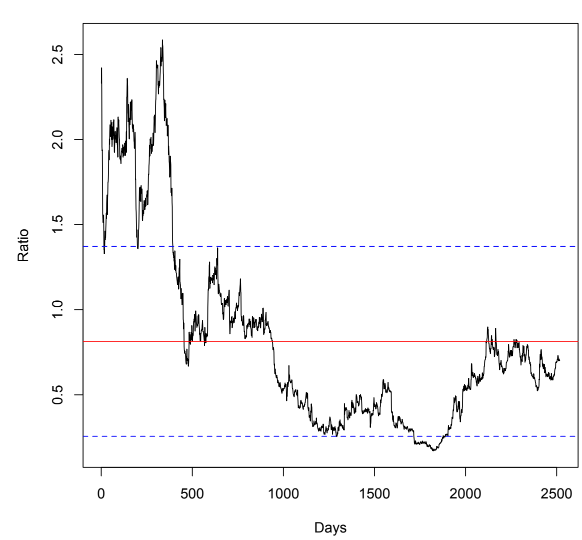
AMD and NVDA between January 1st, 2011 and January 1st, 2021.
The correlation between the stocks is 0.91613.

AMD is Advanced Micro Devices and NVDA is NVIDIA Corporation. They are both leading manufactures of computer hardware. In particular, AMD and NVIDIA have dominated the market for high-preforming graphics cards (GPU). However, AMD has become well-known for producing many different computer parts. They have become the Intel's primary competitor in the processor (CPU) market. On the other hand, NVIDIA is more hyper-focused on producing high-performing graphics hardware.

a. Prices of AMD (red), NVDA (blue)



b. Ratio



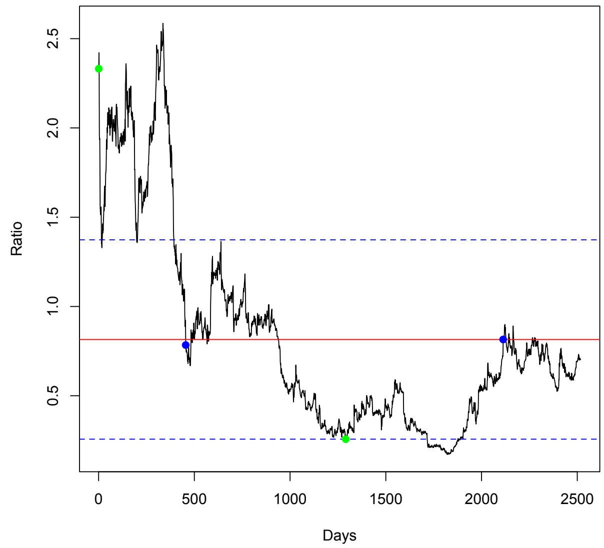
2. Opening / Closing Positions

Over 10 years, we can trade AMD & NVDA using Pairs Trading. We do this by selling one stock when it's overpriced, while buying the other when it's underpriced.

Let AMD be stock 1 and NVDA be stock 2. There are two pairs of open / close positions.

1st pair: High
Open: Day 1. Bought stock 1. Sold stock 2.
Close: Day 455. Sold stock 1. Bought stock 2.

2nd pair: Low
Open: Day 1291. Bought stock 2. Sold stock 1.
Close: Day 2113. Sold stock 2. Bought stock 1.



3. Profits

Calculations for 1st pair

Let x1 be stock 1 on opening day. Let x2 be stock 2 on opening day. Let y1 be stock1 on closing day. Let y2 be stock2 on closing day. [x1 = 8.470. x2 = 3.632. y1 = 2.180. y2 = 2.780.]

1. Calculate how many shares we buy for $1 on opening day.
1/x1 = 1/8.470 = 0.118
1/x2 = 1/3.630 = 0.275

2. Calculate how much our shares are worth on closing day.
1/y1 = 0.118 * 2.180 = 0.257
1/y2 = 0.275 * 2.780 = 0.765

3. We buy back stock 1 and sell stock 2. So, calculate the difference between stock 2 and stock 1.
0.765 - 0.257 = 0.508

4. Calculate the fees from opening and closing transaction with a fixed rate of 0.3%.
0.003 * (1 + 1 + 0.257 + 0.765) = 0.00906

5. Final profit is the difference between our stock profit and the fees.
0.508 - 0.00906 = 0.499

Profit of each pair
1st pair: 0.49903, 2nd pair: 10.35278

Net profit
The total net profit, summing each pair is 10.85182.

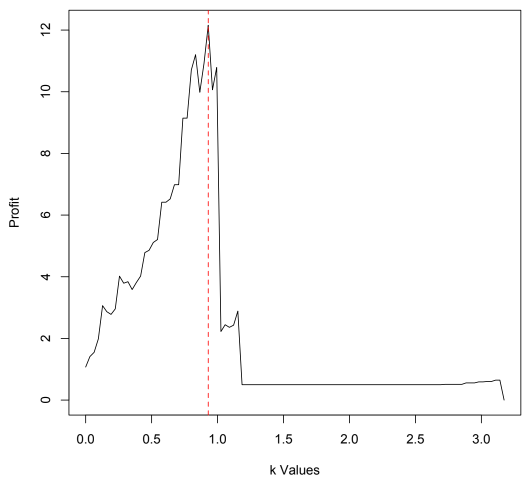
4. Optimal k

We usually trade each pair when the ratio crosses 1 standard deviation from the mean. However, if we can optimize how many k standard deviations we should go above the ratio.

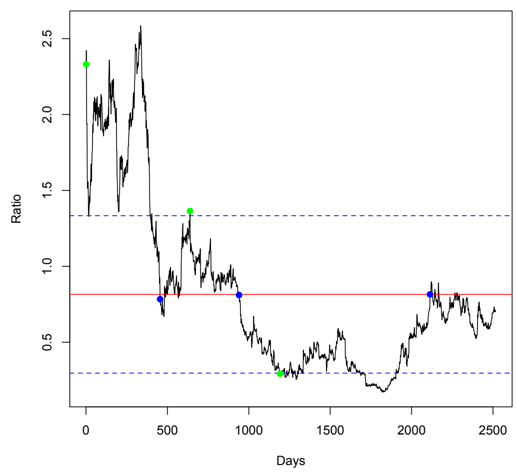
The optimal k value is 0.92928. With this optimal k value, it gives slightly more profit than the default, k = 1. It also gives three pairs, instead of of two.

With the optimal k value, the net profit is 12.15695.
[1st pair: 0.49903, 2nd pair: 0.51770, 3rd pair: 11.14021]

a. Profit vs. k Values



b. Profit positions with optimal k value

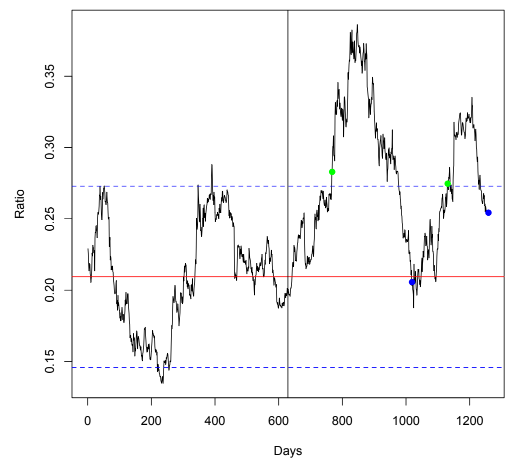


5. Training and Testing Data

The realism of our current strategy is poor. We cannot possibly know the historic mean, while trading in the same time frame. Instead we will divide the trading time frame in half.
One half will be for training the data to determine the historic mean, standard deviation. The other half will be for testing the data to do the trading.

We will test our new strategy on three different kind of stock pairs: positively correlated, not correlated and negatively correlated pairs.

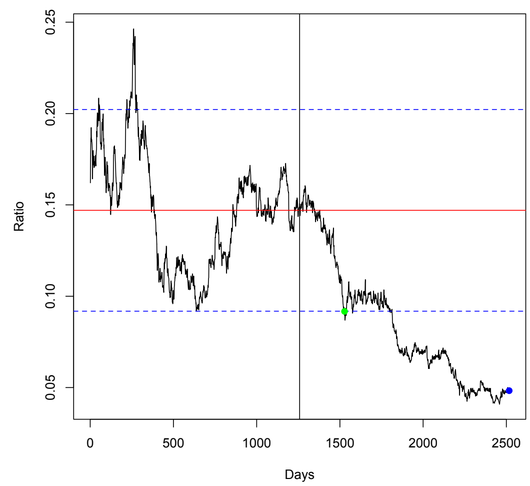
A. Positive
SONY and MCD between 2010 and 2020. The correlation is 0.74291. The net profit is 0.41422.



SONY is Sony. MCD is McDonalds. They are two companies that have a history of working together. They also are big companies that benefit from market growth in 1st world countries.

Thanks to the positive correlation, we are able to utilize pairs trading to make a profit.

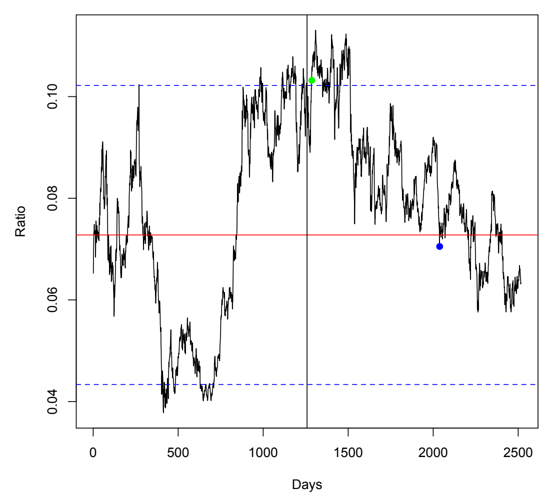
B. None
F and MCD between 2010 and 2020. The correlation is -0.13118. The net profit is -0.87359.



F is Ford. MCD is McDonalds. They don't share a sector. Therefore, they are practically not correlated at all.

Since, there is no correlation, the ratio has an essentially random flow. Therefore, we were not able to make a profit.

C. Negative
F and GLD between 2010 and 2020. The correlation is -0.47680. The net profit is 0.32740.



F is Ford. GLD is SPDR Gold Trust. Gold is said to be a "safe-haven". It is a reliable investment during times of market decline. Therefore, it is negatively correlated with Ford, which suffers during times of market decline.

Since, the stocks are negatively correlated, the stocks are moving in a predictable pattern. So, we were actually able to make a profit.

6. VAR Simulate Distribution

VAR stands for vector autoregressive models. To summarize, we can use various parameters to simulate stock pairs. The parameters are the following:

n: the length of the time series
rho: the within-stock correlation;
psi: the between-stock correlation;
b1/b2: the linear trend of each stock;
sigma1/sigma2: the standard deviation of each stock

* n = 1000
* No temporal trends
* simga1 = sigma2 = 1
* rho = psi =0.9

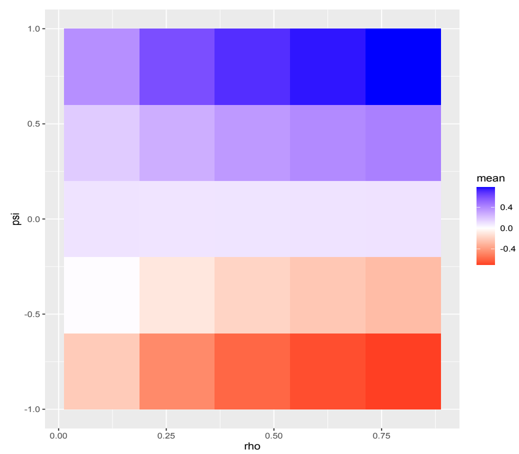
Mean Profit is 7.5334 with a standard error of 0.3295.
We are 95% confident that the true profit is between 6.88674 and 8.18008.
Mean Correlation is 0.83305 with a standard error of 0.001597.
We are 95% that the true correlation is between 0.86017 and 0.86643.

7. Mean Correlation Heatmap

[Rho is between 0.1 and 0.8. Psi is between -0.8 and 0.8.]

The mean profit is the highest when psi is big and rho is big. It lessens as rho gets smaller. It lessens even more strongly as psi gets smaller.

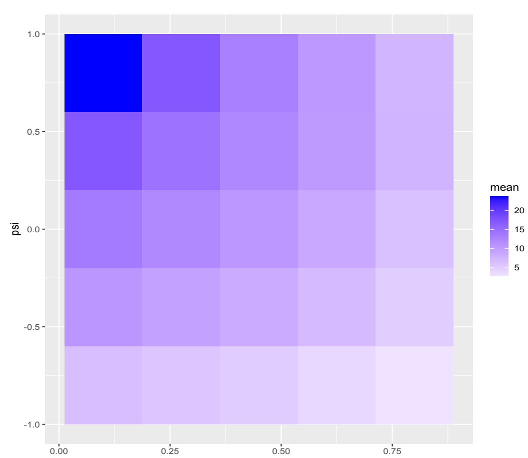
When psi is zero, the rho variable has no effect on the correlation.



8. Mean Profit Heatmap

[Rho is between 0.1 and 0.8. Psi is between -0.8 and 0.8.]

The mean correlation is the highest when psi is big and rho is small. It lessens as psi gets smaller.
It lessens at roughly the same strength as rho gets bigger.

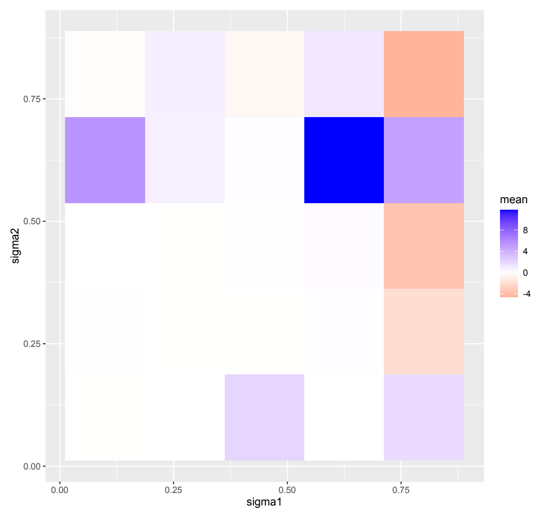


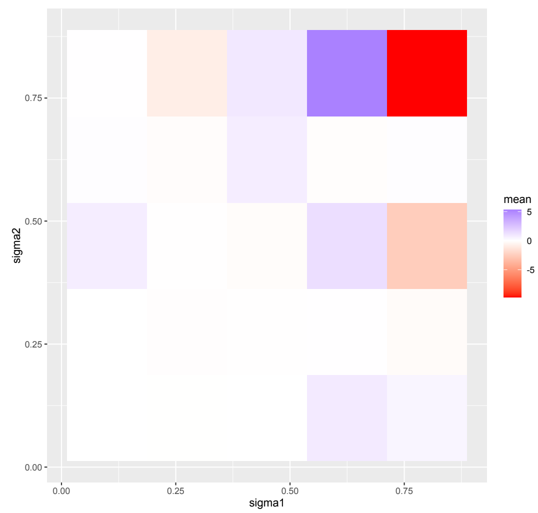
9. Sigma VAR

[Sigma1 and Sigma2 are both between 0.1 and 0.9]

The more we increase either of the sigma values, the nosier the mean profits become.

Sigma is supposed to be random noise value. It is modeled by the rnorm function. Therefore, any increase in the sigma values will always turn a slightly different result. Despite the following heatmaps having the exact same inputs, they produce very different results.





10. Conclusions based on VAR

A high correlation does not necessarily guarantee the highest possible profit. Comparing the mean correlation and the mean profit heatmaps, we can see that the highest correlation is in the upper rightmost corner, whereas the highest profit is in the upper leftmost corner.

Rather for a high profit, we need a high between-stock correlation. On the other hand, the within-stock correlation actually detracts from the profit. This is logical, since pairs trading relies on the stocks being related to each other.

Noise makes the profit less reliable. Maybe we will get a good profit, maybe a bad profit. If all other things considered, we are sure that we will have profit, then it is probably better to have less noise. That way we can guarantee that we will have at least some profit.

Extension
What is the optimal training fraction? How much time do we spend training the data vs. how much time do we spend testing the data?

Obviously, we cannot know what will be the optimal time to train our data in advance. We are using data that we wouldn't know, if we were actually trading stocks. The testing data is meant to simulate the experiences of using our Pairs Trading strategy. So, we cannot use that data to draw conclusions for profit. That would defeat the realism of the project.

But we can do these experiments in retrospect. If we notice trends, then that might help us in the future. If some particular type of stock pair tends to benefit from some particular training training fraction, then we can use that training fraction with other stock pairs.

To figure out what conclusions we can draw about the optimal training fraction:
1. Implement a findOptimalTF(stocksDF, plot = FALSE) function.
2. Test the function on several stock pairs.
3. Create a heatmap to find mean optimal training fractions for different rho, psi values.
4. Draw conclusions by comparing it to the Mean Profit / Mean Correlation heatmaps.

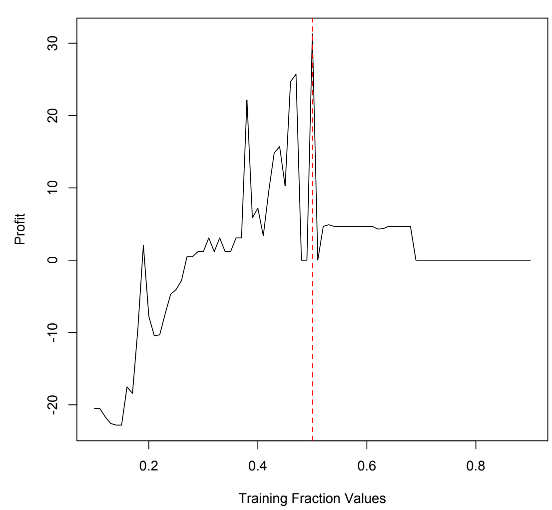
1. Optimal TF
The function to find the optimal training fraction is in the same vein as the function to find the optimal k value.
- We are only using values between 0.1 and 0.9. Training factions beyond those seem unreasonable and would only be used in corner cases. This gives us 81 total values.
- Then we evaluate the pairs trading at each of those values.
- Whichever training fraction produces the highest profit becomes the optimal training fraction.

2. Test
We will reuse the same stocks that we did to test evaluatePairsTrading(), but this time also include the AMD & NVDA stock pair.

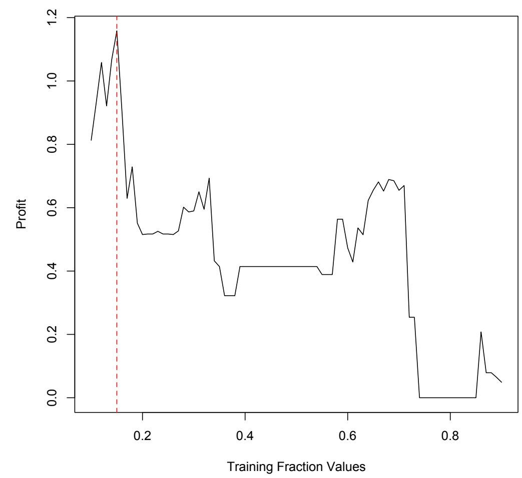
A. Graphics
AMD and NVDA between 2011 and 2021. The correlation is 0.91613. The default net profit is 31.3118.

The optimal training fraction is 0.5. This is the same as the default training fraction! This may hint that highly correlated stocks actually tend to do well with a 0.5 training fraction.

The net profit with optimal training fraction is 31.3118, the same as the default.



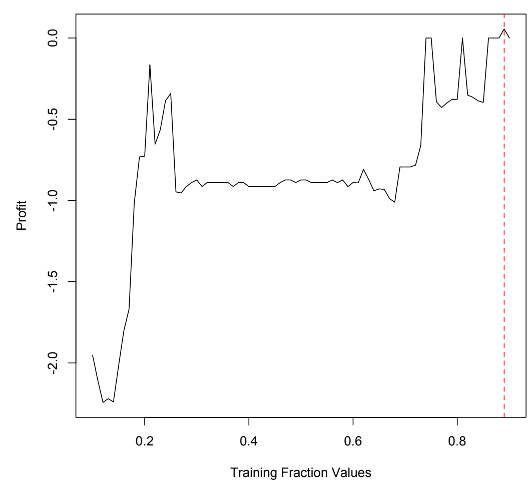
B. Positive
SONY and MCD between 2010 and 2020. The correlation is 0.74291. The default net profit is 0.41422.



The optimal training fraction is 0.15. This is totally different from the last positive stock pair. However, these stocks are also a lot less correlated. From the graph we can tell that less time training tends to be better for this pair.

The net profit with optimal training fraction is 1.158292, more than twice that of the default.

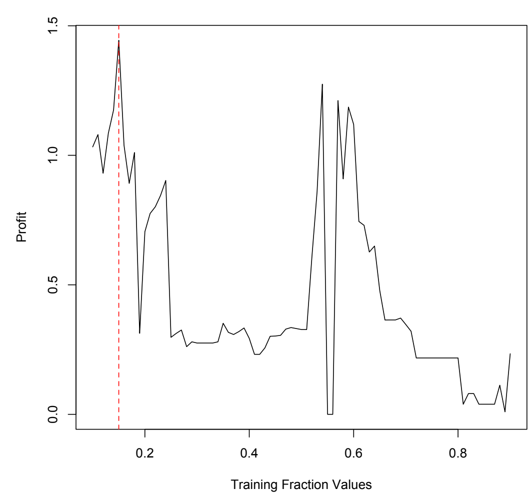
C. None
F and MCD between 2010 and 2020. The correlation is -0.13118. The default net profit is -0.87359.



The optimal training fraction is 0.89. This is higher than the previous two. No correlation means we have very little information about how the stocks will actually do. Thus, spending the most time possible getting information about the data produces the best results.

The net profit with optimal training fraction is 0.055223. Most of the training fractions for this stock pair do not turn a profit, so this is is impressive.

D. Negative
F and GLD between 2010 and 2020. The correlation is -0.47680. The default net profit is 0.32740.



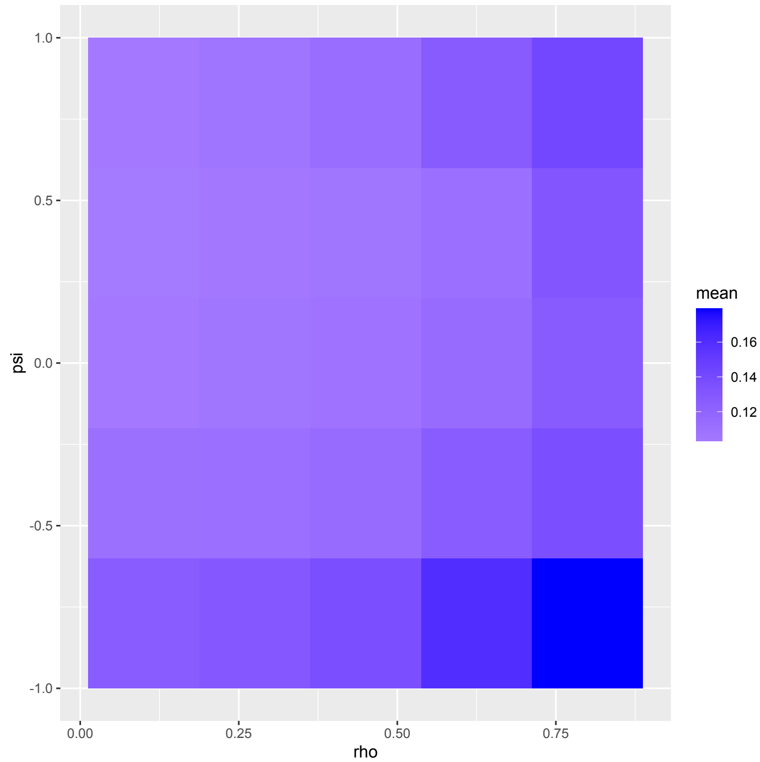
The optimal training fraction is 0.15. This is the same as the positive stock pair. Both are mildly correlated stock pairs. From these two examples, it seems that stock pairs that have a medium stock correlation (0.4 - 0.8) benefit from a lower training fraction.

The new profit with optimal training fraction is 1.44389. That is a little less than 5 times the original.

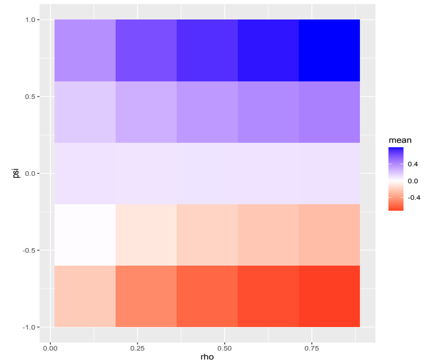
3. Training Fraction Heatmap

[Rho is between 0.1 and 0.8. Psi is between -0.8 and 0.8.]

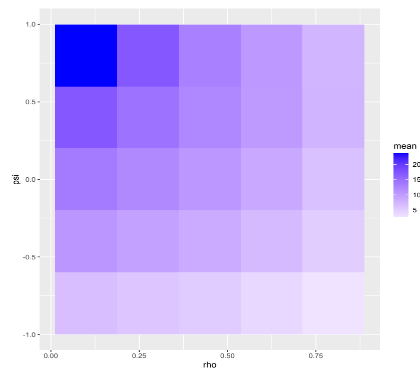
a. Training Fraction Heatmap



b. Mean Correlation Heatmap



c. Mean Profit Heatmap



4. Conclusion

In this simulation, most of the stock pairs tend to prefer a lower training fraction (0.1 - 0.2). But these heatmaps are more accurate for observing general trends rather than drawing exact numbers, since there are other variables, such as sigma1 and sigma2.

Stock pairs that have the highest profits (low rho, high psi) tend to have a lower training fraction. Stock pairs that have a highly negative correlations and the lowest profits tend to have a higher training fraction.
Stock pairs that have highest correlations (high rho, high psi) tend to have training fractions that are about in the middle between higher and lower training fractions.

In summary, the more profit a stock pair produces, the lower the training fraction is.

The next step would be to figure out a way to categorize stock pairs into rho/psi values. Then we could predict what training fraction would be best to use.
One solution would be to map correlation values onto training fractions using corresponding rho/psi values from the heatmaps. During the training phase of the data, we could find out what training fraction is mapped to by their training phase correlation.