Università degli Studi di Torino

Department of Economics, Social Studies, Applied Mathematics and Statistics
SCHOOL OF MANAGEMENT AND ECONOMICS

Degree Course in Quantitative Finance and Insurance

Master Degree Thesis

Exploring Trading Strategies in Cryptocurrency Markets: Statistical Arbitrage, and Directional Strategies

Rapporteur:
Lorenzo Schoenleber

Co-Rapporteur:
Alberto Turigliatto

Candidate:
Emanuele Fogliati

ID number: 858191

ACADEMIC YEAR 2022-2023

Acknowledgments

This thesis marks the culmination of a two-year master's journey in Quantitative Finance, and for this reason, I would like to express my gratitude to all those people who have contributed to my growth, supported me, believed in me, and cared for me. A special thank you goes to my professor and advisor, Lorenzo Schoenleber, who guided me throughout this thesis, as well as to my professor and co-advisor, Alberto Turigliatto, for their invaluable support and guidance in bringing this thesis to fruition.

Abstract

This thesis aims to analyze, implement, and optimize two statistical arbitrage strategies. Specifically, it focuses on the mean-reverting strategy and the pairs trading strategy. Additionally, each strategy is examined for potential optimization opportunities. Finally, a method is explored to reduce overall transaction costs, thereby maximizing profitability.

Contents

List of Tables

List of Figures

1 Introduction

1.1 General Overview

In the dynamic world of financial trading, traders employ a diverse range of strategies to capitalize on pricing discrepancies between assets and generate profits. One such strategy that stands out is arbitrage, a technique that leverages price differences to achieve theoretically riskless gains. At its core, pure arbitrage involves seizing the opportunity presented by price disparities for the exact same security across different exchanges.

For instance, let us consider an example of pure arbitrage. Assume a trader notices that a share of stock X is priced at 25,75 on the Amsterdam exchange, while the same share is being sold for 26,00 on the Frankfurt exchange. Recognizing the price discrepancy, the trader swiftly purchases the stock on the Amsterdam exchange and sells it on the Frankfurt exchange. Since the shares of X are interchangeable, that is, they are the exact same financial instrument, this trade results in a flat position that is a state where the trader has no exposure to a particular asset, eliminating any associated risk.

However, as trading strategies evolved, a concept known as statistical arbitrage emerged. Unlike pure arbitrage, which focuses on price differences across different exchanges, statistical arbitrage aims to profit from relative statistical mispricing between assets within the same market. This strategies leverages statistical analysis and modelling techniques to identify pairs of assets with historically correlated price movements. Two examples of statistical arbitrage strategies that are analysed in this thesis are pairs trading and mean reversion.

Originally developed by Nunzio Tartaglia and his team at Morgan Stanley in the 1980s [1], pairs trading emerged as a profitable strategy. Tartaglia's team, comprised of mathematicians, physicists, and computer scientists, collaborated to create automated trading systems. Notably, David Shaw was among the esteemed members of the team. As knowledge of pairs trading spread throughout the financial industry, traders recognized the potential of using statistical models and methods to exploit relative mispricing while maintaining market neutrality, that is don't having an exposure to market directions, leading to the concept of statistical arbitrage. Algorithmic traders embraced pairs trading due to its straightforward nature, making it a popular strategy since the mid-1980s.

Pairs trading is a strategy employed to maximize profits by exploiting pricing discrepancies between two or more assets or portfolios that exhibit a strong historical correlation. This approach involves the identification of a pair of assets or portfolios and vigilant monitoring of the spread or price differential between them. When the spread deviates from its historical range, traders strategically establish positions, opting to go long on the underperforming asset while simultaneously shorting the outperforming asset. The ultimate objective is to capitalize on the eventual convergence of prices.

At the core of pairs trading lies the task of identifying two instruments, such as assets or portfolios, that have exhibited historically correlated price movements. Traders meticulously observe the spread, which reflects the disparity in prices between the selected instruments. When the spread surpasses a predetermined threshold, traders take decisive action by initiating a long position in the underperforming asset and a short position in the outperforming asset. Profits are derived from the subsequent convergence of prices between the two assets, and the positions are ultimately closed.

Pairs trading operates on the principle of relative pricing, which refers to the comparison of prices between two or more similar securities. Relative pricing is the practice of evaluating the prices of securities in relation to each other rather than focusing on their individual absolute values. This approach considers that securities with similar characteristics should be priced similarly.

In the context of pairs trading, the concept of relative pricing becomes crucial. Instead of attempting to determine the true or intrinsic value of a security, traders exploit discrepancies in relative pricing. When the relative prices between two similar securities diverge, one can be considered overpriced while the other is considered underpriced, or potentially both.

While the general rule in trading suggests selling overvalued securities and purchasing undervalued ones, pairs trading takes a different approach. It places significant importance on relative pricing, enabling traders to bypass the challenge of determining the true value of a security.

To understand the mechanism, let us suppose we have two highly regarded financial instruments like Coca-Cola and Pepsi. Given the nature of their businesses, it is not difficult to imagine that the prices of these two companies would rise and fall with a certain correlation, indicating that the movements of the two securities should be closely related to each other.

Assume historically Pepsi's value is three times that of Coca-Cola. At time t_0, Coca-Cola's price is 50 and Pepsi's price is 150. At t_1, Coca-Cola's price increases to 55, while Pepsi's price remains at 150. If our assumption holds true, either Coca-Cola is overvalued or Pepsi is undervalued. To profit regardless of the situation, we sell three shares of Coca-Cola and buy one share of Pepsi. There are three possible scenarios for profit. First, if Pepsi is undervalued, Coca-Cola's price remains at 55 and Pepsi aligns with the 3:1 ratio, increasing to 165, resulting in a profit of +15 from Pepsi and +0 from Coca-Cola. Second, if Coca-Cola is overvalued, Pepsi's price remains unchanged while Coca-Cola's price decreases to 50, resulting in a profit of +15 again. Finally, in a hybrid scenario, both stocks return to a 3:1 ratio, such as Coca-Cola at 52.5 and Pepsi at 157.5, leading to a profit of +15.

The second strategy that this thesis puts its focus on is Mean reversion trading. This is a trading strategy that focuses on short-term price fluctuations and the observation that prices tend to deviate from their short-term mean values before eventually reverting back to them. Traders employing this strategy aim to identify overextended price levels, whether above or below the mean, and take positions in anticipation of a price correction.

This strategy involves a directional approach, as traders anticipate short-term price deviations and subsequent reversion to the mean. By analyzing statistical indicators, traders can gauge the magnitude of the deviation and assess the likelihood of a price reversion. They rely on these indicators to identify temporary price discrepancies and capitalize on them.

When the price deviates above the mean, mean reversion traders view it as overextended and expect a potential correction back towards the mean. In such cases, they may choose to sell the stock, betting that the price will eventually revert back to the mean level. On the other hand, when the price deviates below the mean, traders see it as an opportunity to buy the stock at a relatively lower price, anticipating that the price will eventually revert back towards the mean.

Consider an example where the stock's price deviates above the mean. If the price rises from \$50 to \$60, mean reversion traders might decide to sell the stock at \$60, anticipating that it will eventually revert back towards the mean of \$50. Once the price returns to the mean value, they close their position, locking in a profit.

On the other hand, if the stock's price deviates below the mean, let us say from \$50 to \$40, mean reversion traders might see it as an opportunity to buy the stock at \$40. Their expectation is that the price will revert back towards the mean of \$50. Once the price reaches the mean level, they close their position, capitalizing on the price reversion and generating a profit.

To implement mean reversion trading, traders use statistical analysis and indicators to identify deviations from the short-term mean. They aim to profit from these temporary price discrepancies by entering positions and closing them when the price returns to the mean level. This strategy allows them to generate profits by taking advantage of short-term price fluctuations.

In summary, pairs trading and mean reversion trading are two distinct strategies that offer unique approaches to generating profits in financial markets.

Pairs trading is a strategic approach that capitalizes on historical correlations between assets and exploits pricing discrepancies. It adopts a market-neutral perspective, aiming to generate profits regardless of the overall market direction. By focusing on relative pricing, pairs trading identifies and takes advantage of price disparities between similar securities.

On the other hand, mean reversion trading is a strategy that exploits short-term price fluctuations by identifying deviations from the mean and anticipating a reversion back towards it. Unlike pairs trading, mean reversion trading focuses on taking a specific directional view on the market. Traders analyze statistical indicators to identify overextended price levels and position themselves in anticipation of a price correction.

Both pairs trading and mean reversion trading offer distinct methodologies for

traders to navigate financial markets. Pairs trading emphasizes historical correlations and relative pricing, while mean reversion trading focuses on short-term price fluctuations and statistical indicators. By understanding the principles and nuances of each strategy, traders can diversify their approaches and potentially enhance their ability to generate consistent profits in dynamic market conditions.
It is crucial to provide a final clarification regarding the strategies encompassed by statistical arbitrage, as mentioned earlier. These strategies aim to capitalize on relative statistical mispricing among assets within the same market. When analyzing pairs trading strategies, it becomes evident that the two assets prices involved are those of two distinct financial instruments (such as price of Coca-Cola and price of Pepsi). However, it is worth noting that the mean-reverting strategy, which may utilize a single financial instrument, can also be categorized under the umbrella of statistical arbitrage. In this case, the price discrepancies occur not between two different assets but between an asset and a statistic derived from the asset itself, such as its short-term mean, which will be also a price (e.g. the average price of the last 20 minutes). This average value effectively functions as a the price of the second asset in the context of the strategy.

1.2 Literature

In the past two decades, numerous researchers have developed various statistical arbitrage strategies that differ in their behaviour and nature. The common goal among these strategies, as implied by the term statistical arbitrage, is to identify profitable strategies based on statistical analysis.[2]
The fundamental logic underlying these strategies remains unchanged. Firstly, two instruments are selected based on their historical price movements, focusing on assets that have exhibited a strong correlation. Then, the spread between their prices is monitored during the trading period. Finally, if the prices diverge, a short position is taken on the asset that has performed better, while the asset that has performed worse is bought. Although the core logic is consistent, there are several elements that may vary across different pairs trading strategies.

One key distinction in the broader realm of statistical arbitrage strategies lies in the selection of pairs to be traded. There are four main macro categories for pairs selection:

1. Univariate (one vs one): In this approach, trading is limited to two individual financial instruments, such as Coca-Cola versus Pepsi.

2. Quasi-Multivariate (one vs many): This approach involves trading one financial instrument against a portfolio of instruments. For example, Apple may be traded against a linear combination of Microsoft and Google.

3. Multivariate (many vs many): Here, trading involves two portfolios.

4. Individual: This approach utilizes a single financial instrument against a statistical model, such as comparing the price of Amazon against its short or long-term mean.

Another important distinction among approaches to statistical arbitrage lies in their spread construction methods. Currently, there are at least five identified categories, but new ideas may emerge in the future. These categories include:

1. Distance Approach[3]: This approach employs simple non-parametric distance metrics to measure the spread between two assets (or their returns). Trading signals are generated based on a set of predefined rules. This approach will be the one used in this thesis for the Pairs trading strategy.

2. Cointegration Approach[4]: This approach requires a stronger co-movement between the pairs being traded and typically involves conducting cointegration tests.

3. Time Series Approach[5]: This approach relies on forecasting or modelling the price behaviour. Trading signals are generated based on the difference between the model price and the actual price. This approach will be the one used in this thesis for the Mean Reverting strategy.

4. Stochastic Control Approach[6]: The core idea behind this approach is to identify optimal trading rules or weights for elements within a pair, taking into account the statistical properties of price or spread processes.

5. Other Approaches[7]: These are novel approaches that incorporate Machine Learning, Copula approach, Principal Component Analysis, and other methodologies.

1.3 Data

This thesis aims to analyze two trading strategies: Pairs Trading and Mean Reverting. When constructing a trading strategy, it is crucial to decide the data on which the strategy will be based from the beginning. This includes determining the financial instruments to use, selecting a data provider, deciding on the timeframe, and specifying the data to be utilized in the strategy.

To begin, this thesis utilizes cryptocurrency data available for trading on the Binance exchange. The choice of cryptocurrency data was made due to its accessibility from various websites and freely downloadable software. It is worth noting that although the data can be downloaded directly from Binance via its specific API, limitations exist regarding the amount of data that can be downloaded simultaneously. Consequently, for backtesting (a technique used to evaluate the performance of a trading strategy using historical data) and analyzing large quantities of data, Binance does not offer an efficient method for downloading extensive datasets. Consequently, reliance was placed on an external free software known as StrategyQuant, which facilitates the downloading of substantial volumes of data directly from the Binance exchange.

Regarding the timeframe for collecting cryptocurrency price data, one must consider the type of strategy to be implemented and the desired frequency of trading operations. It is recalled that the timeframe represents the interval between collecting one data point and the next. For instance, if a daily timeframe is utilized, the collected data will display the daily prices of the specific instrument in a dataframe. Thus, a shorter timeframe will result in more data points being collected, enabling more trading operations within the same time period. Additionally, a shorter timeframe implies a faster duration for each individual trade, i.e. the time between opening and closing the trade. In this thesis, a relatively fast timeframe of one minute was chosen, resulting in 1440 observations per day (the number of minutes in a day). The data was collected from the first minute of January 1, 2022, until the last minute of December 31, 2022, amounting to a total of 525600 data points (1440x365). The cryptocurrencies used in this analysis are those tradable on Binance within this timeframe, specifically 285 cryptocurrencies.

Lastly, it is important to specify the data that was downloaded and utilized in the actual implementation of the trading strategy. In this thesis, only the minute-by-minute closing prices of each individual cryptocurrency were employed. It is essential to note that this closing price does not necessarily represent the price at which the cryptocurrency can always be bought or sold. To consider bid and ask prices, one would have had to download such data. However, for simplicity and for the purposes of this thesis, it was assumed that buying and selling could always occur at the closing price.

After downloading all the data from StrategyQuant for each cryptocurrency, the result is a CSV file that can be easily imported into Python as a dataframe with 285 columns (one for each cryptocurrency) and 525600 rows (representing each minute in the year 2022).

1.4 Thesis Objective

The main objective of this thesis is to comprehensively study and analyze two prominent strategies in Statistical Arbitrage. To achieve this, two different strategies were developed from scratch, the first is a Pairs Trading strategy and the second is a Mean Reverting strategy; the development was made by using the Python programming language. Initially, the code was written in a less efficient but highly detailed manner to facilitate the in-depth study and analysis of each strategy individually. This approach provided the opportunity to thoroughly examine the strategies and verify their correctness. Additionally, it allowed for the analysis of each cryptocurrency one at a time, enabling a comprehensive understanding of their individual performance.

Subsequently, the focus shifted towards finding the most optimal parameters for these strategies. In order to further enhance their effectiveness, various Machine Learning techniques were employed. This integration of Machine Learning methodologies aimed to fine-tune the strategies and improve their overall performance.

Furthermore, the strategy codes were later rewritten in a more efficient and vectorized manner. This modification facilitated the simultaneous implementation of both strategies across the wide range of cryptocurrencies available on the Binance Exchange. By leveraging vectorization, the execution time of the strategies was significantly reduced, ensuring faster and more efficient analysis.

The obtained results were analyzed in detail, taking into consideration the transaction costs involved. Evaluating the profitability of these strategies while accounting for transaction costs was another key objective of this thesis. By incorporating these costs into the analysis, a more realistic assessment of the strategies' effectiveness and potential profitability was achieved.

Lastly, a final code was developed with a different logic that allowed for the simultaneous implementation of both strategies. This novel approach aimed to reduce transaction costs even further, thereby increasing the overall profitability.

1.5 Thesis Structure

This thesis, following this introductory chapter, consists of a chapter entitled "Research Methodology" that aims to provide the basic theoretical knowledge necessary to understand this thesis. Subsequently, there are two chapters explaining the two mean reverting strategies in Chapter 3, aptly titled "Mean Reverting Strategy," and then the chapter on pairs trading strategy titled "Pairs Trading Strategy." Following these chapters, the chapter entitled "The Portfolio Methodology" is developed, which aims to illustrate a methodology that allows us to reduce transaction costs on one hand and increase profits on the other. Lastly, there is a chapter containing the conclusions.

2 Research Methodology

In this chapter, we will explain the statistical and econometric concepts underlying the strategies of statistical arbitrage, such as pairs trading and mean reverting. The purpose of this explanation is not to provide a comprehensive coverage of the topic but rather to provide a level of detail that is relevant to this thesis. In particular, it is appropriate to provide a reminder of the key concepts on which these strategies are based. Specifically, we will start by explaining the concept of a stationary time series, leading to the concept of mean reversion, which forms the foundation of the strategies analyzed in this thesis.

A time series is a sequence of observations of a variable over time. Formally, it can be represented as a sequence of data: $x_1, x_2, ..., x_n$, where x_t represents the observed value at time t.

Analyzing a time series requires understanding its fundamental components. The first component is the trend, which represents the overall direction of growth or decline of the data over time. It can be modeled using regression techniques to estimate its trend over time, for example, through linear or non-linear regression. The second component is seasonality, which refers to periodic fluctuations that repeat at regular intervals of time. It can be identified using methods such as additive decomposition or Fourier decomposition. The third component is the cycle, which represents long-term fluctuations above or below the trend, generally with longer periods compared to seasonality. Finally, the residuals represent the part of the time series that is not explained by the trend, seasonality, and cycle, and contain the residual uncertainty or noise in the time series.

Understanding the components of a time series is crucial for selecting the appropriate model for analysis. For example, if the time series exhibits a trend, it may be necessary to apply smoothing techniques such as moving averages or exponential moving averages to remove it before proceeding with the analysis. If seasonality is present, specific models such as seasonal ARIMA or structural component models can be used. The presence of a cycle may require the use of models such as ARIMA with a cyclical component. The residuals, if present, can be analyzed to check for residual patterns or autocorrelation.

The goal of time series analysis is to extract information and make forecasts. Understanding the components of the time series provides a basis for selecting appropriate statistical models that capture the data structure over time. The choice of a suitable model depends on the specific characteristics of the time series and the objectives of the analysis, such as short-term forecasting or understanding underlying mechanisms.

At this point, it is important to define what a stationary process is. A stationary process is a type of stochastic process in which its statistical properties remain constant over time. Formally, a process X_t is stationary if it satisfies the following conditions:

1. Constant mean: The mean of the process, $E[X_t]$, is constant for every time instant t. This implies that there are no systematic trends or patterns of growth or decline over time.

2. Constant variance: The variance of the process, $Var(X_t)$, is constant for every time instant t. This indicates that the dispersion of the data around the mean remains stable over time.

3. Constant autocovariance: The autocovariance between two observations X_t and X_s depends only on the time lag $|t - s|$ and not on specific time points t and s. Formally, $Cov(X_t, X_s)$ depends only on $|t - s|$ and not on t and s individually. This implies that the dependence structure of the data is constant over time.

The stationarity of processes is important because it simplifies the analysis and modeling of time series. In particular, stationary processes allow for the application of traditional statistical methods, such as parameter estimation, construction of confidence intervals, and hypothesis testing, as their statistical properties remain constant over time. Additionally, many classical models, such as autoregression (AR) and moving average (MA), assume data stationarity.

However, in practice, many time series do not satisfy the conditions of stationarity. Trends, systematic variations, or structural changes over time can be observed. In such cases, more sophisticated approaches are necessary, such as data transformation or the use of specific models for handling non-stationarity, such as autoregressive integrated moving average (ARIMA) models or stochastic trend models (STL).

To assess the stationarity of a time series, various statistical tests are available. The Dickey-Fuller test and the Augmented Dickey-Fuller test are two common tests used to check the hypothesis of data stationarity. These tests compare the statistical properties of the process against the hypothesis of non-stationarity, allowing for determining whether a time series is stationary or requires the application of specific models for non-stationarity.

On the other hand, when working with non-stationary series, it is important to consider the difference between successive observations rather than the individual observations themselves. This approach is referred to as "first differencing," which corresponds to the notion of integration of order 1, denoted as I(1). Therefore, the use of an I(1) model indicates that we are dealing with a non-stationary time series, but we are analyzing the differences between successive observations to make the series stationary.

The idea behind I(1) models is that non-stationary time series can often be made stationary by applying the first difference. Once a stationary series is obtained, we can use appropriate models and analytical techniques for analysis and forecasting.

The use of I(1) models is particularly relevant when working with phenomena that exhibit trends or patterns of growth or decline over time. For example, if we have a series representing the price of a cryptocurrency over time, and this series shows a long-term upward or downward trend, an I(1) model may be suitable to address such non-stationary behavior.

These concepts just explained will be used in the construction of trading strategies. Specifically, if we know that the price movements of an asset follow a stationary process, then when the price deviates from the mean, we can profit by buying the asset when its price is below the mean or selling short when the price is above the mean. However, in practice, the price movements of an asset are rarely stationary.

What we can do is analyze the price returns, which often follow I(1) processes.
It is important to note that for studying time series of cryptocurrencies, a less "strict" concept of stationarity is sufficient, namely the concept of mean stationarity. This refers to a process with a constant mean over time but not necessarily constant variance and autocovariance.
In general, a stationary process is a statistical process in which its statistical properties remain constant over time. This means that the mean and variance of the process do not change with time. However, in a process with a constant mean but not constant variance and autocovariance, the variance and autocovariance of the process can vary over time.
More specifically, if a process has a constant mean over time, it means that the average value of the process observations remains the same regardless of when the measurements are taken. For example, if the mean of a process is 0, then the expected value of its observations will always be 0.
However, the variance of a process with a non-constant variance changes over time. This means that the dispersion of the process observations may be different depending on when the measurements are taken. For example, heteroscedasticity may occur, where the variance of the process increases or decreases over time.
Similarly, the autocovariance of a process with a non-constant autocovariance changes over time. Autocovariance measures the correlation between the process observations at different time intervals. If the autocovariance is not constant over time, it means that the correlation between the process observations can vary over time.
In conclusion, the concepts of stationary time series, stationary processes, and non-stationary processes provide a solid conceptual foundation for the development of statistical arbitrage strategies, particularly for pairs trading and mean-reverting strategies.
By using time series analysis, it is possible to identify patterns and dynamics in the prices of an asset over time. If the prices follow a stationary process, one can exploit deviations from the mean to profit from pairs trading. This strategy involves identifying two or more correlated assets and buying the asset that has become undervalued relative to their historical average, while short-selling the asset that has

become overvalued relative to their historical average. The objective is to capitalize on the hypothesis that prices tend to revert to their mean over time, thus generating a profit.

On the other hand, when prices follow a non-stationary process, as often happens in reality, a mean-reverting strategy can be applied. In this case, instead of seeking correlated pairs of assets, the goal is to identify a specific asset that deviates from its historical mean. If the price deviates from the mean, it can be assumed that there will be a return to the mean in the future, allowing for targeted trading actions. For example, if the price of an asset is below its mean, one might consider buying that asset, while if the price is above the mean, short-selling could be considered.

To conclude this chapter, there is one last concept that is worth mentioning: correlation. Correlation, from a theoretical standpoint, is a statistical measure that indicates the relationship or association between two or more variables. It represents the linear dependency between variables and measures the strength and direction of the relationship between them. Correlation can take a value between -1 and 1, where -1 represents a perfect negative correlation, 1 indicates a perfect positive correlation, and 0 indicates no correlation.

To better understand the concept of correlation, let us take the example of two cryptocurrencies: Bitcoin (BTC) and Ethereum (ETH). Suppose we have a historical series of data representing the daily prices of both cryptocurrencies over a period of time.

To calculate the correlation between BTC and ETH, various methods can be used, including the Pearson correlation coefficient. This coefficient measures the linear correlation between the two variables. Suppose that after calculating the correlation coefficient between BTC and ETH, we obtain a value of 0.8. This indicates a strong positive correlation between the two cryptocurrencies. It means that when the price of Bitcoin increases, it is likely that the price of Ethereum will also increase proportionally, and vice versa. This positive correlation suggests that the two cryptocurrencies follow a similar trend in the market.

On the other hand, if the correlation coefficient had been -0.3, it would indicate a weak negative correlation between BTC and ETH. In this case, when the price of Bitcoin increases, the price of Ethereum tends to decrease slightly. A negative correlation suggests that the two cryptocurrencies may exhibit opposite behaviors in the market.

It is important to note that correlation does not necessarily imply a cause-and-effect relationship between variables. It may simply indicate a statistical relationship without one variable directly influencing the other.

Understanding the correlation between cryptocurrencies or any set of financial assets can be useful for various purposes. For example, for investors looking to create a diversified portfolio, correlation can help identify assets that move independently or have a negative correlation, thereby reducing the overall portfolio risk. Similarly, for traders using pair trading or arbitrage strategies, correlation can be a tool to spot profit opportunities by exploiting price differences between correlated assets.

3 Mean reverting strategy

3.1 Linking Theory to Practice

In this chapter, the mean-reverting strategy will be examined, implemented, and optimized. This strategy, as the name suggests, is based on the concept of mean reversion, particularly on the concept of mean stationarity of a time series. This concept has been analyzed in detail in the "Research Methodology" chapter, but in brief, it refers to the property of a time series to maintain a constant mean in the long run, even if the variance and autocorrelation may not be constant. Therefore, if this property is present in a particular time series, we can infer that if it deviates from the mean, it will tend to return to it in the near future due to this characteristic.

Transferring this theoretical concept to the world of trading, if we consider the prices of cryptocurrencies as a time series, we might consider going long or short in anticipation of a return to the mean. However, in general, cryptocurrency prices do not have the characteristic of being a stationary or mean-reverting time series (figure 1). For this reason, instead of working with the price series itself, we will work with the time series of price returns, i.e., a type I(1) time series that, instead, turns out to be stationary. (figura 2). The return time series, denoted as $r(t)$, will be constructed as follows:

$$r(t) = \frac{p_t - p_{t-1}}{p_{t-1}} \tag{1}$$

where p_t represents the price of the cryptocurrency at time t.

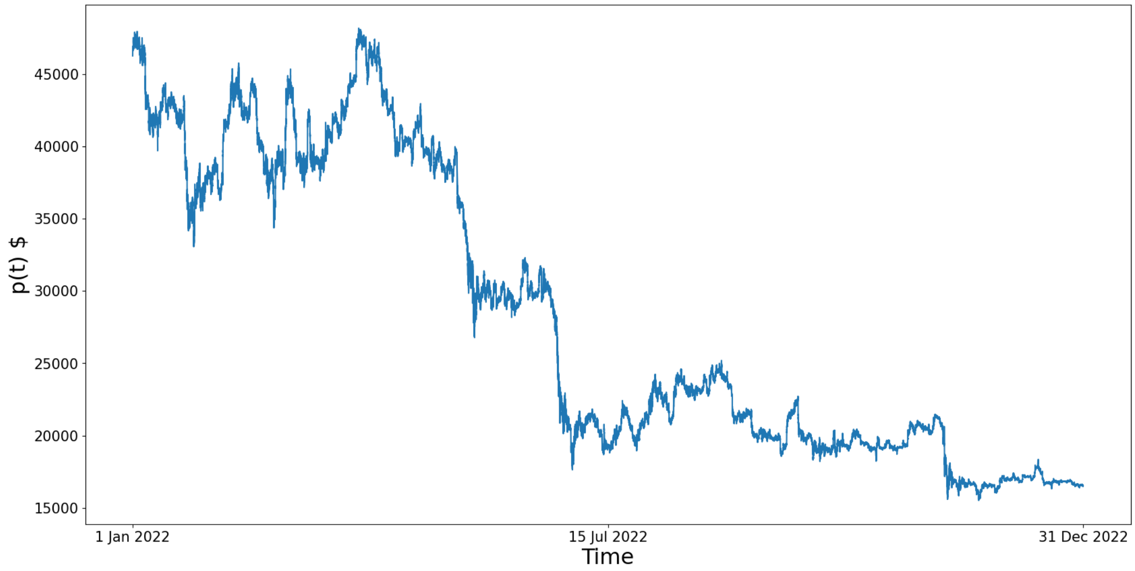

Figure 1: Bitcoin prices, clearly not a stationary time series

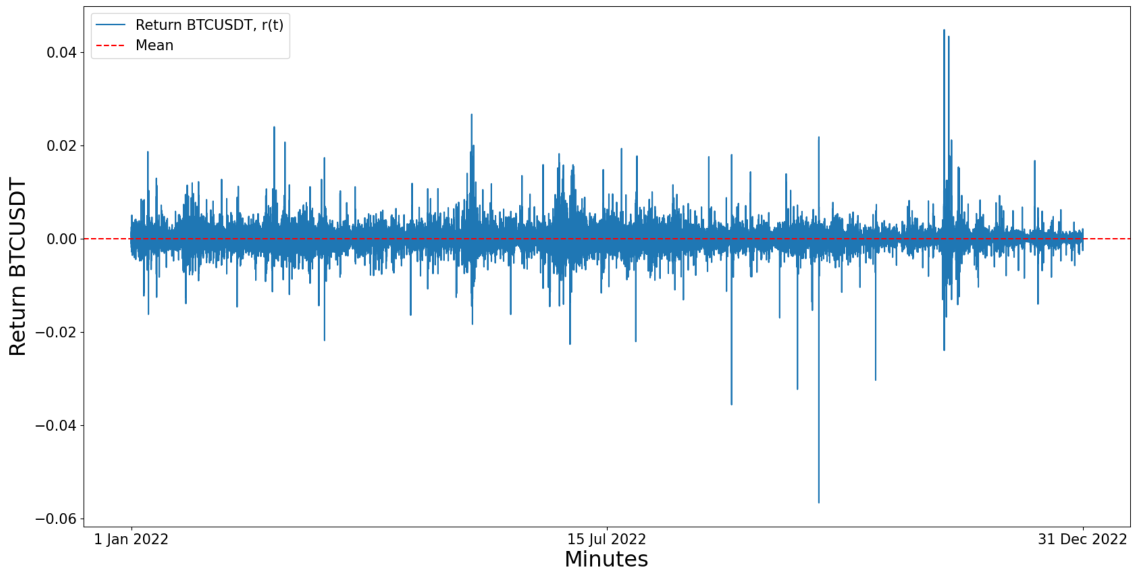

Figure 2: Bitcoin return, clearly a mean stationary time series

Therefore having an mean stationary time series, the concept behind the strategy will be as follows if the returns at time t, $r(t)$, are lower than the average returns, then we will buy the underlying asset with the expectation that the returns will be positive in the near future because only with positive returns will the time series $r(t)$ converge back to the mean. If we expect positive returns, it implies that we expect the price of the cryptocurrency to rise in the near future, and for this reason, to profit from a potential price increase, we will buy the cryptocurrency in question.

On the other hand, if the returns at time t, $r(t)$, are higher than the average returns, then we will sell the underlying asset with the expectation that the returns will be negative in the near future because only with negative returns will the time series $r(t)$ converge back to the mean. If we expect negative returns, it implies that we expect the price of the cryptocurrency to decline in the near future, and for this reason, to profit from a potential price drop, we will sell the cryptocurrency in question.

At this point, it is important to specify that for all the 285 cryptocurrencies considered, the average returns in all cases can be safely approximated to zero. There are several studies that explain this concept and attribute it to the theoretical hypothesis of efficient markets. The result of the work in this thesis indicates that these average returns can be approximated to zero; however, this is not by virtue of the studies recalled earlier, but simply based on the results obtained by calculating the sample mean of the returns.

Before continuing, it is important to briefly introduce and explain what transaction costs are and how they are calculated, as they are always taken into account in this thesis. Transaction costs represent the expenses that a trader incurs when executing buy or sell operations of assets such as stocks or cryptocurrencies.

Transaction costs are usually expressed as a percentage of the transaction amount or as a fixed amount per transaction. Regarding the Binance exchange, the transaction costs are expressed as a percentage of the transaction amount, specifically 0.2%.

For example, let us say you want to buy \$1000 worth of Bitcoin with a transaction fees of 0.2%. In this case, the formula to calculate the transaction costs would be:

$$\text{Total transaction costs} = \text{Transaction amount} \times \text{Transaction fees}$$
$$\text{Total transaction costs} = \$1000 \times 0.002 = \$2$$

Therefore, you would need to pay \$2 as total transaction costs for buying \$1000 worth of Bitcoin. It's important to note that the same applies when closing the position. For instance, if you want to close the previous trade and the price of Bitcoin has doubled so now your position is worth \$2000, the transaction costs for closing the position would be $\$2000 \times 0.002 = \$4$. In total, \$6 would have been spent on transaction costs for executing this trade.

Another important element to consider when executing this strategy is to determine the threshold at which returns should deviate from the mean (zero) in order to enter a trade. This threshold will allow us to automatically decide whether it is opportune to enter a trade or not. One might think that every time returns are, for example, below zero, it is appropriate to open a long position. However, if we were to do that, we wouldn't always generate a profit due to transaction costs. Specifically, it's easy to imagine that the further returns deviate from the mean, the larger the price movement should be in order for returns to converge back to the mean. This can be illustrated with the following example: suppose the return time series indicates -0.1%, suggesting that we should enter a long position, anticipating a profit of almost 0.1% on the invested capital (e.g., \$1000). Once the trade is closed, the pre-tax profit (that is the profit made without taking into consideration the total transaction costs) would be around \$1 (1000 * 0.1%). However, assuming that the transaction fees for opening and closing the position are 0.07%, the costs to open the trade would amount to \$0.7 (1000 $\times$ 0.07%), and the costs to close the trade would be around \$0.7007 (1001 $\times$ 0.07%), resulting in a total of \$1.4007 in total transaction costs. This completely erodes the profits and leads to a loss of \$0.4007. Note that at the closure of the trade, the fees are calculated based on the final value in \$ of Bitcoin, assuming a profit of 0.1%: 1000 $\times$ (1 + 0.001) = \$1001.

Therefore, as we have just seen, it is important that entering the trade guarantees us a sufficiently large gain to make a profit despite transaction costs. In the specific case of this thesis, various values were considered, and the most promising one was to enter a position when the average returns deviated by at least one standard deviation of the returns calculated over the previous 20 observations. To do this, a new time series called "rolling_SD" was created, which tracks the rolling standard deviation over 20 periods. With this time series, it was possible to determine the so-called upper bound (UB) and lower bound (LB) that allow us to determine when to enter a trade. Therefore, UB = rolling_SD, and LB = -rolling_SD. In this way, if the returns are higher than the UB, we will enter a short position as we expect a downward correction in prices. Conversely, if the returns are lower than the LB, we will enter a long position as we expect an upward correction in prices.

3.2 The Basic Mean Reverting Strategy

Start by explaining the theoretical strategy presented by Hanqin Zhanga and Qing Zhang in their paper[8]. Later on, an optimization method will be illustrated. Consider one of the 285 cryptocurrencies available on Binance: 1INCHUSDT. The first thing we need to do is create the time series of returns $r(t)$ (blue), the upper bound (red), and the lower bound (green), as explained earlier and shown in figure 3.

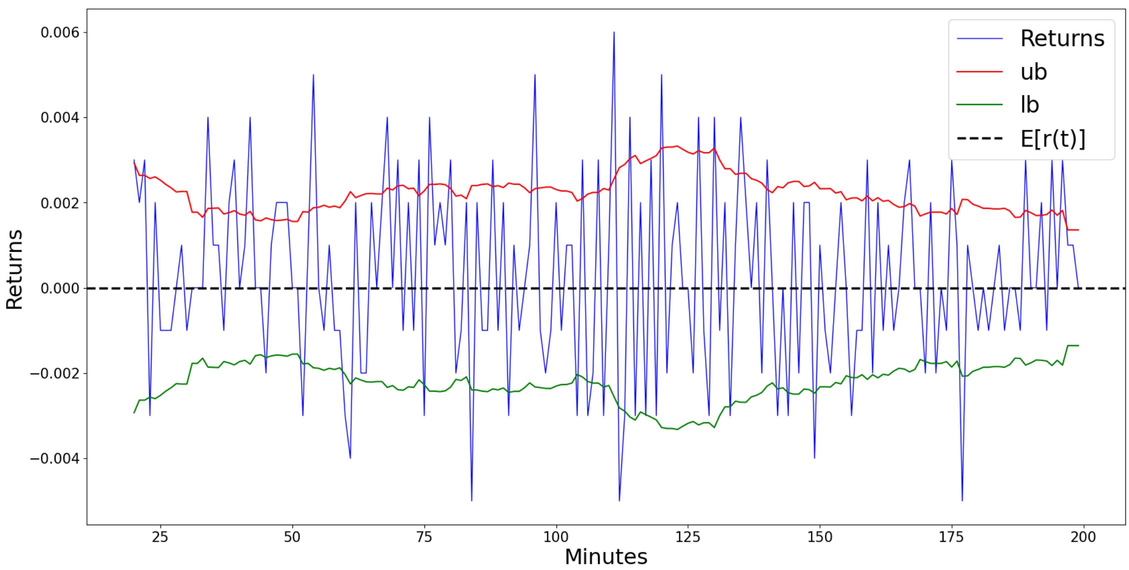

Figure 3: 1INCHUSDT returns (blue), upper bound (red), lower bound (green)

At this point, as soon as $r(t)$ crosses either the UB or the LB, we will enter a position. Referring to the image of figure 4, we notice that one of these situations has occurred, namely $r(t)$ ¡ LB at Date = 10034 (marked by the first green x in the image). At that moment, we have $r(t)$ = -0.005, and LB = -0.003605, so we can open a long position. To determine the quantity of the instrument to buy, we need two additional pieces of data: the first is to determine how much capital to invest per trade, for example, c = \$1000, and the second is to know the price at that moment, in this case, $p_{initial}$ = \$2.159. Therefore, we can calculate the quantity to buy as $qty = c/p_{initial} = \$1000/\$2.159 = 463.18$.

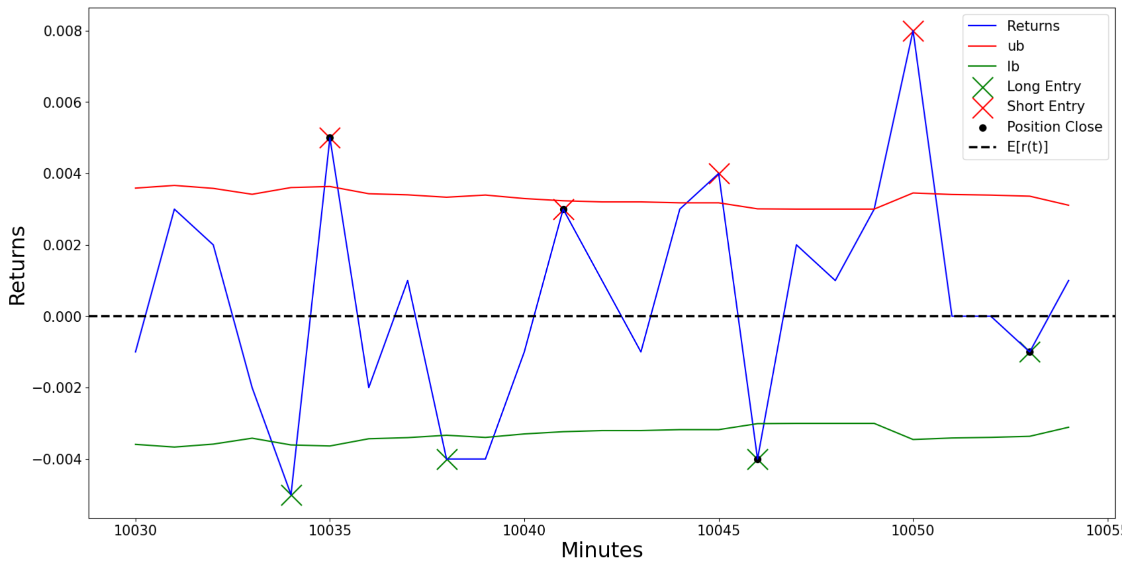

Figure 4: Trading signals 1INCHUSDT

So now we are in a long position with exactly 463.18 units of the cryptocurrency 1INCHUSDT, with a market value of \$1000. All that's left is to wait for $r(t)$ to return to or exceed the mean. In particular, we can already observe that in the next minute, $r(t) = 0.005 > 0$, so we can close the trade. At that moment, the price of 1INCHUSDT is $p = \$2.164$, and we can calculate the profit as follows:

$$profit = qty \times (p_{final} - p_{initial}) = 463.18 \times (\$2.164 - \$2.159) = \$2.32$$

Additionally, we also need to calculate the transaction costs for the trade. The transaction costs for the purchase were $TransactionCosts = 0.0002 \times \$1000 = \$0.2$, and for the sale, $TransactionCosts = 0.0002 \times (463.18 \times \$2.164) = \$0.2005$, resulting in a total transaction cost of $TotalTransactionCosts = \$0.2 + \$0.2005 = \$0.4005$; this slightly reduced the overall after-tax profits to gainaftertax = \$2.32 - \$0.4005 = \$1.9195. At this point, once the trade is closed, the same process restarts with the next observation, and as soon as $r(t)$ crosses either the UB or the LB, we will enter a new position.

In the following figure 5 let us show the final results by applying this strategy to the cryptocurrency 1INCHUSDT. It is possible to observe that the strategy on 1INCHUSDT has been a profitable one. Both considering and not considering transaction costs, we achieved a positive return at the end of the year. In the simulation, the following values were considered: Capital invested per trade c = \$1000 and Transaction costs tax = 0.2%

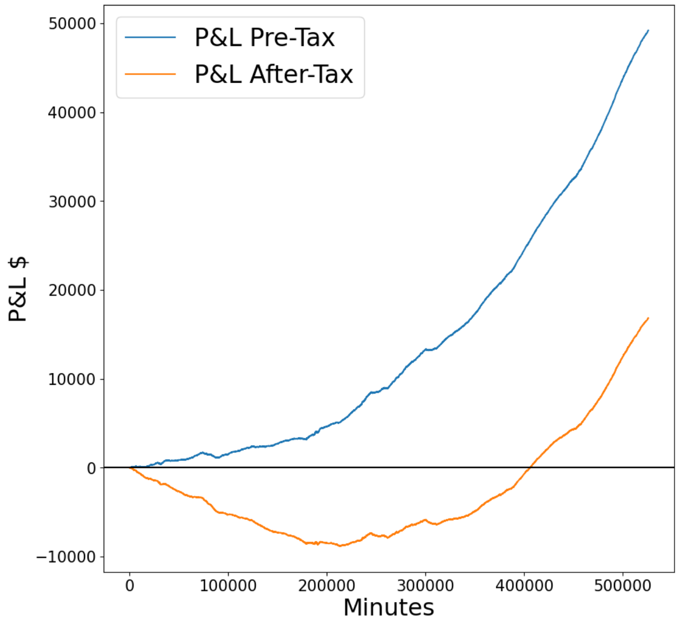

Figure 5: P&L 1INCHUSDT with the Basic Mean Reverting Strategy

The results obtained were as follows:

- P&L pre-tax = \$49192.90
- Total fees = \$32370.54
- P&L after-tax = \$16822.36
- #trades = 70324
- E[P&L per trade pre-tax] = \$0.70
- E[P&L per trade after-tax] = \$0.24

It is important to note that it is necessary to track both pre-tax and after-tax results. Although taxes are always an element to consider since they need to be paid to execute the trades, we will demonstrate a way to significantly reduce the overall cost resulting from fees at the end of the thesis.

3.3 Results of the Basic Strategy

At this point, the results obtained, although encouraging, concern only one of the cryptocurrencies available on Binance. Therefore, the overall results are presented below, where the basic mean-reverting strategy is applied to all 285 cryptocurrencies.

The graph shown in figure 6 shows the profit distributions for each cryptocurrency, both pre-tax and after-tax.

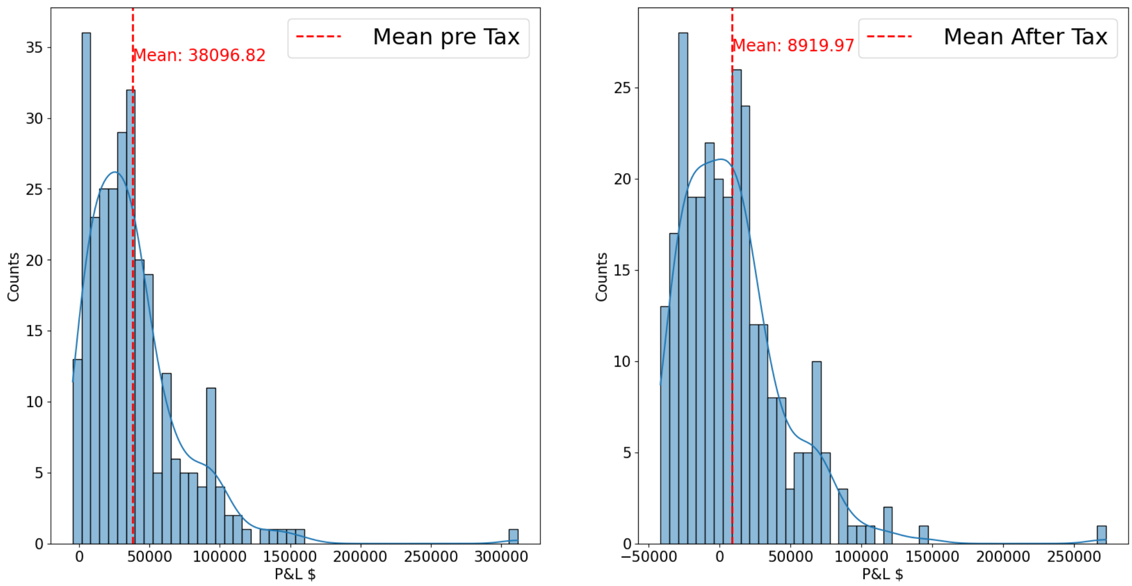

Figure 6: P&L per crypto Pre-Tax(left) and After-Tax (righy) with the Basic Strategy

As anticipated, the returns are significantly eroded after taxes, but the average remains positive even after-tax. Specifically $E[P\&L_{pre-tax}] = \$38096.82$ and $E[P\&L_{after-tax}] = \$8919.97$

Furthermore:

Basic Mean Reverting Strategy	**Pre-Tax**	**After-Tax**
# cryptocurrency with $P\&L > 0$	274	151
# cryptocurrency with $P\&L < 0$	11	134

Table 1: # crypto with > or < P&L Basic Mean Reverting Strategy

From Table 1 it is clear once again that the fees have a big impact on the profitability of the strategy.
In figure 7 the distributions of daily P&Ls are presented, followed by the cumulative profits over time on a daily basis (figure 8).

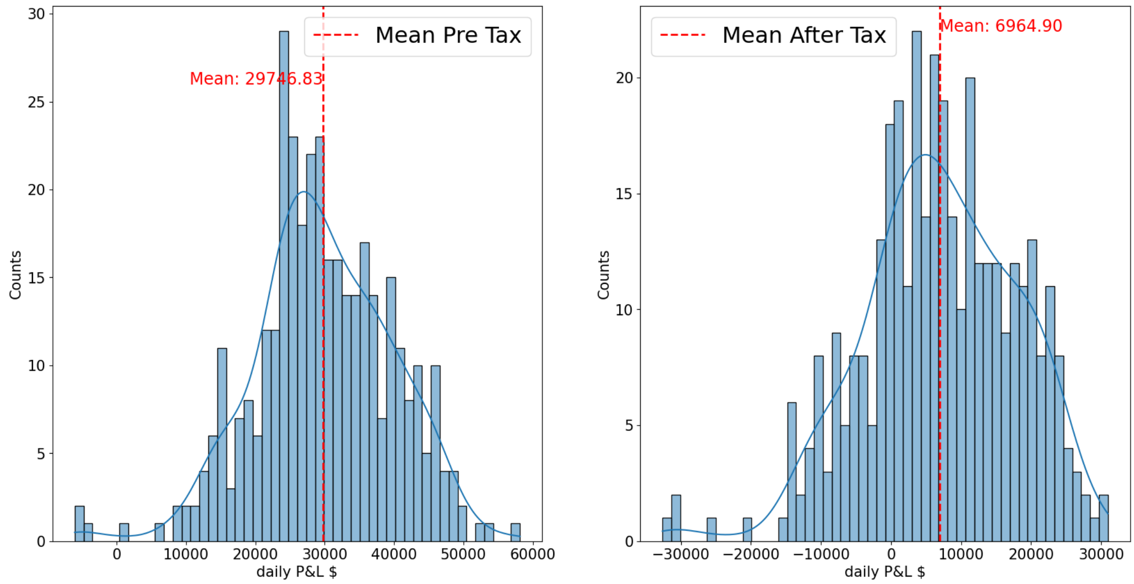

Figure 7: Daily P&L Pre-Tax(left) and After-Tax (right) with Basic Strategy

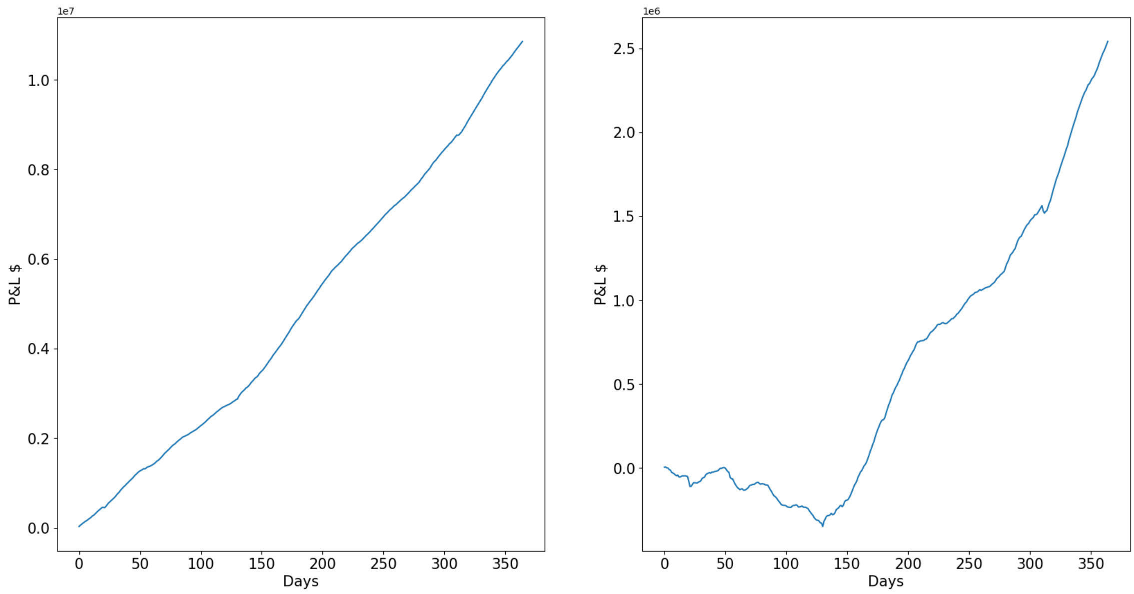

Figure 8: Cumulative P&L Pre-Tax(left) and After-Tax (right) with Basic Strategy

The final results are as follows:

Basic Mean Reverting Strategy	**Pre-Tax**	**After-Tax**
P&L Tot.	10857594.45	2542190.06
E[P&L per day]	29746.83	6964.9
# days with $P\&L > 0$	362	3
# days with $P\&L < 0$	274	91
Highest daily P&L	58153.92	-6058.82
Lowest daily P&L	30951.23	-32629.23

Table 2: Summary of the Basic Mean Reverting Strategy

3.4 Optimized Mean Reverting Strategy

In this section, the processes undertaken to optimize and make the mean-reverting strategy more profitable and robust are analyzed.
The first attempt made was to determine in advance which trades, even if they met the entry conditions, would result in a losing trade. For this purpose, a neural network was constructed, taking as input the trading volume in the previous minute, the values of Open, High, Low, Close from the previous minute, the values of UB and LB, and the P&L of that operation. The neural network's objective was to classify with a "1" if the trade should be executed or with a "0" if the trade should not be executed. The hope was that the network would predominantly output 1 for positive P&L and 0 for negative P&L. The neural network aimed to optimize the sum of P&Ls derived from each trade. Unfortunately, the result was very disappointing as, despite building several neural networks, all of them consistently outputted 1 for every trade. This indicates that the neural network failed to find any pattern among the given inputs. However, since the strategy was still profitable even with negative trades, the network simply returned 1 for every trade.
The second attempt to optimize the strategy, fortunately, was more successful. The basic idea is straightforward, and once understood, it is not surprising that implementing this improvement led to better results than before. Specifically, the basic mean-reverting strategy proposed by Hanqin Zhanga and Qing Zhang is designed in such a way that once a trade is closed, there is no immediate opportunity to enter a new trade in the opposite direction. Instead, before opening a new trade, one must wait until the next observation. Therefore, the fundamental idea of the optimized

strategy remains the same, but the possibility of opening a new trade is allowed directly after the closure of a previous trade on the same observation of the closure of previous trade.
To illustrate this with an example, consider the previously described scenario, where at the end of the operation, it was noted that $r(t) = 0.005 > 0$, and thus the trade could be closed. However, it was not mentioned that at that moment, the upper bound had a value of UB $= 0.003635$. Since $0.005 > 0.003635$, it would have been possible, but was not done previously, to immediately open a new short trade.

3.5 Results of the Optimized Strategy

Before presenting the overall results of testing the optimized strategy on all 285 cryptocurrencies available on the Binance exchange, let us first examine the results of the strategy on the 1INCHUSDT cryptocurrency to make an initial comparison (figure 9).

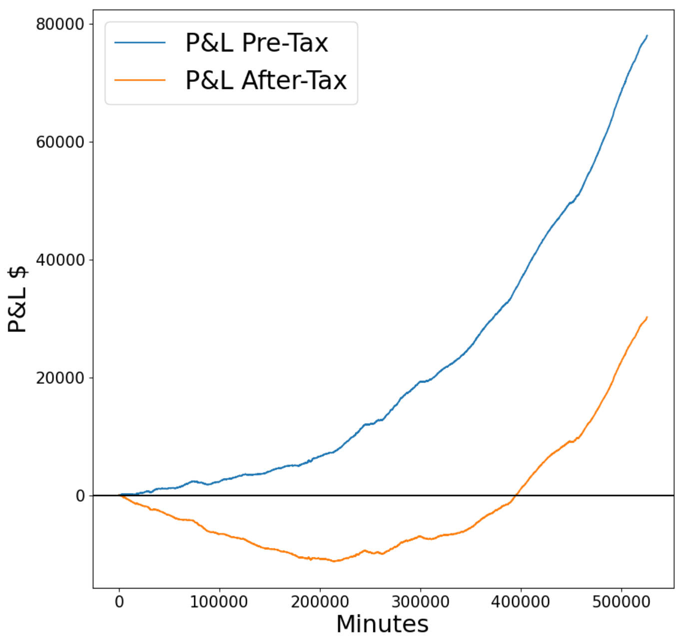

Figure 9: P&L 1INCHUSDT with the Optimized Mean Reverting Strategyy

At first glance, the results seem very similar to the once of the basic strategy (figure 5). However, let us compare the previous results with those of the optimized strategy:

	Basic Strategy	**Optimized Strategy**
P&L pre-tax	49192.90$	77937.50$
Tot. Fees	32370.54$	47739.54$
P&L after-tax	16822.36$	30197.96$
# trades	70324	102386
E[P&L per trade pre-tax]	0.7$	0.76$
E[P&L per trade after-tax]	0.24$	0.3$

Table 3: Comparison Basic Strategy and Optimized Strategy.

From the table 3, it is clear that practically all the metrics used to evaluate the strategy are better in the case of the optimized strategy. Now let us see the general results for all 285 cryptocurrencies (figure 10).

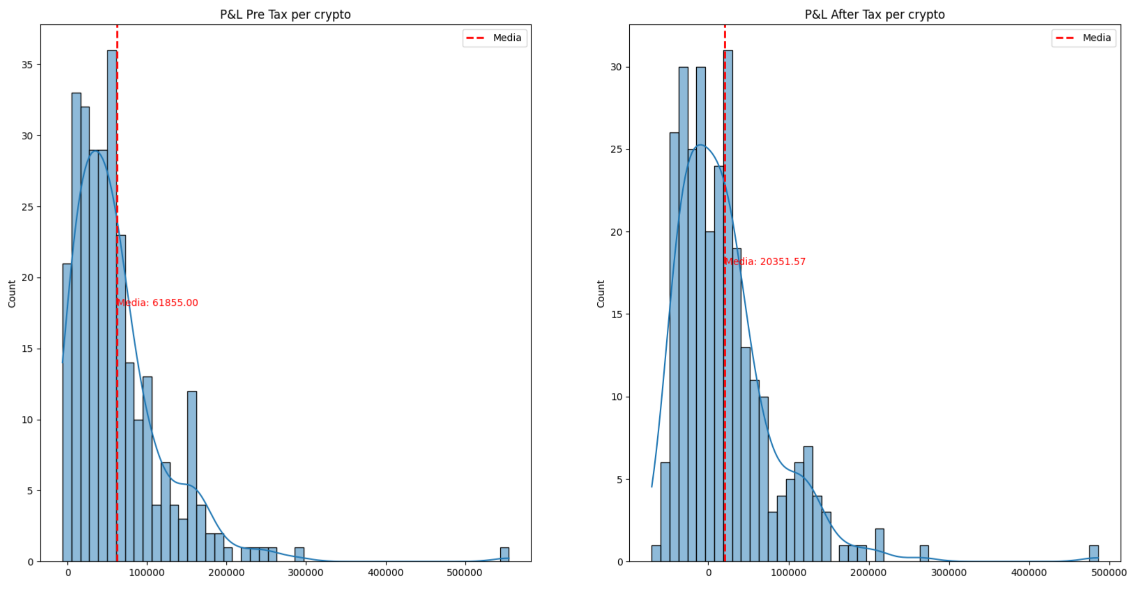

Figure 10: P&L Distribution per crypto Pre-Tax(left) and After-Tax (righy) with the Optimized Strategy

As expected, even with the optimized strategy, returns are significantly eroded by taxes. However, the average remains positive even after-tax. Specifically, comparing the base strategy with the optimized strategy, we have:

	Basic Strategy	Optimized Strategy
E[P&L pre-tax]	38096.82$	61855.00$
E[P&L after-tax]	8919.97$	20351.57$

Table 4: Comparison of E[P&L] between Basic Strategy and Optimized Strategy.

And then we can compare the results in Table 1 with the new one of the optimized strategy in Table 5:

Optimized Strategy	Pre-Tax	After-Tax
# cryptocurrency with $P\&L > 0$	275	161
# cryptocurrency with $P\&L < 0$	10	124

Table 5: # crypto with > or < P&L Optimized Mean Reverting Strategy.

Even from these metrics, we can see that the optimized strategy outperforms the base strategy across the board. Now let us examine the distributions of daily P&Ls and cumulative profits over time, with a daily frequency (figure 11 e 12).

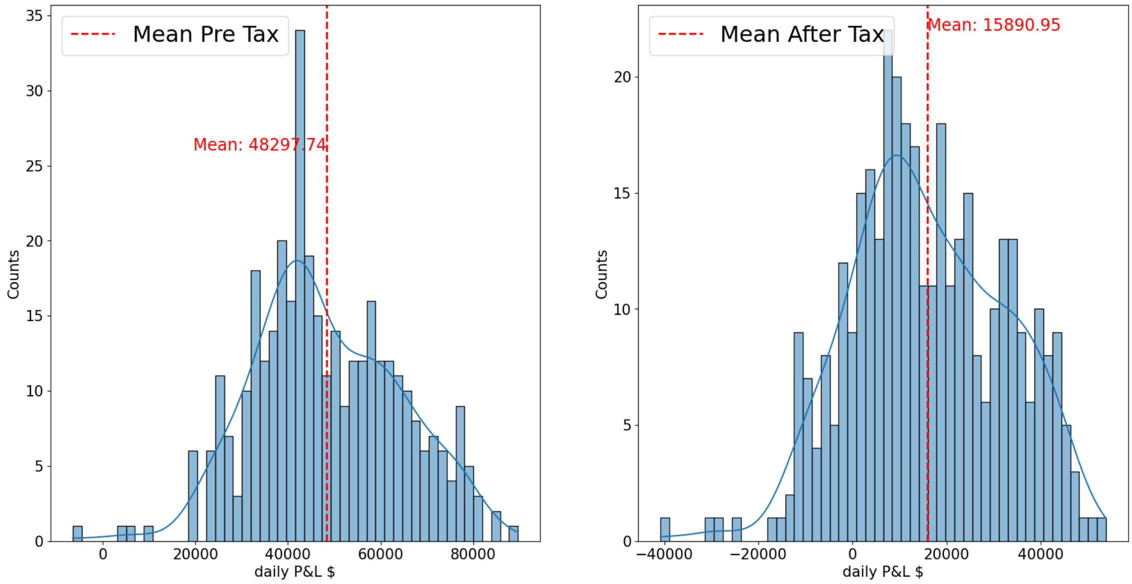

Figure 11: daily P&L Pre-Tax(left) and After-Tax (right) with Optimized Strategy

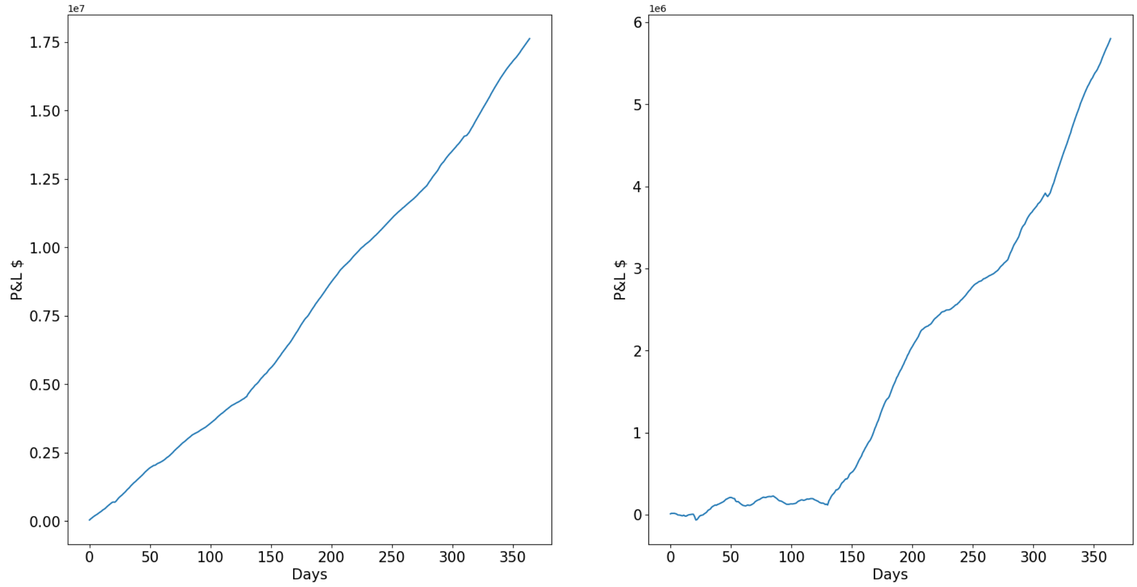

Figure 12: cumulative P&L Pre-Tax(left) and After-Tax (right) with Optimized Strategy

The final results compared to the base strategy are as follows:

Basic Strategy	**Pre-Tax**	**After-Tax**
P&L Tot.	10857594.45$	2542190.06$
E[P&L per day]	29746.83$	6964.9$
# days with P&L $>$ 0	362	274
# days with P&L $<$ 0	3	91
Highest daily P&L	58153.92$	-6058.82$
Lowest daily P&L	30951.23$	-32629.23$

Table 6: Comparison of P&L and trading metrics in Basic Strategy.

Optimized Strategy	**Pre-Tax**	**After-Tax**
P&L Tot.	17628674.27$	5800197.71$
E[P&L per day]	48297.74$	15890.95$
# days with P&L $>$ 0	364	307
# days with P&L $<$ 0	1	58
Highest daily P&L	89644.25$	53886.56$
Lowest daily P&L	-6458.51$	-40801.56$

Table 7: Comparison of P&L and trading metrics in Optimized Strategy.

Once again, we can see that almost all the metrics favor the optimized strategy.

3.6 Conclusions on the Mean Reverting

In conclusion, we can state that the mean-reverting strategy based on the statistical and econometric concept of mean reversion in a mean-stationary time series is generally a profitable strategy, whether or not it is optimized as presented earlier. Another conclusion is that using machine learning models to determine whether each individual trade should be executed or not is useless. However, machine learning can still be implemented in other ways. For example, we could explore using machine learning to identify which cryptocurrencies, as a whole, are profitable during a specific historical period and which ones should be avoided for trading.

4 Pairs trading strategy

4.1 Linking Theory to Practice

In this chapter, the pairs trading strategy will be examined, implemented and optimized. As the name suggests, this strategy does not rely on trading with a single cryptocurrency at a time, like the mean-reverting strategy, but rather on a pair. The concept is similar to the mean reverting strategy, although slightly more complex because each trade involves both a long position on one cryptocurrency and a short position on the other. Therefore, while the mean-reverting strategy was directional, as it involved either going long or short, and expected a specific price movement of the underlying asset to generate profits, the pairs trading strategy does not have a directional bias. Profit can be generated regardless of the movement of the two cryptocurrencies. In this strategy, it is crucial to adopt a different perspective, recognizing that an operation involves more than just two separate trades on two cryptocurrencies. Instead, it is a combined operation where both long and short positions are simultaneously taken on the two pairs. As a result, the profit or loss of the operation is not solely determined by individual trades but rather by the cumulative sum of profits and/or losses from both trades combined. This understanding is essential for a comprehensive analysis of the strategy. Thus, in this case, a trade can be considered successful whether gains are obtained from both the long and short positions or if one cryptocurrency yields a greater profit compared to the loss on the other, as the overall profit outweighs the loss, resulting in a positive balance. At this point, the natural question arises: which pairs can be selected and why? The pairs under consideration should have a certain relationship, meaning they should exhibit a high statistical correlation. This strategy relies on simultaneously taking long and short positions on two different cryptocurrencies. If these cryptocurrencies have a correlation close to zero, we are essentially betting that one will rise in price while the other falls. However, if the cryptocurrencies are highly correlated, we can expect them to move in almost the same direction. Assuming, for the sake of argument, that we find two perfectly correlated cryptocurrencies with a correlation coefficient of 1, going long and short on them would completely offset the profits from one with the losses from the other (disregarding transaction costs, which will be considered later on). Therefore, assuming that a correlation of 1 is impossible to find in markets, let us consider a pair of cryptocurrencies with a correlation of 0.98. In contrast to the previous scenario, even though these two cryptocurrencies almost always move in the same direction, there will be moments when this relationship is not preserved due to the correlation not being exactly 1. The underlying idea is to identify those moments when the pair of cryptocurrencies deviates from their usual behavior and bet on the assumption that it is just a temporary anomaly, expecting the relationship of moving almost identically to continue over time, thus aligning the price movements of the two cryptocurrencies again. As explained in the previous chapter, using cryptocurrency prices directly is not ideal because the price time series of cryptocurrencies are neither stationary nor mean stationary, that can be seen in Figure 13. To address this issue, similar to the mean-reverting strategy, we will use the return time series of the cryptocurrencies. Therefore, in the previous example, if the two cryptocurrencies have a high correlation such as 0.98, we expect the return time series to be nearly identical, with any differences indicating temporary deviations in their relationship, which would be the entry points for the trades.

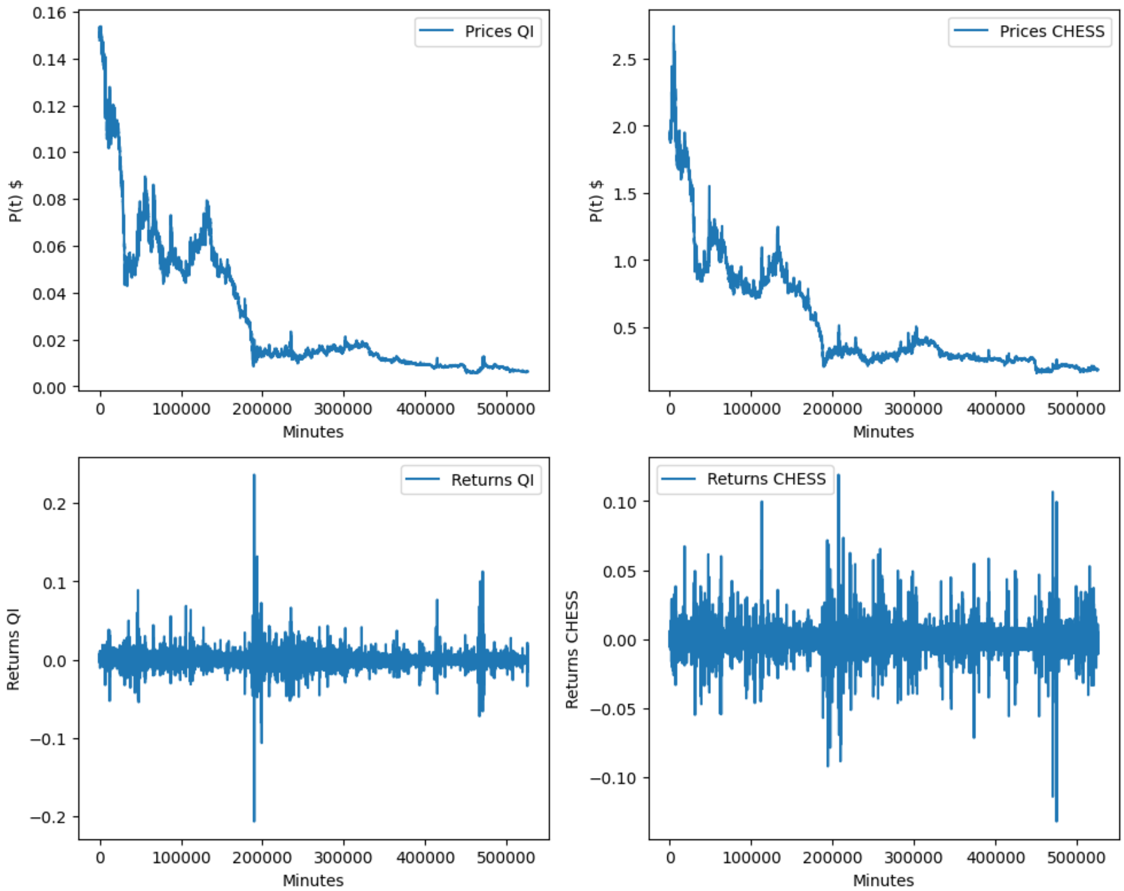

Figure 13: Top left and top right prices of two cryptocurrency(non stationary), bottom left and right, their returns(mean statioanry)

Furthermore, to facilitate the comparison of the two returns time series, we will subtract one from the other as can be seen in Figure 14. Assuming the two cryptocurrencies are A and B, the relationship that links them will be:

$$d(t) = \left(\frac{p_A(t) - p_A(t-1)}{p_A(t-1)}\right) - \left(\frac{p_B(t) - p_B(t-1)}{p_B(t-1)}\right) \tag{2}$$

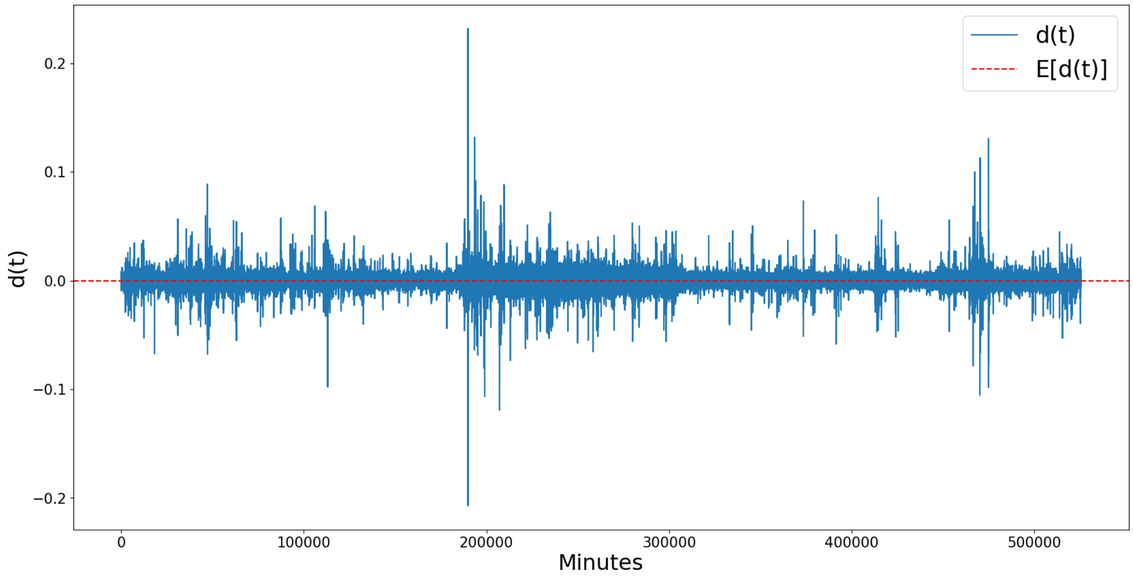

Figure 14: Differences between returns of the two cryptocurrency

By doing this, when the returns of the two individual cryptocurrency are equal, they will cancel each other out, and the resulting time series will be close to zero. However, when the relationship between the two cryptocurrencies deviates, it implies that one is experiencing higher or lower returns compared to the other. During those moments, the approach is to buy the cryptocurrency with lower returns and sell the one with higher returns. In this way, when the return time series reverts to the mean of the differences (zero), which is our basic assumption, we will close the trade.

For example, consider two cryptocurrencies, A and B, with a correlation of 0.98. Initially, we assume that both A and B have the same returns, resulting in a time series value of zero due to the subtraction of two identical values. However, in the next observation, we see that the returns of A are +2%, while the returns of B are only +1%. In this case, we would buy B because it is underperforming compared to A, and sell A because it is outperforming compared to B.

As mentioned earlier, in order to determine whether a pair is suitable for pairs trading and can generate positive results, we need a high correlation between the two cryptocurrencies. In the following sections of this chapter, we will identify a threshold below which it is no longer worthwhile to consider a pair. For further insights into the concept of correlation, please refer to the chapter "Research Methodology," where this concept was explained in detail.

Another aspect in which the pairs trading strategy significantly differs from the previous mean-reverting strategy is that the latter is implemented on one cryptocurrency at a time. Therefore, considering all the 285 cryptocurrencies available on Binance, the application of the mean-reverting strategy would allow for a maximum of 285 simultaneous operations, one for each cryptocurrency. In pairs trading, on the other hand, as it does not involve individual cryptocurrencies but rather pairs traded simultaneously, we have the opportunity to execute a maximum number of pairs, precisely $\frac{285 \cdot 284}{2} = 40470$ pairs. However, as mentioned earlier, not all pairs will be suitable for pairs trading; only those with sufficiently high correlation will be considered.

Now, taking transaction costs into account, similar to the mean-reverting strategy, we realize that it may not be opportune to open a trade every time there is a discrepancy in the returns of the two cryptocurrencies, as all the gains could be eroded by transaction costs.

Similarly to the mean-reverting strategy, we need to identify a threshold beyond which it is not worth opening a trade because the profits would be completely eroded by transaction costs. The threshold established is to open a trade only when the time series of the difference between returns difference, d(t), deviates from the mean (zero) by at least one rolling standard deviation calculated based on the previous 20 observations, just like the mean-reverting strategy.

4.2 The Pairs Trading Strategy

As done for the mean-reverting strategy, we will now explain the pairs trading strategy step by step using a concrete example. First, consider two cryptocurrencies with one of the highest correlations among all possible pairs, namely the pair formed by QI and CHESS, two of the 285 cryptocurrencies available on Binance. It has been calculated that this pair has a correlation greater than 0.99, making it an excellent candidate. It is important to specify that the correlation was calculated on the time series of all closing prices data of January 2022 on 1-min timeframe, as we will be using this information on the 2022 data later on. To understand in detail how trading develops, observe how it evolves over a short period lasting a few tens of minutes as shown in Figure 15.

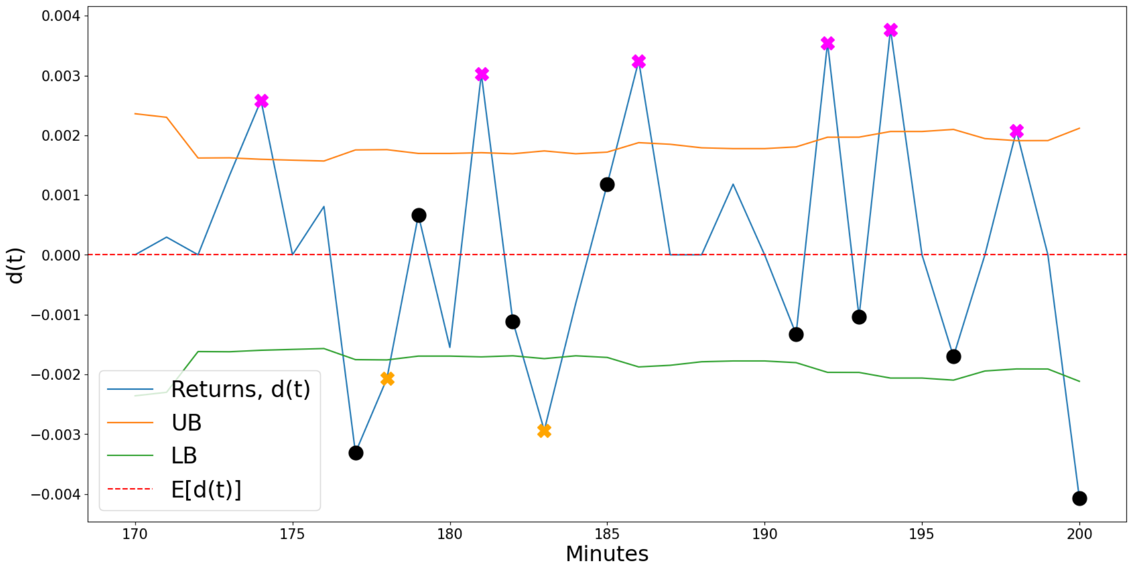

Figure 15: Differences of returns between QI and CHESS, with the "X" where the positions are opened and the black circles where they are closed

First, let us specify how d(t) is constructed in this specific case:

$$d(t) = \left(\frac{p_{QI}(t) - p_{QI}(t-1)}{p_{QI}(t-1)}\right) - \left(\frac{p_{CHESS}(t) - p_{CHESS}(t-1)}{p_{CHESS}(t-1)}\right) \tag{3}$$

Observe Figure 15. As described above, we create the upper bound (UB, orange line), the lower bound (LB, green line) and the mean (zero, red line). At this point, the first trade will open as soon as the differences of returns, d(t) (blue line), cross either of these two bounds. This event can be observed at minute 174, coinciding with a magenta "x." Therefore, at that moment, d(174) = 0.002584, while UB(174) = 0.001598, indicating that $d(174) > UB(174)$. Having crossed the UB, we will buy the undervalued cryptocurrency and sell the overvalued one. In this specific case, since $d(174) > 0$ due to crossing the UB, the overvalued cryptocurrency is QI, and we will sell it, while CHESS is the undervalued cryptocurrency, and we will buy it. Suppose we want to invest \$1000 for each individual trade, implying that we will buy \$500 worth of CHESS and sell \$500 worth of QI. At that moment, the price of

QI is $p_{QI}(174) = 0.1508$, while the price of CHESS is $p_{CHESS}(174) = 1.930$. Similar to calculating quantities for the mean-reverting strategy, in this case, we will go long on CHESS with 259.067358 units and short QI with -3315.649867 units. Additionally, we calculate the total fees resulting from these two initial trades, which amount to \$0.2 (1000\$ · 0.0002). At this point, we are in a position and wait for the trade to unfold. Three minutes later, at minute 177, we can see from Figure 15 that d(177) has again fallen below the mean (zero), specifically $d(177) = -0.003311 < 0$. At that moment, we will close the trade. At that instant, $p_{QI}(177) = 0.1505$, $p_{CHESS}(177) = 1.931$, and we will observe a P&L equal to:

$$P\&L_{total} = P\&L_{QI} + P\&L_{CHESS} - Total\_Fees \tag{4}$$

Performing the calculations and recalling that

$$P\&L = quantity \cdot (p_{final} - p_{initial}) \tag{5}$$

we have:

- $P\&L_{QI} = -3315.649867\$ \cdot (0.1505 - 0.1508) = 0.995\$$
- $P\&L_{CHESS} = 259.067358\$ \cdot (1.931 - 1.930) = 0.26\$$
- $Total\_Fees = 0.2\$ + (3315.649867\$ \cdot 0.1505 \cdot 0.0002) + (259.067358\$ \cdot 1.931 \cdot 0.0002) = 0.399853\$$

Therefore, the total P&L is 0.86\$.

At this point, once the trade is closed, the same process restarts with the next observation, and as soon as d(t) crosses either the UB or the LB, we will enter a new position.

In the following Figure 16 we present the final results obtained by applying this strategy to the pairs formed by QI and CHESS. It can be observed that the strategy on this pair has been profitable. Whether considering transaction costs or not, a positive return was achieved at the end of the year. The simulation considered the following values: Capital invested per trade = \$1000 and Transaction costs = 0.2%.

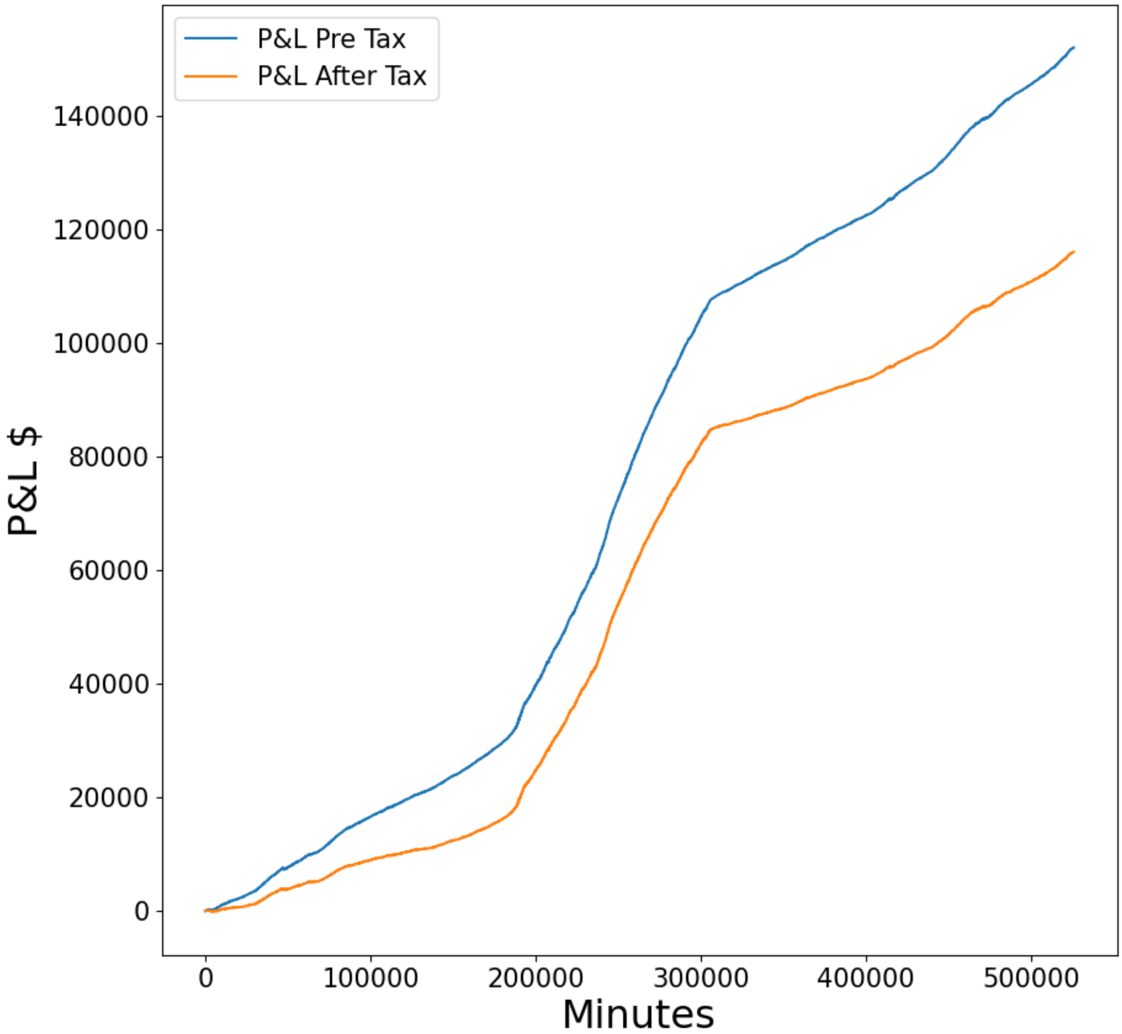

Figure 16: Cumulative profit pre-tax(blue) and after-tax(orange)

The results obtained were as follows:

- P&L pre-tax = \$151989.35
- Total fees = \$34326.63
- P&L after-tax = \$117662
- #trades = 57740
- E[P&L per trade pre-tax] = \$2.63
- E[P&L per trade after-tax] = \$2.04

Just like the mean-reverting strategy, the results are significantly affected by transaction costs. For this reason, a method aimed at reducing the impact of these costs on the final P&L will be exposed at the end of the thesis.

4.3 Results of the Pairs Trading Strategy

Unlike the mean-reverting strategy, where the overall results could be obtained by executing the strategy on all 285 cryptocurrencies and taking the cumulative results, for pairs trading, we cannot do the same because one of the assumptions made from the beginning was that the pairs consist of two cryptocurrencies with a high correlation. Therefore, it would not make sense to apply the strategy to a pair with a very low correlation. However, the concept of how high the correlation should be is not universally recognized, so it is a value to be determined.
Initially, the strategy was run on all currency pairs with a correlation greater than 0.99, as it will be illustrate. Then the same process was repeated for all pairs with correlations between 0.98 and 0.99, and so on for all the correlation intervals until the last interval between 0.70 and 0.71.
Now let us present the results obtained by running the strategy on all cryptocurrency pairs with a correlation greater than 0.99. It should be noted that the number of pairs with such correlation is 154. The Figure 17 shows the cumulative P&L before and after-tax. In the Figure 18 we can also see the distribution day by day of the PLs.

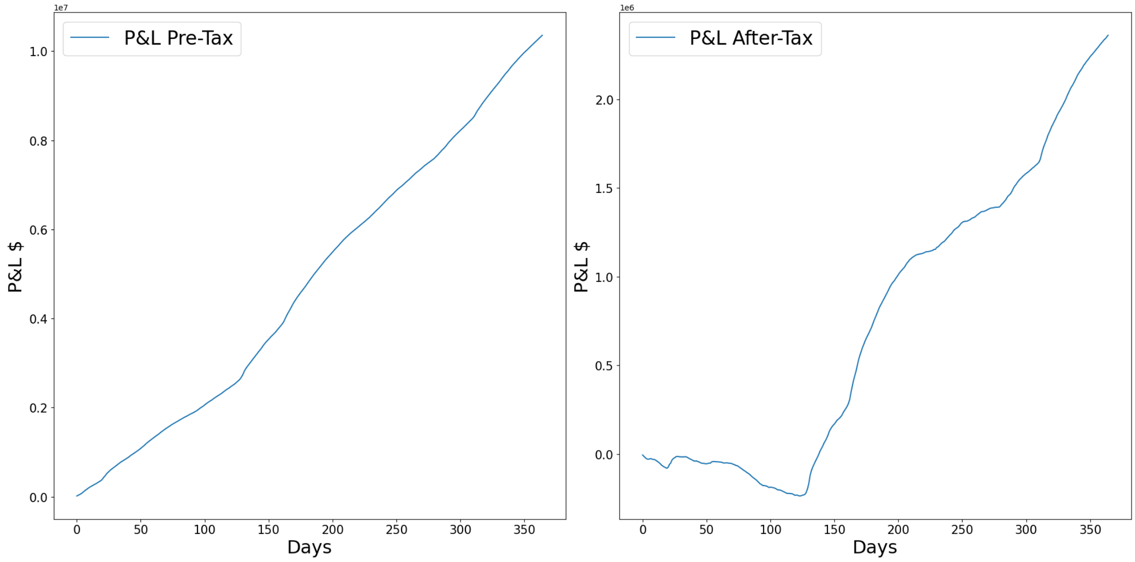

Figure 17: Cumulative P&L pre-tax(left) and after-tax(right) on all the currency pairs with a correlation greater than 0.99

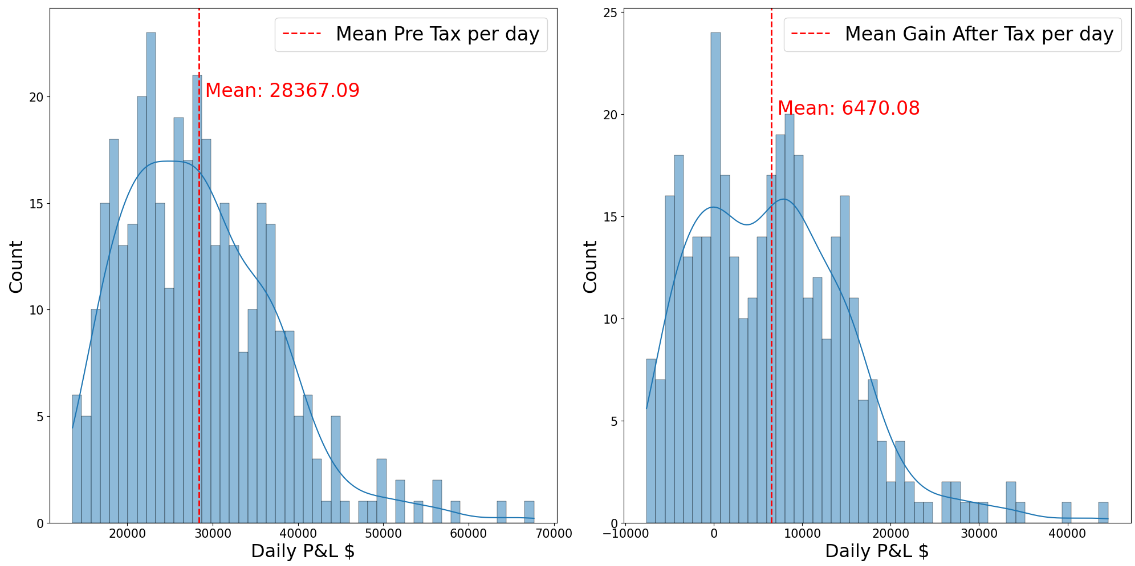

Figure 18: Distribution day by day of the P&Ls on all the currency pairs with a correlation greater than 0.99

Furthermore, the overall results (after-tax) for all pairs with a correlation greater than 0.99 are:

- #pairs: 154
- Total profit after-tax: $2361580.80

- E[P&L per crypto in a year]: $15334.94
- Positive days: 267 (73%)
- Negative days: 98
- Min P&L in a day: -$7625.62
- Max P&L in a day: $44580.14
- Max drawdown: $231041.39
- Final P&L after-tax/Max drawdown: 10.22

At this point, focus on the last metric reported, which is the Final P&L after-tax/Max drawdown. This metric determines how the final P&L relates to the largest loss incurred during the year. For example, if this value is approximately 10, as in the case of a correlation greater than 0.99, it means that for each unit of drawdown, the gains at the end of the year are approximately 10 times greater. The graph in image 19 shows this value for each correlation interval.

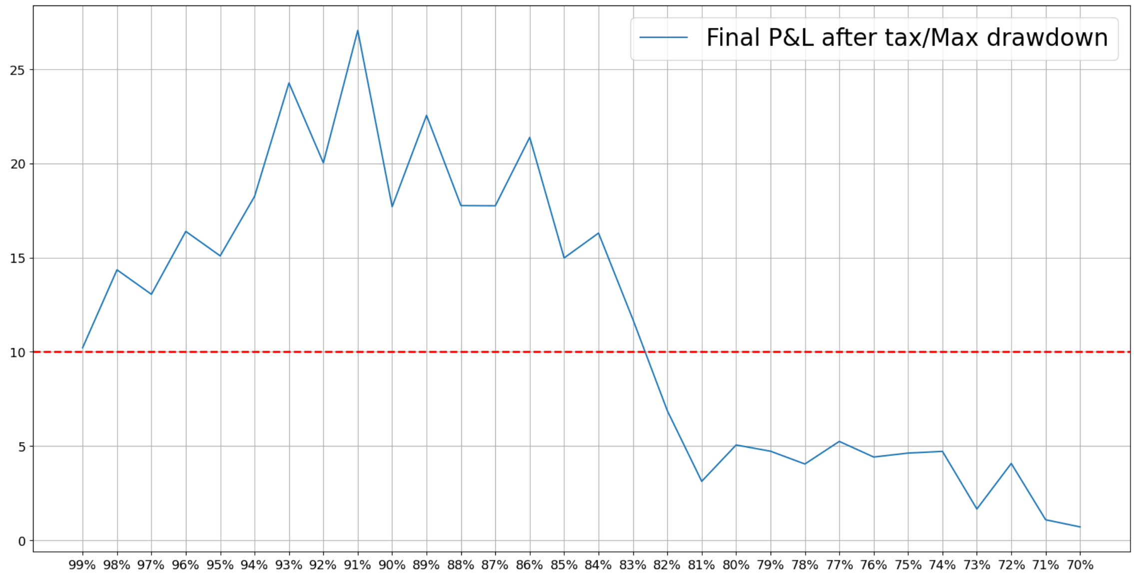

Figure 19: Plot of all the values of "Final P&L after-tax/Max" drawdown for each interval

It can be seen from the graph that the selected metric follows a bell shape. For the purpose of this thesis, it has been decided that the minimum acceptable value for the pairs to be considered is 10. This implies that a minimum correlation of 83% is required for the pairs to be considered.

Now we can move on to the last part of this section, which involves presenting the results of applying the strategy to all pairs with a correlation greater than 0.83. First, let us show the cumulative profit before and after-tax in image 20, and the daily distribution of P&L before and after-tax in image 21.

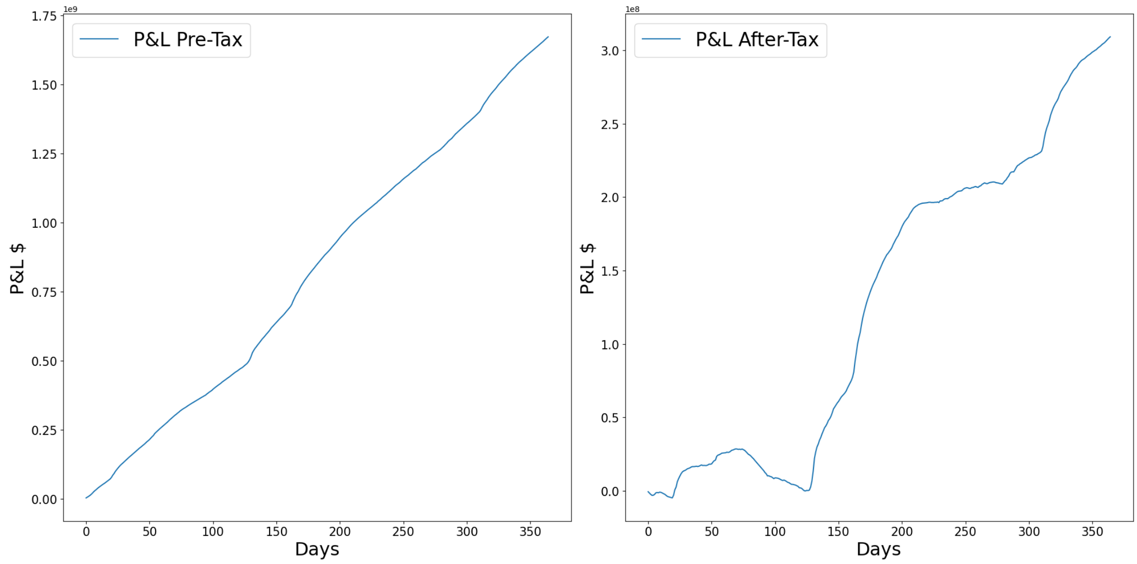

Figure 20: Cumulative P&L pre-tax(left) and after-tax(right) on all the currency pairs with a correlation greater than 0.83

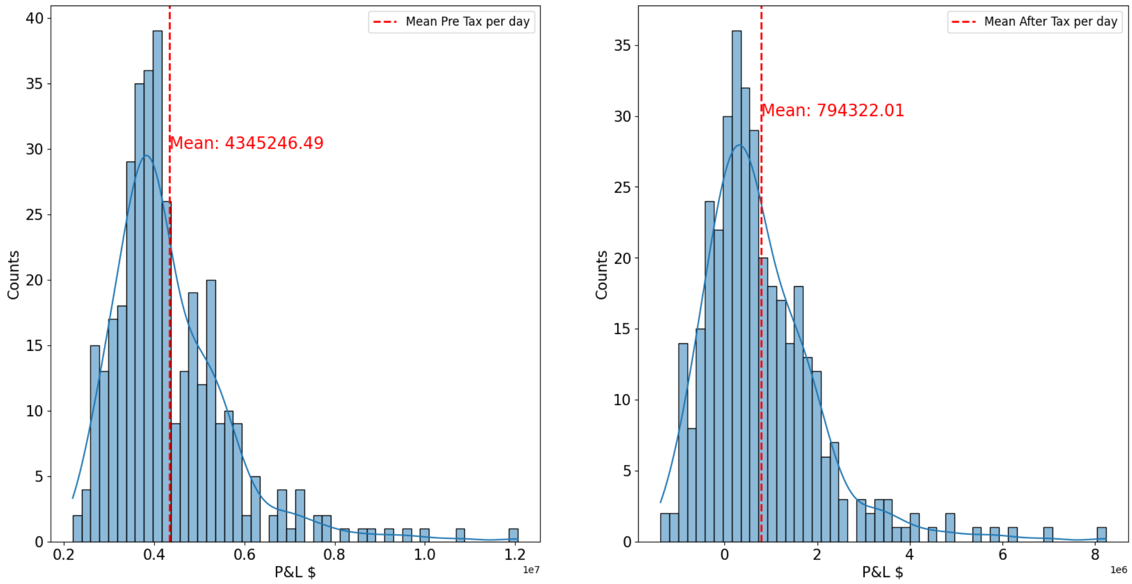

Figure 21: Distribution day by day of the P&Ls on all the currency pairs with a correlation greater than 0.83

Furthermore, the overall results (after-tax) for all pairs with a correlation greater than 0.83 are:

- #pairs: 26346
- #trades in a year: 3407905882
- #trades per minutes: 6483
- Total profit pre-tax: $1586014969.07
- Total fees: $1296087434.05
- Total profit after-tax: $289927535.02
- E[P&L per pair in a year]: $11004.61
- Positive days: 274 (75%)
- Negative days: 91
- Min P&L in a day: -$1381154.35

- Max P&L in a day: $8230891.33
- Max drawdown: $27591379.50
- Final P&L after-tax/Max drawdown: 10.5

Following these results, an important clarification needs to be made. The results obtained are based on specific choices, such as the investment of $1000 per trade. If we were to invest solely in a single pair, as demonstrated in the introductory example between CHESS and QI, an investment of $1000 per trade could still be considered a reasonable amount, considering that we observed an average return of approximately $2 per trade. However, it is crucial to emphasize that considering we have over 27000 pairs, if we hypothetically assume that all these pairs activate the setup to open a trade simultaneously, it would require a capital of at least 27000 x $1000 = $27000000. It is evident that, unless having millions of dollars, it would not be feasible to invest such an amount. This clarification serves to highlight that the results obtained are purely theoretical and are based on unrealistic assumptions, such as the ability to buy and sell infinite quantities at the closing price.
Therefore, although the results obtained from applying the pairs trading strategy to cryptocurrency pairs with high correlations show promising returns, it is essential to consider the practical limitations and constraints involved in implementing such a strategy in a real-world scenario.

4.4 Optimized Pairs Trading Strategy

Now that we have demonstrated the basic functioning of the pairs trading strategy, as we did for the mean-reverting strategy, we wonder if it is possible to optimize this strategy. First and foremost, it is important to mention that, just like with the previous strategy, an attempt was made to construct a neural network to identify specific patterns that could determine whether a trade would be profitable or not. However, once again, the results were unsatisfactory as the network did not discover any patterns and therefore did not exclude any operations from being performed.
A second unsuccessful attempt was made by enforcing that the closure of each

operation would not occur if the previously given rules were respected, but rather would occur no matter what in the following minute.
Finally, a more satisfactory approach was pursued by revising the logic of pair formation. Specifically, in the strategy outlined in the preceding sections, a minimum correlation threshold of 0.83 was established, and all pairs exhibiting a correlation exceeding 0.83 in the time series of January 2022 closing prices were actively traded throughout the entirety of 2022. The new idea was to divide the year into 12 months and create pairs based on the correlation calculated the month before the month of operativity: for example, in August, only pairs with a correlation higher than 0.83 from the previous month, i.e., July, were selected. This was done for all months, aiming to ensure a sufficiently high correlation over a shorter time period compared to the entire year. The results obtained by making this modification are reported in Table 8.

Month	Profit before Taxes $	Fees $	Profit after Taxes $	#Pairs	#Trades
Feb	127896088	109864937	18031151	27458	288876626
Mar	49452811	56823631	-7370820	14315	149410897
Apr	61564043	70558837	-8994794	17759	185525968
May	149045550	93553482	55492068	23527	245987619
Jun	207109178	113931761	93177417	28793	299569850
Jul	105317624	73773769	31543855	18927	193979244
Aug	39176672	35806154	3370518	9116	94147971
Sep	85215991	79179079	6036912	21039	208191857
Oct	52328336	43300113	9028223	11360	113852435
Nov	26044298	17668482	8375816	4319	46457147
Dec	119989503	88982864	31006639	25249	233969728
TOT	1023140094	783443109	239696985		2059969342

Table 8: Results month per month, year 2022

Now, to analyze the results, we need to take into account that the number of operations and the operations executed, as explained, are not identical. Therefore, we can compare the new results by considering these differences, which can be summarized in Table 9.

These data are very satisfying because, pre-tax, we have increased the profit

	Old Method	New Method
E[P&L per trade pre-tax]	$0.47	$0.5
E[P&L per trade after-tax]	$0.09	$0.12

Table 9: Comparison of Old and New Methods

per trade by approximately 7%. This implies that the executed operations were statistically trades with a higher return. Moreover, we can see that after-tax, the expected P&L has increased by approximately 37%. This latter result is very good as it indicates that, overall, the operations executed using this new method involve better pairs. These pairs have a higher probability of leading us to a profit rather than a loss.

4.5 Conclusions on Pairs Trading

In conclusion, we can affirm that the pairs trading strategy, like the mean-reverting strategy, is a profitable approach. The significant difference is that in this strategy, there is no directional bias; instead, it focuses solely on identifying small and temporary mispricing between two highly correlated currencies.

5 The portfolio methodology

Thanks to the previous chapters, titled "Mean Reverting Strategy" and "Pairs Trading Strategy," we demonstrated the effectiveness of these two strategies on a 1-minute timeframe, covering the period from January 1, 2022, to December 31, 2022. In those chapters, we also took into account transaction costs and found that both strategies were profitable.

In this chapter, our goal is to illustrate a method that can elevate these strategies to make them even more profitable. While optimizing the strategies was already an objective in the previous chapters, where we sought to find optimal parameters to maximize returns, in this chapter we will focus on reducing transaction costs. By reducing these costs, we can benefit from an increase in overall final profits. However, the only applicable method to reduce these costs is by reducing the number of executed trades.

At this point, the reader may naturally question how it is possible to reduce "fixed" costs such as transaction costs and, consequently, reduce the total number of trades while achieving the same, if not better, results as before. To answer this question, we need to introduce a new concept. Previously, we viewed the individual strategies as separate entities that, when applied to cryptocurrencies or cryptocurrency pairs, statistically led to generate profits. Specifically, in the previous chapters, the results obtained were the outcome of applying the two strategies to all cryptocurrencies on Binance for the mean-reverting strategy or to all pairs with a correlation higher than 0.83 for the pairs trading strategy. However, not all cryptocurrencies or pairs were consistently profitable at the same time. Nevertheless, the overall results always led to profits within the considered timeframe.

Now, the new concept to grasp is as follows: ultimately, the combination of these two strategies implies that at any given moment, one must either buy or sell or remain flat on the 285 cryptocurrencies on Binance for a certain underlying quantity (volume). Consequently, it is possible, for example, that for a cryptocurrency X, one needs to go long using the mean-reverting strategy while simultaneously going short with the pairs trading strategy. Assume, for example, that we need to go long with 20 units of cryptocurrency X using the mean-reverting strategy while going short with 10 units using the pairs trading strategy. The desired final outcome is that we only want to go long with a total of 10 units (+20 -10 = +10). In this example, assuming the transaction costs remain the same as previously considered, i.e. c=0.02%, if we had executed the strategies individually, we would have incurred transaction costs of exactly 20x0.02% for the mean-reverting strategy and 10x0.02% for the pairs trading strategy, resulting in a total of 30x0.02%. However, by executing the combined operation, we would have only incurred transaction costs of 10x0.02%. The example provided is straightforward, but we can assume that the considered cryptocurrency X is involved in 100 simultaneous operations, such as 50 long positions with certain pairs and 50 short positions with other pairs. The final result will always be the sum of the positions we want to represent. In this latter case, paradoxically, it may be possible to avoid executing any operations and save 100% of transaction costs.

To concretize this reasoning, we need to shift from viewing two strategies (and potentially many others) as separate entities with rules dictating whether to buy or sell a certain asset at a given moment, to viewing the strategies as components of a portfolio of strategies. In this portfolio, each strategy contributes by determining the volume to be bought or sold for a given cryptocurrency. The final result will be a portfolio of strategies that precisely indicates, at any given moment, the quantity of each cryptocurrency to be held in order to execute all the strategies simultaneously. It is important to emphasize that by adopting this approach, we are not reducing the number of trades by choosing not to execute some of them. Rather, we are simply trying to avoid executing useless trades, as they would involve simultaneous buying and selling of the same instrument, resulting in unnecessary payment of fees. Using the concept of this portfolio methodology instead of individual strategies, it becomes evident that the more profitable strategies we possess statistically, the more trades we will need to execute simultaneously. Consequently, the frequency of encountering situations where opposite trades must be executed simultaneously on the same cryptocurrency increases, allowing for their mutual cancellation. Therefore, one method to effectively apply this concept is to implement multiple strategies simultaneously.

Let us now present the new results of the pairs trading strategy explained in the previous chapter using this new portfolio methodology, and compare them with the previous results (see Table 10).

	Old Method	Portfolio Methodology
Total Profit Pre-tax	\$1586 Mln	\$1586 Mln
Total Fees	\$1296 Mln	\$934 Mln
Total Profit After-tax	\$290 Mln	\$652 Mln

Table 10: Results of Pairs Trading Strategy on a 1-Minute Timeframe: Old Method vs. Portfolio Methodology

These numbers must be considered as the result of specific assumptions, including the investment of \$1000 per operation, which, as previously noted, would be excessively expensive. Therefore the numbers should not be evaluated in an absolute manner but primarily focusing on the difference before and after implementation and not on the numerical values themselves. What can be observed is that, as expected, pre-tax profits remain unchanged from before to after (they were calculated mainly to verify if the two codes yielded the same results, as any difference would indicate an issue). Transaction costs have been reduced by approximately 30% by transitioning from the old method to the portfolio of strategies methodology. As a result, after-tax profits have increased, more than doubling the overall profits.

It is now wanted to mention that, in addition to the one just discussed, there are also other methods for increase the number of simultaneous trades in order to statistically reduce transaction costs. For example, we can employ the same strategy with different parameters. Another approach is to apply the same strategy but on a different timeframe. For instance, in the previous two chapters, we used closing data for each minute, making those strategies profitable on a 1-minute timeframe. However, we can divide the time series into sub-time series with a longer timeframe. For example, if we split the 1-minute timeframe data into two time series, one con-

taining only even minutes and the other containing only odd minutes, we would obtain two time series with a 2-minute timeframe, each being half the length of the 1-minute timeframe. By iterating this idea, we can use timeframes of three, four, ten minutes, and so on.
In short, we realized that there are numerous ways to increase the number of trades and, following the previous promising results, we now present an additional method to increase the number of trades and take full advantage of the strategy portfolio methodology. Specifically, we will now start showing the results of the pairs trading strategy by implementing the portfolio methodology on all trades executed on the 2-minute timeframe (whereas the previous case all trades were executed on a 1-minute timeframe). First, in Table 11, we present the results of the pairs trading strategy alone with a 2-minute timeframe using the old method, compared with the results obtained using the portfolio methodology.

	Old Method	Portfolio Methodology
Total Profit Pre-tax	\$1905 Mln	\$1905 Mln
Total Fees	\$1347 Mln	\$976 Mln
Total Profit After-tax	\$558 Mln	\$927 Mln

Table 11: Results of Pairs Trading Strategy on a 2-Minute Timeframe: Old Method vs. Portfolio Methodology

Comparing Table 10 with Table 11 one can immediately see that the pairs trading strategy is more profitable when implemented on a 2-minute timeframe than on a 1-minute timeframe. This is an excellent indication of the goodness of the strategy and also indicates that although a profitable strategy was previously identified, this does not guarantee that an even more profitable and successful strategy cannot exist by appropriately adjusting certain parameters or, as in this case, the timeframe. From this comparison, in particular, we want to highlight one aspect; with the portfolio methodology in the case of both the 1-minute timeframe-based strategy (table 10) and the 2-minute timeframe-based strategy (table 11) transaction costs were reduced and in particular with 2-minute timeframes transaction costs decreased by 28%, resulting in a 66% increase in total after-tax profits.

Now we present in Table 12 the comparison between the results obtained by combining the two pair trading strategies on 1-minute and 2-minute timeframes into a single portfolio, and the results obtained without using the portfolio methodology.

	Two Separate Portfolios	Portfolio Methodology
Total Profit Pre-tax	$3491 Mln	$3491 Mln
Total Fees	$2643 Mln	$1910 Mln
Total Profit After-tax	$848 Mln	$1581 Mln

Table 12: Combined results of Pairs Trading Strategy on a 1 and 2 Minute Timeframe with Portfolio Methodology

As the results show, combining the two portfolios further increases total profit after-tax by further reducing total fees. Specifically transaction costs have been reduced by approximately 30% resulting in an increase of total after-tax profits of almost 90%.

To further demonstrate the effectiveness of the portfolio methodology, some previously used statistics are presented in Table 13.

	E[P&L per trade pre-tax]	E[P&L per trade after-tax] (old method)	E[P&L per trade after-tax] (portfolio methodology)
Pairs Trading on 1 min timeframe	$0.47	$0.09	$0.19
Pairs Trading on 2 min timeframe	$0.57	$0.17	$0.28
Pairs Trading on 1 and 2 min timeframe	$0.53	$0.13	$0.24

Table 13: Comparison of statistics between the three portfolios

In order to highlight the difference between the two methods, it is important to highlight the comparison of after-tax results with pre-tax results, as this allows us to really understand how much each transaction is affected by fees.

The closer the E[P&L per transaction after tax] is to the E[P&L per transaction before tax], the more we were able to eliminate transaction costs from each transaction.

Before moving on to the final part of this section, we present the results of the optimized pairs trading strategy on a 1-minute timeframe, as illustrated in the chapter "Pairs Trading Strategy" in Table 8, implemented with the portfolio of strategies methodology in Table 14.

	Profit before-taxes	Fees	Profit after-taxes	#Pairs	#Trades
Feb	127896088	84168919	43727169	27458	288876626
Mar	49452811	43712560	5740251	14315	149410897
April	61564043	54306269	7257773	17759	185525968
May	149045550	71910287	77135262	23527	245987619
Jun	207109178	86739180	120369998	28793	299569850
Jul	105317624	55693289	49624335	18927	193979244
Aug	39176672	27582942	11593729	9116	94147971
Sept	85215991	60359319	24856672	21039	208191857
Oct	52328336	32517758	19810577	11360	113852435
Nov	26044298	13649979	12394319	4319	46457147
Dec	119989503	65149707	54839796	25249	233969728
TOT	1023140094	595790209	427349881		2059969342

Table 14: Results of the optimized pairs trading strategy on a 1-minute timeframe

Even with the optimized strategy, we can observe significant improvements in the final P&L by using the portfolio methodology. Comparing the results in Table 14 with the previous results in Table 8, we can see that after-tax profits have increased by almost 80% while transaction costs have decreased by approximately 25%.
It is evident that this portfolio methodology is generally effective in reducing transaction costs. However, in this final section of the thesis, it is important to make one last clarification. The pairs trading and mean-reverting strategies were programmed, backtested, and optimized using data from 2022. However, if one intends to use one or more of these strategies in the future, it is crucial to verify their performance out of sample.
For this reason, at the conclusion of this thesis, exactly that has been done. Specifically, data from January 1, 2023, to June 15, 2023, was downloaded to fully test the strategies out of sample. Furthermore, these two strategies were included in the same portfolio of strategies, with the hope of observing similar profitability results to

what was obtained in the in-sample testing. The results can be observed in Table 15.

	Profit before-taxes	Fees	Profit after-taxes	#Trades	E[P&L per trade after-tax]
Jan	$216 M	$107 M	$108 M	448,337,596	$0.24
Feb	$234 M	$145 M	$89 M	614,400,706	$0.14
Mar	$25 M	$14 M	$11 M	59,307,986	$0.19
Apr	$48 M	$32 M	$16 M	132,627,270	$0.12
May	$72 M	$46 M	$26 M	189,021,311	$0.14
TOT	$595 M	$344 M	$251 M	1,443,694,869	$0.17

Table 15: Results month per month, from January 2023 to May 2023

Based on these latest results, we can state that, using the concept of portfolio methodology, the combination of the mean-reverting strategy with pairs trading strategies on 1 and 2 minute timeframes is a profitable approach even when trading out of the sample.

As recalled several times, the assumption of investing a capital of $1000 per operation is economically unrealistic. Therefore, to conclude this chapter, we propose a final out-of-sample test in which, instead of employing a fixed capital of $1000 for each transaction, we decide to vary the capital "c" according to the number of pairs considered (namely, those that, as seen, have a correlation greater than 0.83).

Specifically, on a single timeframe (in our case 1 minute) one wants to ensure that the capital 'c' invested in each individual trade multiplied by the total number of trades does not exceed certain maximum limits. Therefore, the parameter "c" to be invested in each trade is calculated by hyperbolic interpolation so that, on average, each minute an investment of about $5 is made, to be divided among all the trades as a whole. This means that every minute, on average, when trading configurations are activated, the maximum capital to be invested among all trades will be about $5.

Finally, again for this last test, the correlation was no longer calculated on a monthly basis but on a weekly basis; thus considering a hypothetical week X the correlation was calculated in the time interval of week X-1.

The final results of the portfolio consisting of pairs trading strategies on 1 and 2 minute timeframes, together with the mean-reverting strategy, for the period from January 1, 2023 to May 31, 2023, are shown in Table 16, while the cumulative profits over time can be observed in Figure 22.

Week	Profit before-taxes	Fees	Profit after-taxes	#Trades	Mean $ invested per minute	E[P&L per day]
1	36515	20640	15875	40137200	5.31	2268
2	27577	13309	14268	151302987	3.97	2038
3	37544	20492	17051	41400079	5.47	2436
4	42100	22939	19160	11068660	6.10	2737
5	33862	20339	13522	40057960	4.94	1932
6	30397	19494	10902	50559404	4.36	1557
7	25069	16749	8319	89856567	3.67	1188
8	28852	21475	7377	24641336	4.16	1054
9	31080	17292	13787	90502405	4.48	1970
10	29774	14557	15217	138859958	4.26	2174
11	29418	15652	13766	119906934	4.25	1967
12	32593	20098	12494	48502579	4.68	1785
13	29244	20595	8648	37123333	4.19	1235
14	33187	25082	8104	5242603	4.83	1158
15	30314	21646	8668	30571393	4.43	1238
16	23378	14997	8380	122470635	3.35	1197
17	32280	23595	8685	13965485	4.65	1241
18	37307	21272	16035	35074018	5.41	2291
19	22230	15308	6921	104115284	3.16	989
20	26253	22486	3766	12935321	3.75	538
21	27471	20738	6732	26829510	3.93	962
22	34869	21825	13044	19802873	5.06	1863
TOT	681314	430580	250721	1254926524	4.47	1628

Table 16: Results week per week, from January 1, 2023, to May 31, 2023

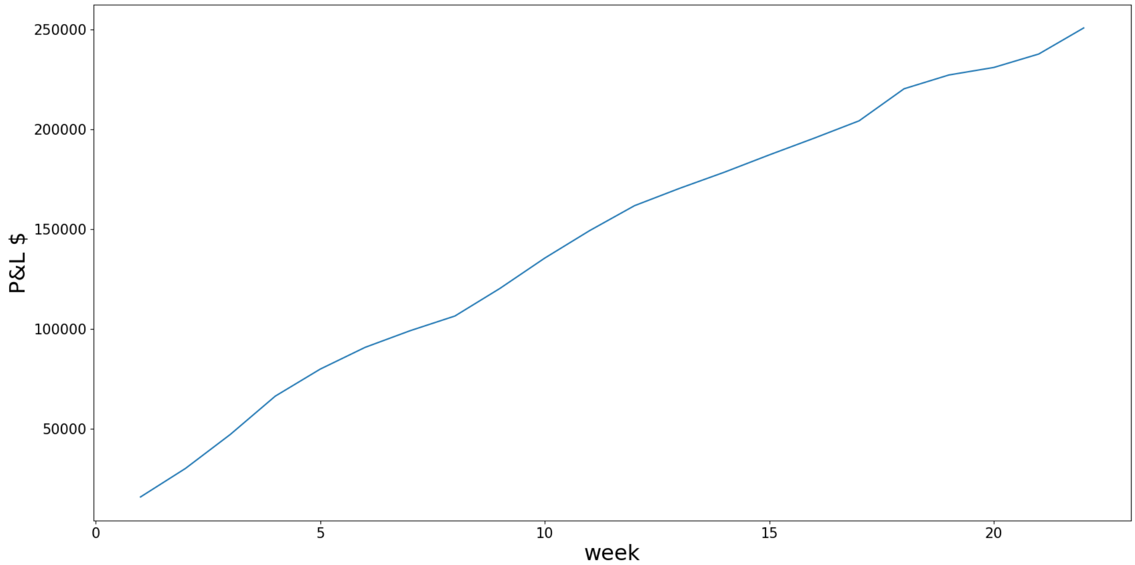

Figure 22: Cumulative profit by using portfolio methodology from January 1, 2023, to May 31, 2023.

Based on the results obtained, we can conclude that the hyperbolic interpolation function used to determine the capital to be invested for each trade, denoted "c," was effective. In fact, during the entire period under consideration, in each minute the invested capital remained very close to the desired target value of \$5. Moreover, by analyzing the cumulative P&L curve, we can see that it has a consistently positive slope, indicating minimal drawdowns. However, this portfolio of strategies can still be improved. In fact, looking at the total results shown in Table 16, we see that an average capital of \$4.47 per minute was invested, while the returns per minute were \$1628/1440 = \$1.13. To increase this value, we can apply the approach explained earlier of increasing the number of trades that can offset each other to reduce transaction costs. To this end, as mentioned above, it may be appropriate to expand the range of timeframes on which to execute pairs trading and mean-reverting strategies. Finally, exploring different logics and/or developing new strategies could also help to increase the number of trades.

6 Conclusions

This thesis aimed to delve into the topic of statistical arbitrage strategies such as mean-reverting and pairs trading. Initially, the mean-reverting strategy was implemented and optimized. The mean-reverting strategy capitalizes on the statistical and econometric concept of a time series being mean stationary to predict whether a specific cryptocurrency is overpriced or underpriced. Therefore, this strategy is directional, as it involves buying the cryptocurrency only when its returns deviate negatively by more than one standard deviation from their mean, or selling the cryptocurrency when its returns deviate positively by more than one standard deviation from their mean.

Subsequently, the pairs trading strategy was implemented and optimized. Unlike the mean-reverting strategy, pairs trading is a non-directional strategy, as profits can be generated regardless of the price movements of the cryptocurrency pairs. This strategy relies on two statistical concepts. The first concept is correlation, where it was found that a correlation above 0.83 between two cryptocurrencies could guarantee a profit with an acceptable drawdown. The second concept, similar to the previous one, is mean reversion, but this time not based on the returns of a single cryptocurrency, but on the difference between the pair of cryptocurrencies.

Lastly, this thesis aimed to analyze how transaction costs could be reduced by introducing the concept of a portfolio methodology. Rather than viewing the strategies as separate entities with rules dictating when to buy or sell a certain asset at a given moment, the strategies are considered as components of a portfolio. In this portfolio, each strategy contributes by determining the volume to be bought or sold for a given cryptocurrency. The final outcome is a portfolio of strategies that precisely indicates, at any given moment, the quantity of each cryptocurrency to be held in order to execute all the strategies simultaneously.

The results of the final out-of-sample test indicate that this created portfolio of strategies, which included pairs trading strategies on 1-minute and 2-minute timeframes, along with the mean-reverting strategy, was profitable.

7 Bibliography

[1] Gesina Gorter, Pairs trading (2006)

[2] Gatev et al., Pairs Trading: Performance of a Relative Value Arbitrage Rule (2006)

[3] Gatev, E., Goetz-mann, W. N., and Rouwenhorst Pairs trading: Performance of a relative-value arbitrage rule (2006)

[4] Vidyamurthy, G., Pairs trading: Quantitative methods and analysis (2004)

[5] Elliott, R. J., Van Der Hoek, John, and Malcolm, W. P., Pairs trading (2005)

[6] Mudchanatongsuk, S.,Primbs, J. A., and Wong, W., Optimal pairs trading: A stochastic control approach (2008)

[7] Liew, R. Q. and Wu, Y., Pairs trading: A copula approach (2013)

[8]Hanqin Zhanga, Qing Zhangb, Trading a mean-reverting asset: Buy low and sell high (2008)