Time Dependent Method for Short Term (daily) Trading with potential extensions to HFT

Aviv Shalit [a]
James Guito [b]
Charles Godin [c]

[a] Department of Particle Physics and Astrophysics, Weizmann Institute of Science, 234 Herzl Street, Rehovot 7610001, Israel

[a] *shalitaviv@gmail.com*
[b] *Jamesguito@gmail.com*
[c] *charles.godin@gmail.com*

Abstract

During the "Development" stage, we focused on practically attempting to find the best representation of Regression-based, High-Frequency Trading (HFT) strategy through the bisection of a) filtering several "slow" algorithms (in the sense of trade execution), then b) back-testing in a "fast" environment. We review the concept of statistical arbitrage and discuss on how it is used in HFT. We have thematically narrowed down into pair-trading, complemented by a staking algorithm, in this paper, few past works are discussed to cast perspective. In what follows, we present a short introduction to the field then we discuss the theoretical framework of statistical arbitrage and pair-trading. Later, we dive into the methodology. We include the techniques and tools we used to conduct our research. We then show some preliminary results or our research. Lastly, we discuss and conclude our results and findings.

Keywords: Negative correlation, pair-trading, staking algorithm, statistical arbitrage, high-frequency trading, regression

I. Introduction

A state of the art in High Frequency Trading means ever-accelerating reactions based on focused, one-pass algorithms, specialized hardware, ever-faster networks and trading platform colocation.

The leading firms in High Frequency Trading publish little, protect their intellectual property aggressively, and update their systems often as the arbitrage loophole is diminishing: being a slightly slower than everyone else does not make you a little less profitable than the average – it makes you loss-making. In another word, only the few top performers come out profitable in the pursuit of searching, taking advantages, and closing the arbitrage opportunities.

These techniques tend to have three stages: firstly, an offline analysis of relevant trading data to find arbitrage (such as bid-ask crossing [6]) or statistical patterns. Secondly, an implementation of said analyses on a live system – where speed really matters. Finally, live performance monitoring, as most of these approaches have a finite life expectancy – once others have caught up with the approach and/or the speed. In particular, the final step is essential as the back-testing on historical data does not capture the adaptation of market participants.

Our effort seeks to inject more fundamental, low-frequency fundamental analysis into rapid trading to compensate for our lack of rapid trading system.

2. Theoretical Framework

Based on multiple sources, pair trading was firstly and formally introduced by Gerry Bamberger, Nunzio Tartaglia, and David Shaw of Morgan Stanley in early 1980's, continued by Peter Muller, in 1985 post founder departures. Given the lopsided risk/reward, the strategy quickly rose to fame then this popularity caused the arbitrage opportunity to wane by the early 2000's. The disciples then branched out into other types of statistical arbitrage (the formal nomenclature of pair-trading) by developing their own Black Boxes.

With all the technologies, developments, and talent in quantitative methods, pair-trading can seem rudimentary, the proven key to success revolves around the technical and fundamental interpretation of the pair relationship. Synonymously, the interpretation is far more important than the metric(s) chosen to measure the strength of the relationships between the 2 poles (either 2 underlying securities, or groups of securities), which signifies a level of "predictability" and "repeatability" of the identified trends.

Apart from the core mean-reversion algorithm, our team would like to strongly consider how the impact of staking algorithms incrementally benefits the trading performances. For this, we would explore, test, and choose amongst a set of narrowed choices.

We hypothesize that the combinations of the approaches delineated above could potentially be more effective in the search of arbitrage through statistical lenses.

3. Methodology

Our goal is uniquely to improve risk management in HFT systems. As known, statistical arbitrage [1] is one of the key features in HFT. HFT portfolios take advantage from market fluctuations to profit in short periods of time. Risk management in HFT systems is usually related to a downwards trend of assets value. When an asset's value is reduced below some threshold, the HFT system stops the trading of that asset in order to prevent losses. Our goal, is to try to predict this downwards behavior in advance, before reaching the low threshold. Even though pair trading has been studied by many researchers [2], usually the analysis has been done on similar time periods. For an example of such pair correlation, see *Figure 1* below:

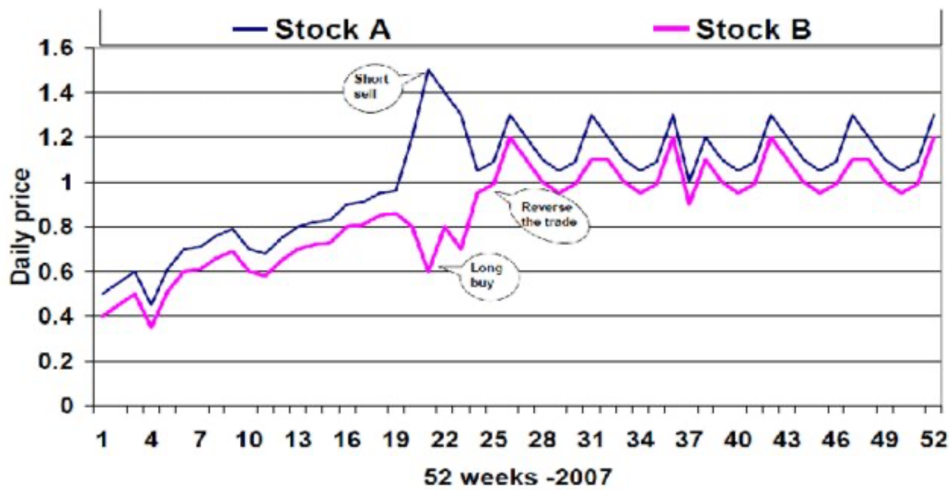


Figure 1. *An example of negative correlation between Stock A and Stock B on the same time period [7]*

In the example above, asset A and asset B are highly correlated with each other. In particular, one should expect that the Pearson correlation **over the entire time** will be close to 1:

$$\rho_{Y_t^A, Y_t^B} \approx 1 \qquad (1)$$

Moreover, if we do not include the runway behavior from week 19 to week 25 as seen in *Fig. 1*, the overall correlation will be even positively stronger.

Unfortunately, in practice, things are much more complications. For example, consider the following two function:

$$Y_t^A = f(t) \qquad for\ \ 0 \le t \le T \qquad (2)$$

$$Y_t^B = \begin{cases} f(t) & for\ \ 0 \le t \le \frac{T}{2} \\ -f(t) & for\ \frac{T}{2} \le t \le T \end{cases} \qquad (3)$$

where $f(t)$ is a simple continuous function of t with $f\left(\frac{T}{2}\right) = 0$.
It is obvious that knowing $\frac{T}{2}$ and $Y_t^A = f(t)$ will reveal the entire information about Y_t^B. However, traditional correlation function over the entire time (or sample path) will result in no correlation

$$\rho_{Y_t^A, Y_t^B} \approx 0 \qquad (4)$$

Moreover, the market is affected by short time period microeconomics and sudden events. This may lead to an increasing trend in some assets or a decreasing trend in other assets. However, the time of the impact of such events may change from different assets.

Our proposal aims to address pair trading, and in particular negative correlation, on **different time periods**. This strategy involves the decision on selling or buying assets based on the behavior of a different asset. Our task is to find

$$\Delta t \equiv t - t' \qquad (5)$$

In which two assets exhibit inverse/negative correlation between one another, i.e.

$$\rho_{Y_t^1, Y_{t'}^2} \approx -1 \qquad (6)$$

Where Y_t^1 is the value of one asset at time t , and $Y_{t'}^2$ is the value of second asset at could be a different time t'.

The Algorithm

We have created a full automatic system, given an input of any two paired-assets, will download the historical data and compute the difference time period from a decreasing trend in one asset to a decreasing trend in another asset.

The algorithm is established as follows:

1) The code download the historical data and save Closing price of the assets of interest. Then it creates a smooth version of the Closing price of each stock by

taking the avarage of 9 days, using the function: .rolling(9, min_periods=1, center=True).mean().
Lastly, it calculates the return of each stock at any given day after the coarse graining (smoothing the data)[1]

2) We find the downwards trends of asset A, given a sharp return threshold for decreasing trend and minimal days of decreasing behavior.

3) We look for increasing trend in asset B, that happened **before** the decreasing trend in asset A and are with minimal correlation. Moreover, we keep only such trend in both B and A that are both strongest negatively correlated and exhibit increasing trend above some return threshold.

4) We repeat the same logic of step (3) above only for increasing trends in B that happened **after** the decreasing trend in asset A

5) We plot the result of $\left(\Delta t_{before}\right)_i \equiv t^A_{ith\ decreasing\ trend} - t^B_{ith\ increasing\ before}$ and the result of $\left(\Delta t_{after}\right)_i \equiv t^B_{ith\ increasing\ after} - t^A_{ith\ decreasing\ trend}$. where i represent the date (in days) of decreasing trend.

6) We fit a model called model_before to find a trend in Δt_{before} as a function of date (in days). Moreover, we fit a model called model_after to find a trend in Δt_{after} as a function of date (in days).

7) If there is a meaningful fit with high enough statistical score (statistical significant) we

4. Results

The following example exhibit such phenomena of inverse pair-trading:

Consider the SodaStream International Ltd. stock historical data from January 1st, 2013 – January 1st, 2018 [1], and also consider the Keurig Dr Pepper Inc. stock historical data from January 1st, 2013 – January 1st, 2018 [2].

We fundamentally expect the stock of SodaStream company to have an inverse correlation with the stock of the Keurig Dr Pepper Inc. This is because SodaStream provides an alternative to soft drinks with lots of sugar as sold by the Dr Pepper company. SodaStream even advertise itself as a competitor to companies such as Dr Pepper, calling the public to consume its products over the products of soft drink companies. Therefore, even if SodaStream International Ltd. is a much smaller company by sales and market capitalization than Keurig Dr Pepper Inc., a rise on SodaStream's stock value may indicate a general trend in the public and a potential decrease in Dr Pepper's stock value. However, an increasing value of SodaStream's stock might only be the start of the trend, thus it may take some more time before it will influence Dr Pepper's stock value.
The algorithm works as follows:

1. We Download and smooth the Data. We find the time of decreasing trends in KDP below some return threshold as shown in the figures below:

1. We thank Professor Saad Zaman for pointing out that basing our analysis in return-space over price-space would yield better results for detecting increasing/decreasing trends.

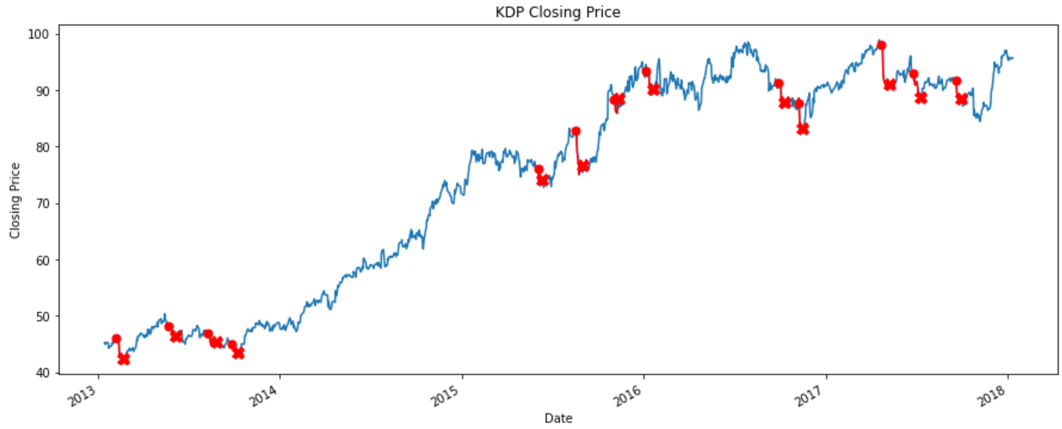


Figure 1. *The daily closing price of Keurig Dr Pepper Inc. with detected times of decreasing trends. In red are the time periods of decreasing trend in KDP*

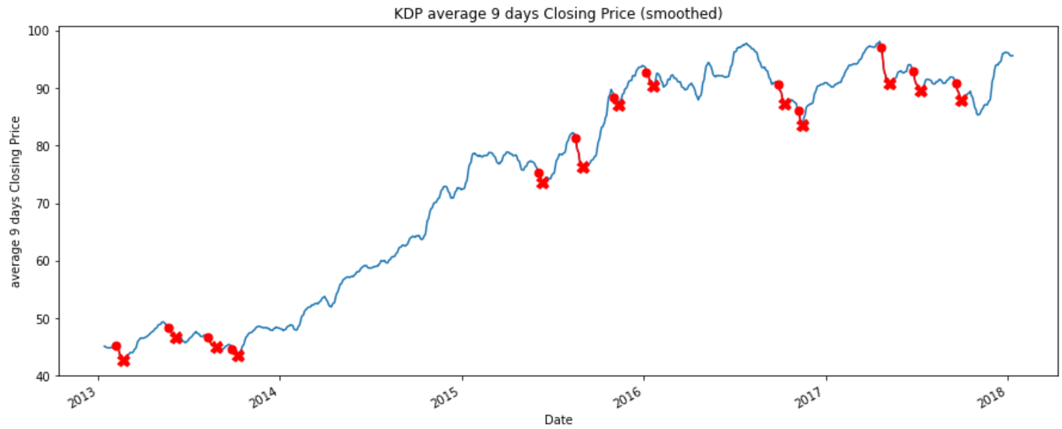


Figure 2. *The daily smoothed price of Keurig Dr Pepper Inc. with detected times of decreasing trends. In red are the time periods of decreasing trend in KDP*

2. We, show a significant negative correlation between Y_t^1 - SodaStream's stock at time t **before** t', and $Y_{t'}^2$ - Dr Pepper's stock at time t'. We repeat the same analysis for time t **after** t'

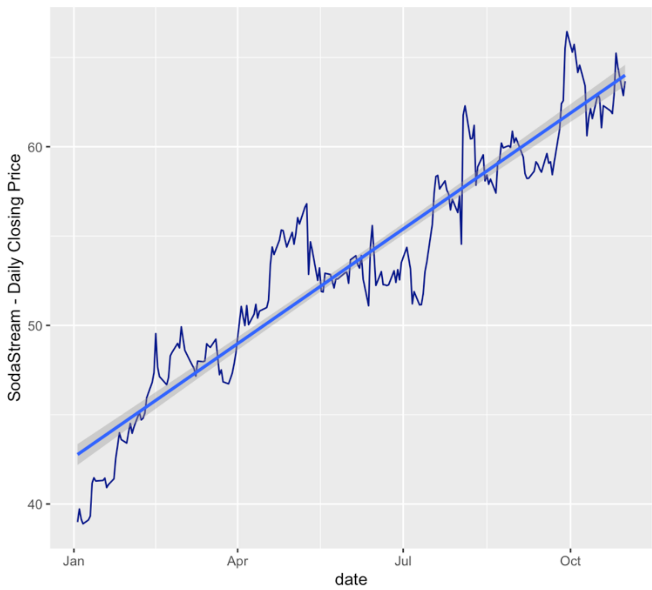
Figure 3. *The daily price of SodaStream LTD. with detected times of increasing trends. In red are the time periods of decreasing trend in KDP. In green are the time periods of increasing trend in SODA* ***before*** *the decrease in KDP. In purple are the time periods of increasing trend in SODA* ***after*** *the decrease in KDP.*

3. We only pick a period of time for which Y_t^1 (SODA) shows substantial growth, above some return threshold.

This can be simply done by linear regression model which shows a positive coefficient between time and the stoke value

$$Y_t^1 = \alpha \cdot t + c \qquad (2)$$

Where:
Y_t^1 - SodaStream's stock closed value at time t
t – the specific time of evaluation
α, c – constants of the fitting




```
Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-4.8503 -1.5322 -0.1924  1.1468  5.1742

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 42.683293   0.298372  143.05   <2e-16 ***
x            0.101733   0.002452   41.49   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '

Residual standard error: 2.154 on 208 degrees of freedom
Multiple R-squared:  0.8922,	Adjusted R-squared:  0.8917
F-statistic:  1721 on 1 and 208 DF,  p-value: < 2.2e-16
```

Figure 4. *The daily closing price of SodaStream stock, including a linear regression analysis which evaluates the growth of the stock price.*

For example if $Y_t^1 = 1.3 * Y_{t_{start}}^1$, i.e. SodaStream's stock price increased by 30% since a starting time frame and the analysis predicts that it will increase. At this point the trend in SodaStream's stock hint that Dr Pepper's stock will start to fall after $(\Delta t_{before})_{i=t}$ days.

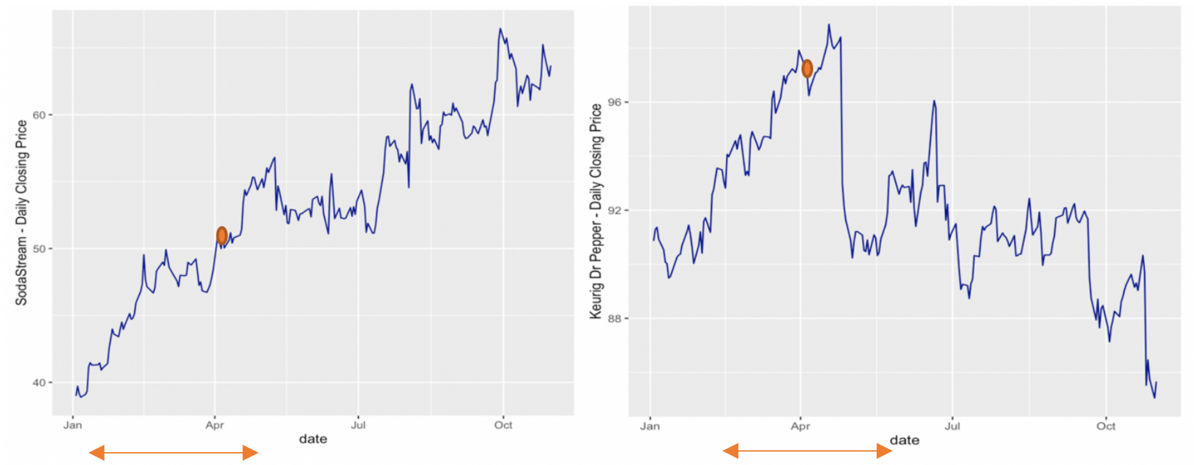
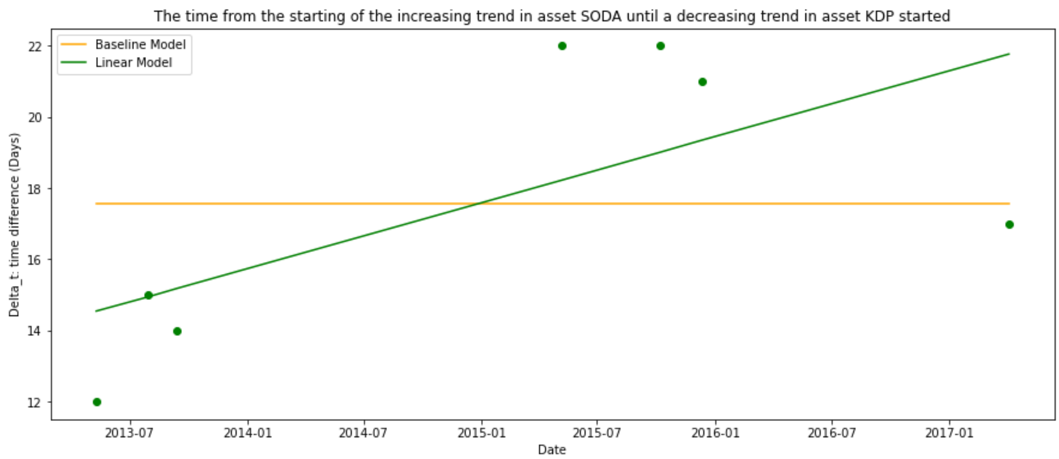


Figure 5. *On the left the 2017 daily closing price of SodaStream stock including the critical point where the price increased by* 30% *since we collected the data. On the right the 2017 daily closing price of Dr Pepper Inc. stock including the critical point at which we should sell before the predicted fall of the stock's price.*

4. We plot both $\left(\Delta t_{before}\right)_i \equiv t^{A}_{ith\ decreasing\ trend} - t^{B}_{ith\ increasing\ before}$ and $\left(\Delta t_{after}\right)_i \equiv t^{B}_{ith\ increasing\ after} - t^{A}_{ith\ decreasing\ trend}$, and try to find and fit trend. As seen by the figures below:

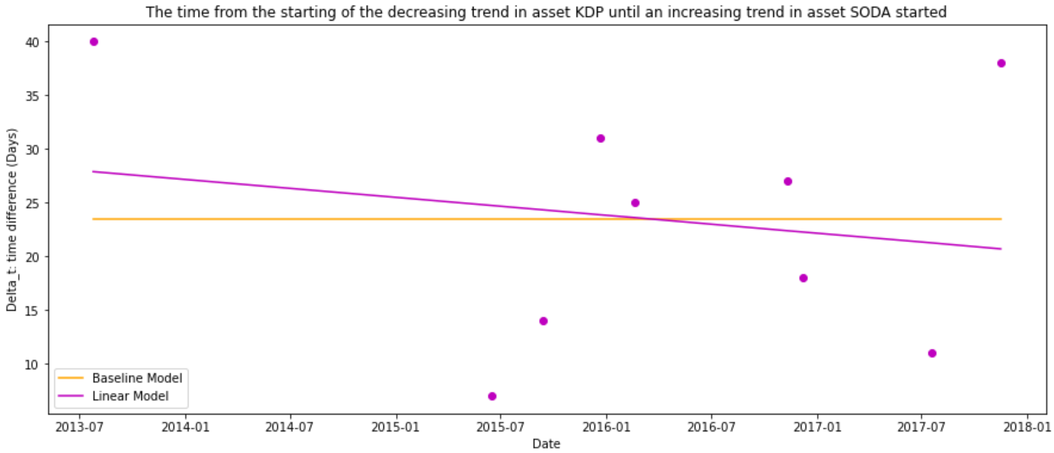
$\left(\Delta t^{SODA\rightarrow KDP}_{before}\right) =$ `13.9425(days) + 0.0074 * (current date in days)`




```
Baseline MAE: 3.5102040816326534
Training MAE for the model_before: 2.43
the model accuracy score is:  0.4434360153640923
```

Figure 6. *The time from the starting of the increasing trend in SODA until a decreasing trend in asset KDP started.*

$\left(\Delta t^{KDP\rightarrow SODA}_{after}\right) =$ `28.7707(days) + -0.0066 * (current date in days)`




```
Baseline MAE: 9.728395061728396
Training MAE for the model_after: 9.46
the model accuracy score is:  0.03385195469714941
```

Figure 7. *The time from the starting of the decreasing trend in KDP until an increasing trend in SODA started.*

The model performed better on predicting $\left(\Delta t_{after}^{KDP\rightarrow SODA}\right)$, with $R^2 = 0.443$. Indeed, this is still far from a good fit, of $R^2 \rightarrow 1$, but it still signifies a trend in the impact time as a function of dates (even if the linear fit is far from perfect). The code is completely automatic and can run on any two assets given that their historical data can be found on Yahoo website. We leave a more comprehensive search for a better fit of two different assets for future investigations. In principle such two assets should exists such that the extrapolation of there impact times $\left(\Delta t_{after}^{A\rightarrow B}\right)$ and $\left(\Delta t_{after}^{B\rightarrow A}\right)$ results in a better accuracy.

4. Discussion

As stated earlier, our main goal is to employ the defined pair-trading strategy in the lens of risk management. With the 3-step, delineated in details above, our team has been able to come to a level of "predictability" of Dr. Pepper's stock (NASDAQ: KDP), avoiding a ~13% drawdown for the next 150 days. Although our approach seems fundamental (in reference to the sampled linear algorithm) and unstable (in reference to the error-prone choosing of hyperparameter m [gradient]), we believe that this is a unique angle that builds on and adds value to trades built on pre-generated pairs (either fundamentally or statistically).

The current world of HFT mostly focuses on profit-making strategies with risk management being a novel utilization. Hence, our team is hopeful that we cast a different light on a popular topic with this paper.

5. Conclusions

In summary, our team starts with the idea of developing a set of HFT-sequel algorithms in a rather "slow" environment. Once successful, we would back-test and monitor the performances in the "fast" environment, proxying the high-frequency characteristics. In specific, our goal is to employ the algorithm from the risk management perspective, avoiding permanent loss of capital.

Acknowledgments

We would like to thank WQU for the opportunity to study and gain the skills to analyze financial markets for a better world of data science.

References

[1] Andrew W. Lo (2010). Hedge Funds: An Analytic Perspective (Revised and expanded ed.). Princeton University Press. p. 260. ISBN 978-0-691-14598-3.

[2] *Rad, Hossein; Low, Rand Kwong Yew; Faff, Robert (2016-04-27). "The profitability of pairs trading strategies: distance, cointegration and copula methods". Quantitative Finance. 16 (10): 1541–1558.* doi:10.1080/14697688.2016.1164337

[3] SodaStream International Ltd. (SODA) 2017-2018 financial data. https://finance.yahoo.com/quote/SODA/history?period1=1483228800&period2=1509494400&interval=1d&filter=history&frequency=1d

[4] Keurig Dr Pepper Inc. (KDP) 2017-2018 financial data. https://finance.yahoo.com/quote/KDP/history?period1=1483228800&period2=1509494400&interval=1d&filter=history&frequency=1d

[5] Morgan Stanley and the Birth of Statistical Arbitrage. https://www.risklatte.xyz/Articles/History/H01.php

[6] Irene Aldridge (2013). High-Frequency Trading (2nd edition). Wiley. ISBN 978-111-834-3500.

[7] Al-Naymat, Ghazi. (2013). Mining Pairs-Trading Patterns: A Framework. International Journal of Database Theory and Application. 6. 19-28. 10.14257/ijdta.2013.6.6.02.