# Cryptocurrency trading-pair forecasting, using machine learning and deep learning Technique

Osifo Ernest

*Osifo.e@gmail.com*

***Abstract***

*With the volume of activities associated with trading, it has become a very tedious task. The advent of the algorithmic trading has brought with it some positive change such as reduced latency and increase in liquidity in the Financial Market. The Algorithmic Trading also came with some high demands for the technological know-how and the resources to run it. This has put the retail trader in a seemingly disadvantaged position as these algo-programs are carefully guided secrets by those that have access to it.*
*Crypto-currency can no longer be ignored as the concept is forming the bedrock for future transactions. Although highly publicized, the concept of these smart contracts is not really known.*
*Looking into the future where the cryptocurrencies dominates over the traditional currencies, my desire with this project is to give the individual trader/ retailer an additional tool to demystify the "black-Box" of the trading crypto-pairs with algorithmic trading strategy. The intended techniques are: Long Short-Term Memory (LSTM), Autoregressive integrated moving average (ARIMA), Moving Average (MA), Cumulative Moving Average (CMA), and Artificial Neural Networks (ANN).*
*The models performance will be measured via correlation, Mean Percentage Error (MPE), Percentage Error (MAPE), Mean Square Error (RMSE) standard deviation and Sharpe ratio (for our trading models)*

***Keywords:*** *Crypto-currency, Long Short-Term Memory (LSTM), Autoregressive integrated moving average (ARIMA), Moving Average (MA), Artificial Neural Networks (ANN), Mean Percentage Error (MPE), Percentage Error (MAPE), Mean Square Error (RMSE) standard deviation and Sharpe ratio*
.

## 1. Introduction

In recent times, the universe of virtual market trading has encountered impressive expansion. Online trading platforms have changed the manner by which speculators inquire about and execute orders. The idea of algorithmic trading presents the next phase of this development, and one that is as of now affecting entities of high volume and liquidity.
In this virtual world, Cryptocurrencies can no longer be ignored. "Cryptocurrencies are digital, straightforward and simple to use in comparison with the traditional currencies" (Tony Simonovsky,20 Nov 2018) The individual investor, the big cooperation and even governments (who seek ways to control their influence in their financial markets) are interested in them.

## 2. Problem statement

First, we consider cryptocurrencies. Although new but block chain technology (which powers the cryptocurrencies) is seriously considered as the future of transactions. This ensures open and credible ledgers, and smart contracts. Cryptocurrencies though at their infant stage has made an impression in the financial world and the advantages can no longer be ignored. For this Project we will be considering the Bitcoin as our base currency.

Trading in the financial market has experienced a significant shift since the advent of algorithmic trading. The introduction of computer aided trades has improved the liquidity in the market, reduced latency and cost of trading. According to Randy Tate in his article: The Impact of Algorithmic Trading on Financial Markets of 11 June 2019; “Currently, algorithmic trading represents up to 70% of all U.S equity trading”. However, these algorithms are often kept as trade secrets and are often never published. This poses a great difficulty for the individual investor in the market that lacks the technology and resources.

We are aware that the gap between the big players and the individual investors is not one we can completely close, but my intention with this project is to offer template or guide for the individual investor. With the use of machine learning and deep learning techniques, the price movement could be forecasted and offer a pair trading strategy suggestion that could be earn them profitable trades.

## 3. Goals and Objective

The goal of this project is to help the individual crypto-currency investor that might not have the needed resource to compete in the fast-paced financial market. We intend to do this by building a model that would help an investor in cryptocurrency pick suitable pair cryptocurrencies for pair trade, and help the forecast the price movement of the currencies and suggest a possible trading strategy that could earn the trader/ investor some profitable trade.

Our objectives are as follows

- Find suitable pair for pair cryptocurrency trade
- Predict cryptocurrency price movement
- Develop and suggest a pair trading strategy

Our expected input and output are summarized in the table below.

| Input | Output |
|---|---|
| Trading data for cryptocurrencies (Sourced free from yahoo finance that contains the Open, Low, High, Close and adjusted Close prices) | • Best correlated pair<br>• Forecast Adjusted price for selected pair<br>• Develop a proposed trading strategy |

## 4. Literature Review

We will begin our probe into trading, by first considering the concept of pricing. According to Ingersoll (1987), “The law of one price communicates that that in the event that the profits from two ventures are indistinguishable in each state, at that point the present estimation of the two speculations must be the equivalent”. Also, for business sectors to be seamlessly coordinated (which is usually expected), two portfolios made from two markets can't exist with dissimilar values but yet similar payoffs (Chen and Knez 1995) in any situation where these conditions are not fulfilled, this could birth a price distortion and an arbitrage opportunity is created.  giving financial stakeholders chances to make hazard free benefits by purchasing under-valued stocks and short-selling the overrated ones (Lamont and Thaler 2003). In an impeccably proficient market, the costs completely respond to the accessible data consistently (Fama 1970).

The genesis of pairs Trading can be credited to Nunzio Tartalia, He was a quant in Wall Street in the mid 1980’s when amassed a group of mathematicians, physicists and Computer Scientists researchers. With the aid of the new powerful computers then, Nunzio Tartalia and his team were able build technical trading models that were void of human sentiments. A large number of these calculations were effective for brief timeframes however didn't show incredible consistency. however, it was reported that this unique group made a $50 million benefit for the Morgan Stanley bunch dependent on the pairs trading strategy. In spite of the fact that the group had a couple of long periods of terrible execution, the pairs trading strategy gained a decent notoriety in the monetary markets and has since become an undeniably mainstream "advertise impartial" speculation methodology utilized by people and institutional dealers just as multifaceted investments.

There are four principle techniques to execute Pairs Trading. These techniques are as follows

- the cointegration technique,
- the distance strategy,
- the stochastic spread strategy and
- the stochastic residual spread technique.

For empirical testing, Gatev, Goetzmann and Rouwenhorst (1999, 2007) [5] and [7] and Nath (2003) [9] publicized the non-parametric distance technique. Recently Elliot, van Der Hoek and Malcolm (2005) [4] and Do, Faff and Hamza (2006) [3] separately proposed the stochastic spread and the stochastic residual spread technique. Faff and Hamza (2006) [3] methods was mainly on mean -reverting models. This they did in order to introduce specifications to behavior of the spread for Pairs Trading technique.

Here are some examples of the pair Trade

- Gatev et. al. (2006), "Pairs Trading: Performance a Relative - Value Arbitrage Rule"
- Caldeira and Moura (2013), "Selection of a Portfolio of Pairs Based on Cointegration: A Statistical Arbitrage Strategy"
- Haque and Haque (2014), "Pairs Trading Strategy in Dhaka Stock Exchange: Implementation and Protability Analysis"

Since pair trading was introduced in the 1980’s has been used, fine-tuned and reused again because on a short term, it could be a profitable trade

## 5. Competitor Analysis

As indicated by Binh Do et al (2009) pair trading methodology (value union exchanging procedure) was seen as productive over a significant stretch of time despite the fact that at a declining rate. The examination demonstrated that the mean return for the period 1989-2002 was 60% not exactly the mean return for the period 1962-1988. By broadening crafted by Gatev et al (1999), Binh Do et al (2009) found no proof to recommend that the profits decrease was because of expanded challenge in the support investments industry. The principle explanation behind the negative turn on profitability, was seen as a diminishing number of offers that didn't join inside the exchanging time frame. This was credited to a violate down in the Law of One Price whereupon this exchanging methodology is based. It is a prerequisite of the Law of One Price that two resources that are close financial substitutes in the preparation time frame keep on being so in the exchanging time frame. It was likewise discovered that there was an expanding likelihood that nearby monetary substitutes characterized in the chronicled value space didn't stay close substitutes in the exchanging time frame. The assertion was made that the expanded risk was due to the fact that experts avoided this procedure. It was proposed in Binh Do et al (2009) that one should shape sets of key closeness which keeps away from pointless expenses as well as decrease non-combination dangers. By and by exchanging calculations ought to contain chance relieving instruments like stop misfortune which will limit the effect of different exchanges. One of the downsides of the basic sets exchanging system is that a couple may have a high recorded spread yet is as yet utilized a perceived pair because of the way that it has perhaps the most minimal spread for the preparation time frame for example the pair are close monetary substitutes in the preparation time frame yet in the exchanging time frame the pair have a higher spread. It was recommended that a better technique would be than structure sets dependent on various criteria and the best procedure here and there is to sit idle. It was found in Binh Do et al (2009) that times of high market unpredictability could bring about uniqueness being driven further. The disparity rate was relapsed against the market unpredictability proxied by the significant half year return standard deviation of the S&P 500 index. Along these lines a positive connection would bolster this recommendation. A negative connection between the uniqueness rate and market unpredictability remained constant for the generally. Along these lines, while unpredictability may have had a little part to play, it can't be viewed as a significant driver.

An expansion in innovation and preparing intensity of computers has driven specialists into different new headings Pair Trading. The pair trading technique talked about right now can be portrayed as a traditional methodology. A portion of the new strategies are depicted as follows:

- With the utilization of Bollinger bands, shares that were not recently thought of as a pair would now be able to be utilized for pair trading; Bollinger groups are a specialized investigation apparatus created by John Bollinger by utilizing a moving average with two exchanging groups above and underneath it and basically includes and subtracts a standard deviation estimation. Bollinger bands change themselves to economic situations by estimating value instability

- Take advantage of new technology, such as advanced execution systems and the use of trading algorithms

- Exploiting time zones for trading productivity

- Utilizing different stocks rather than single sets; the numerous stocks would need to meet all the prerequisites of the preparation period so as to frame various sets which would be exchanged over some exchanging period.

- New techniques, of which co-integration is maybe the most notable different territories of research incorporate the utilization of new separation measures, the incorporation of specialized investigation inside thorough statistical systems.

Advantages of Pair Trading

- Manage Drawdowns

An extra advantage to pair trading, especially for informal investors who should be prepared to move cash all through positions, is that they ordinarily have littler record drawdowns than individual long positions. This is by and by on the grounds that regardless of whether one position is losing cash briefly, drawing down a money market fund's worth, the other position can be picking up and supplant a portion of that incentive for the time being

- Mitigating Risk

One of the principle focal points for pair trading is that each pair exchange intrinsically hedges against risk. Since there are two exchanges included, regardless of whether one stock acts in a surprising manner (goes down), the other stock can make up a portion of the misfortunes. An auxiliary bit of leeway to this is pair trade limit risk from directional movement in the market. For instance, if a whole area drops in view of some huge news, the short position will pick up esteem counterbalancing loss from the drop in the estimation of the long position Hedging Risk with Pair Trades

- Profit in Any Market Conditions

Pair trading relies just upon the connection between the two stocks being exchanged, instead of on the general ascent of decrease of a division or the business sectors comprehensively. That implies that pair traders can discover and benefit on circumstances whether or not the market is picking up, losing, or moving sideways, or whether conditions are truly steady or exceptionally unpredictable

Disadvantages of Pair Trading

- High/ Double Commissions

The possibility of a single trade result to double (that means double commission) is a major disadvantage. For brokers working on moderately tight edges, that distinction in commissions can be the contrast between a benefit and a misfortune. In this way, most pair dealers are compelled to exchange generally high volumes, which requires increasingly capital and can expand chance.

- Model Failure

The most significant thing to be careful with when pair trading is the suspicion that a connection is genuine, and that two stocks will come back to that corresponded relationship after any disparity. Because two stocks have been connected verifiably doesn't imply that they will keep on being correlated. Recognizing feeble correlation model can be incredibly troublesome, and with a high risk of loss

## 6. Data

It is imperative that the quality of data used in this project is of good standard. The Yahoo Finance was our source for the crypto-currency prices that we investigated in this assignment. Our choice programing language is Python. Using some specific codes, we downloaded these data directly from Yahoo Finance. For this we used a mining function from Python: pdr.get_data_yahoo()

## 7. Design methodology

Data in its raw state could be incomplete and poorly structured. The first task is to clean up the data and structure it properly.

In this project, we attempt to forecast cryptocurrency price and design a strategy that will advantage of the statistical arbitrage via paired crypto-currency trading. Our task is divided into the following three major steps

- Get the Pair currencies; Find the most correlated cryptocurrency to our base Currency (Bitcoin)
- Forecast Adjusted price for selected pair
- Develop a proposed trading strategy

## A. Get the Pair currencies; Find the most correlated cryptocurrency to our base Currency (Bitcoin)

For the base currency: Bitcoin We download the data for the duration 1 JAN 2018 to 1 SEPT 2019

The common Cryptocurrencies according to NATHAN REIFF in his article "The 10 Most Important Cryptocurrencies Other Than Bitcoin", Updated Jan 8, 2020(source: https://www.investopedia.com/tech/most-important-cryptocurrencies-other-than-bitcoin/) the top three are. Ethereum (ETH), Ripple (XRP) and Litecoin (LTC). for the sake of a more robust overview, we will be considering the top ten Cryptocurrencies as at 8 October 2019 noted by Yahoo Finance: Source: https://finance.yahoo.com/news/top-10-cryptocurrencies-market-capitalisation-160046487.html

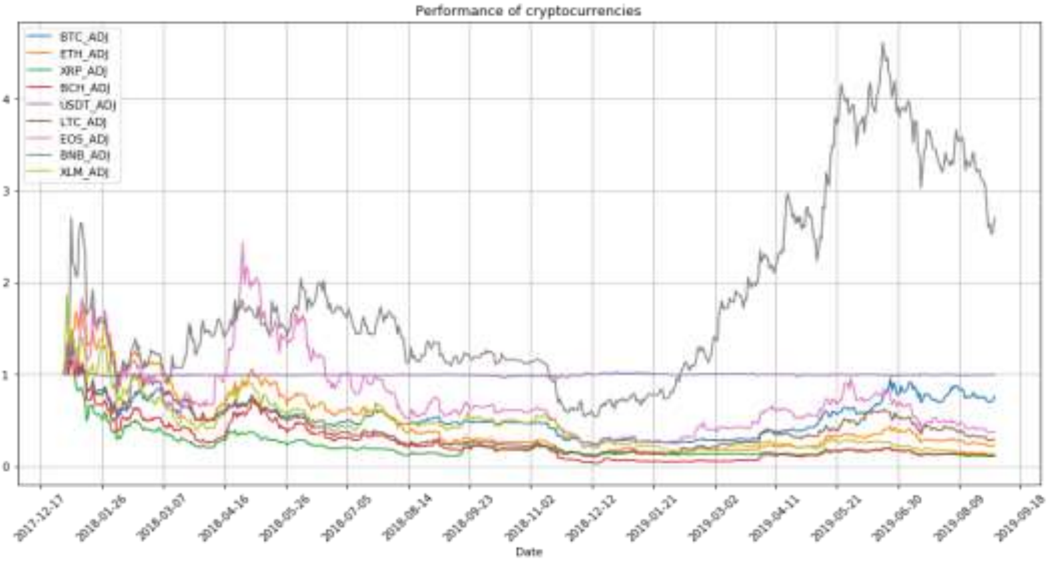
1. Bitcoin (BTC)
2. Ethereum (ETH)
3. XRP (XRP)
4. Bitcoin Cash (BCH)
5. Tether (USDT)
6. Litecoin (LTC)
7. EOS (EOS)
8. Binance Coin (BNB)
9. Bitcoin SV (BSV) the BSV has no unique ticker in yahoo finance ( and claims to be the original version of what Bitcoin was meant to be ; Source: https://bitcoinsv.com/en/learn)
10. Stellar (XLM)

these Eight (save Bitcoin SV (BSV)) would be examined to get the best pair for our base currency (Bitcoin)

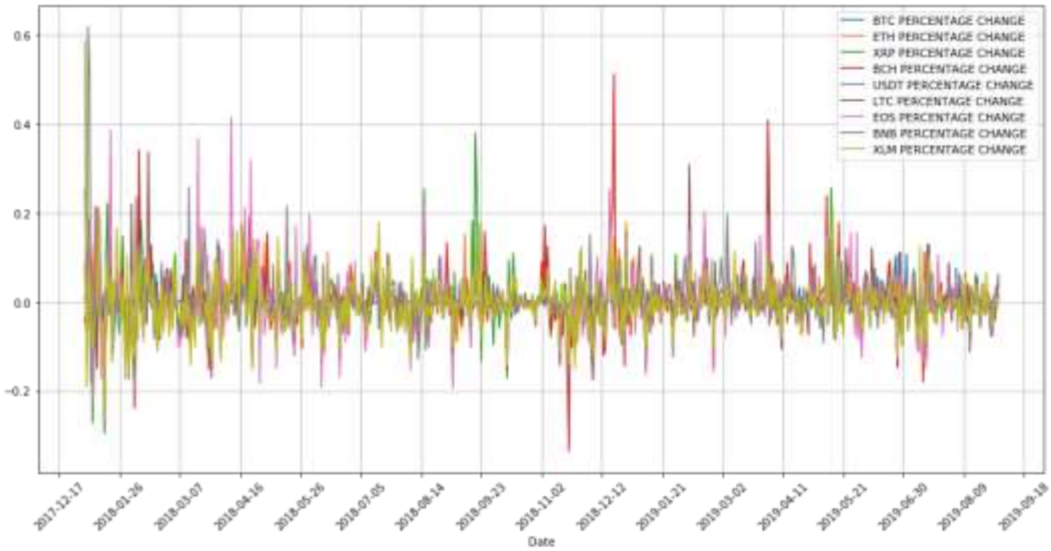
The Adjusted Close for Nine Cryptocurrencies were downloaded and concatenated into a single table as shown below.

| Date | BTC_ADJ | ETH_ADJ | XRP_ADJ | BCH_ADJ | USDT_ADJ | LTC_ADJ | EOS_ADJ | BNB_ADJ | XLM_ADJ |
|---|---|---|---|---|---|---|---|---|---|
| **2018-01-01** | 13657.200195 | 772.640991 | 2.39103 | 2432.540039 | 1.007280 | 229.033005 | 8.84318 | 8.41461 | 0.480008 |
| **2018-01-02** | 14982.099609 | 884.443970 | 2.48090 | 2711.000000 | 1.004900 | 255.684006 | 9.33471 | 8.83777 | 0.564766 |
| **2018-01-03** | 15201.000000 | 962.719971 | 3.10537 | 2608.689941 | 1.013440 | 245.367996 | 10.22790 | 9.53588 | 0.896227 |
| **2018-01-04** | 15599.200195 | 980.921997 | 3.19663 | 2430.179932 | 1.002530 | 241.369995 | 11.27550 | 9.21399 | 0.724050 |
| **2018-01-05** | 17429.500000 | 997.719971 | 3.04871 | 2584.479980 | 0.998634 | 249.270996 | 10.83960 | 14.91720 | 0.662712 |

we then visualized the Performance of the Cryptocurrencies




The Cryptocurrency prices were normalized and again plotted as shown below




With each currency scaled values captured as percentage change.

To better appreciate the progress of individual cryptocurrencies over time, the cumulative percentage change was computed and visualized as seen below:

Two major criteria that would determine the desired pair are the correlation value and Cointegration. The pair with the highest correlation, is the desired pair. High positive correlation indicates that the returns of both currencies are trending in same direction. Additionally, the cointegration of the same pair should be significantly low. A value lower than 0.05 is desired. First, we attempted to get the p-value of all possible pair in our selected (9- Crypto currency Universe)

```
BTC_RET and ETH_RET: p-value = 0.0
BTC_RET and XRP_RET: p-value = 2.470976042943848e-29
BTC_RET and BCH_RET: p-value = 1.9958557257613114e-24
BTC_RET and USDT_RET: p-value = 1.8460483551336274e-23
BTC_RET and LTC_RET: p-value = 7.855462705098844e-28
BTC_RET and EOS_RET: p-value = 1.0541411510509593e-22
BTC_RET and BNB_RET: p-value = 0.0
BTC_RET and XLM_RET: p-value = 7.646644685099445e-25
ETH_RET and BTC_RET: p-value = 3.1037273760209924e-15
ETH_RET and XRP_RET: p-value = 5.403406373617368e-15
ETH_RET and BCH_RET: p-value = 0.0
ETH_RET and USDT_RET: p-value = 5.551978810760889e-23
ETH_RET and LTC_RET: p-value = 3.5690433450567765e-14
ETH_RET and EOS_RET: p-value = 1.7607283333431284e-24
ETH_RET and BNB_RET: p-value = 1.013739230988736e-14
ETH_RET and XLM_RET: p-value = 9.22239159049706e-23
XRP_RET and BTC_RET: p-value = 2.1092846262888587e-29
XRP_RET and ETH_RET: p-value = 0.0
XRP_RET and BCH_RET: p-value = 0.0
XRP_RET and USDT_RET: p-value = 6.335357767950016e-29
XRP_RET and LTC_RET: p-value = 7.208651465169761e-26
XRP_RET and EOS_RET: p-value = 0.0
XRP_RET and BNB_RET: p-value = 0.0
XRP_RET and XLM_RET: p-value = 0.0
BCH_RET and BTC_RET: p-value = 0.0
BCH_RET and ETH_RET: p-value = 0.0
BCH_RET and XRP_RET: p-value = 0.0
BCH_RET and USDT_RET: p-value = 4.795158224967688e-27
BCH_RET and LTC_RET: p-value = 0.0
BCH_RET and EOS_RET: p-value = 0.0
BCH_RET and BNB_RET: p-value = 0.0
BCH_RET and XLM_RET: p-value = 0.0
USDT_RET and BTC_RET: p-value = 1.8869369129583224e-26
USDT_RET and ETH_RET: p-value = 2.5173714575539526e-27
USDT_RET and XRP_RET: p-value = 9.862225227015738e-28
USDT_RET and BCH_RET: p-value = 3.5444811155892606e-14
USDT_RET and LTC_RET: p-value = 1.777988533651183e-27
USDT_RET and EOS_RET: p-value = 1.1886846688942571e-27
USDT_RET and BNB_RET: p-value = 8.269927024799046e-15
USDT_RET and XLM_RET: p-value = 3.634849560475858e-27
LTC_RET and BTC_RET: p-value = 2.532674382332865e-27
LTC_RET and ETH_RET: p-value = 0.0
LTC_RET and XRP_RET: p-value = 0.0
LTC_RET and BCH_RET: p-value = 0.0
LTC_RET and USDT_RET: p-value = 0.0
LTC_RET and EOS_RET: p-value = 0.0
LTC_RET and BNB_RET: p-value = 0.0
LTC_RET and XLM_RET: p-value = 0.0
EOS_RET and BTC_RET: p-value = 1.5940880076424296e-28
EOS_RET and ETH_RET: p-value = 1.0289451215414982e-09
```

```
EOS_RET and XRP_RET: p-value = 2.453256030564739e-28
EOS_RET and BCH_RET: p-value = 0.0
EOS_RET and USDT_RET: p-value = 1.208367391394437e-23
EOS_RET and LTC_RET: p-value = 8.057981753832702e-28
EOS_RET and BNB_RET: p-value = 0.0
EOS_RET and XLM_RET: p-value = 3.776004431523901e-29
BNB_RET and BTC_RET: p-value = 0.0
BNB_RET and ETH_RET: p-value = 0.0
BNB_RET and XRP_RET: p-value = 0.0
BNB_RET and BCH_RET: p-value = 0.0
BNB_RET and USDT_RET: p-value = 0.0
BNB_RET and LTC_RET: p-value = 0.0
BNB_RET and EOS_RET: p-value = 0.0
BNB_RET and XLM_RET: p-value = 0.0
XLM_RET and BTC_RET: p-value = 0.0
XLM_RET and ETH_RET: p-value = 0.0
XLM_RET and XRP_RET: p-value = 0.0
XLM_RET and BCH_RET: p-value = 0.0
XLM_RET and USDT_RET: p-value = 2.907834588463765e-29
XLM_RET and LTC_RET: p-value = 0.0
XLM_RET and EOS_RET: p-value = 0.0
XLM_RET and BNB_RET: p-value = 0.0
```

From the results obtained we saw some very low Cointegration for our Desired Bitcoin Pair.

BTC_RET and ETH_RET: p-value = 0.0

BTC_RET and XRP_RET: p-value = 2.470976042943848e-29

BTC_RET and BCH_RET: p-value = 1.9958557257613114e-24

BTC_RET and USDT_RET: p-value = 1.8460483551336274e-23

BTC_RET and LTC_RET: p-value = 7.855462705098844e-28

BTC_RET and EOS_RET: p-value = 1.0541411510509593e-22

BTC_RET and BNB_RET: p-value = 0.0

BTC_RET /ETH_RET Pair and the BTC_RET / BNB_RET Pair tend to give the lowest P-Values, however since the desired P-value is anything less than 0.05, all of the pairs would be considered to have passed the Cointegration test. Hence, we decided to do the correlation test.

Correlation is the degree to which the pair tend to move together. Below is the correlation matrix for your selected Cryptocurrencies

| | BTC_RET | ETH_RET | XRP_RET | BCH_RET | USDT_RET | LTC_RET | EOS_RET | BNB_RET | XLM_RET |
|---|---|---|---|---|---|---|---|---|---|
| BTC_RET | 1.000000 | 0.814236 | 0.669108 | 0.761667 | 0.172036 | 0.786809 | 0.692565 | 0.584168 | 0.653926 |
| ETH_RET | 0.814236 | 1.000000 | 0.752217 | 0.757754 | 0.102126 | 0.813990 | 0.739462 | 0.558828 | 0.709704 |
| XRP_RET | 0.669108 | 0.752217 | 1.000000 | 0.625250 | 0.023811 | 0.681097 | 0.677222 | 0.457743 | 0.746580 |
| BCH_RET | 0.761667 | 0.757754 | 0.625250 | 1.000000 | 0.091932 | 0.753983 | 0.671952 | 0.483499 | 0.597043 |
| USDT_RET | 0.172036 | 0.102126 | 0.023811 | 0.091932 | 1.000000 | 0.104917 | 0.040406 | 0.088874 | 0.126620 |
| LTC_RET | 0.786809 | 0.813990 | 0.681097 | 0.753983 | 0.104917 | 1.000000 | 0.719158 | 0.587999 | 0.630879 |
| EOS_RET | 0.692565 | 0.739462 | 0.677222 | 0.671952 | 0.040406 | 0.719158 | 1.000000 | 0.508352 | 0.636812 |
| BNB_RET | 0.584168 | 0.558828 | 0.457743 | 0.483499 | 0.088874 | 0.587999 | 0.508352 | 1.000000 | 0.473120 |
| XLM_RET | 0.653926 | 0.709704 | 0.746580 | 0.597043 | 0.126620 | 0.630879 | 0.636812 | 0.473120 | 1.000000 |

Heat map revealing the correlation of our cryptocurrencies


We obtained the following results.

| S/N | PAIRS | CORRELATION | COINTEGRATION |
|---|---|---|---|
| 1 | BTC_RET and ETH_RET | 0.812897 | 0.000000 |
| 2 | BTC_RET and XRP_RET | 0.668008 | 0.000000 |
| 3 | BTC_RET and BCH_RET | 0.760414 | 0.000000 |
| 4 | BTC_RET and USDT_RET | 0.171753 | 0.000000 |
| 5 | BTC_RET and LTC_RET | 0.785515 | 0.000000 |
| 6 | BTC_RET and EOS_RET | 0.691426 | 0.000000 |
| 7 | BTC_RET and BNB_RET | 0.583208 | 0.000000 |
| 8 | BTC_RET and XLM_RET | 0.652851 | 0.000000 |

From the analysis and the values gotten we can see that although pairs all seen to have very low Cointegration value, but we observed that the Ethereum (ETH) had the best correlation value (81.2%). Thus, we will be proceeding with the Bitcoin/Ethereum pair for the next two phases; Price forecasting and developing trade strategy

Since our base currency is Bitcoin, the forecast modelling was done with Bitcoin data.

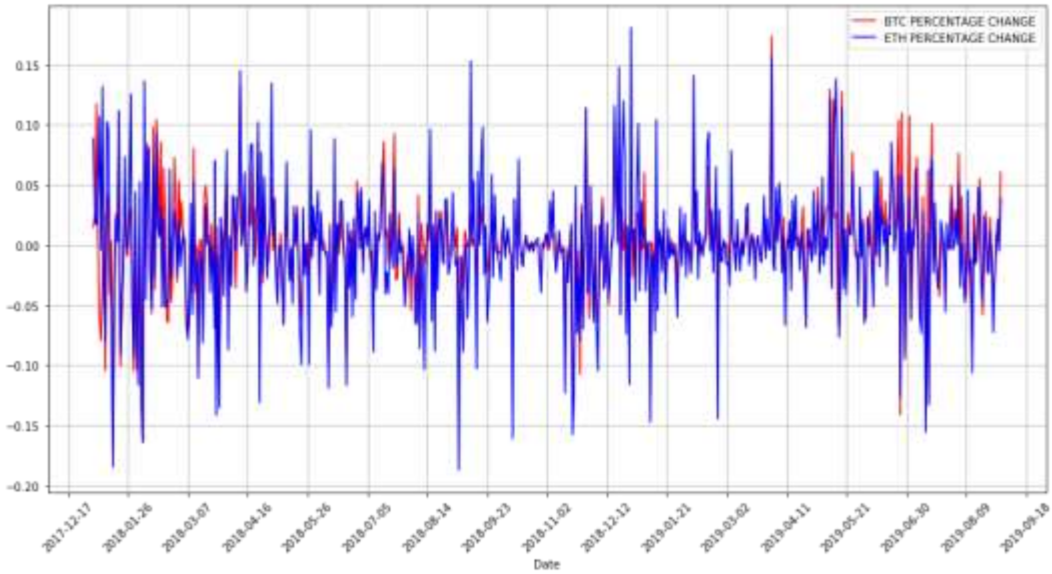
we will now save our data for Bitcoin (BTC) and Ethereum (ETH) as csv files.

Concatenating the adjusted Close for the currencies (Bitcoin (BTC) and Ethereum (ETH))

The Adjusted Closing plot for Bitcoin and Ethereum.

Observing again that values of our individual currencies are far apart, it was imperative for us to see the pattern in their percentage change (or returns).

We calculated the percentage change and below is a graphical representation of our result




The Cumulative returns for both Bitcoin and the Ethereum for the period in view was equally computed and visualized as seen below.


Now the correlation relationship between the Bitcoin returns and the Ethereum returns can be appreciated with this plot.

We took a quick look at the spread between the Bitcoin returns and the Ethereum returns.

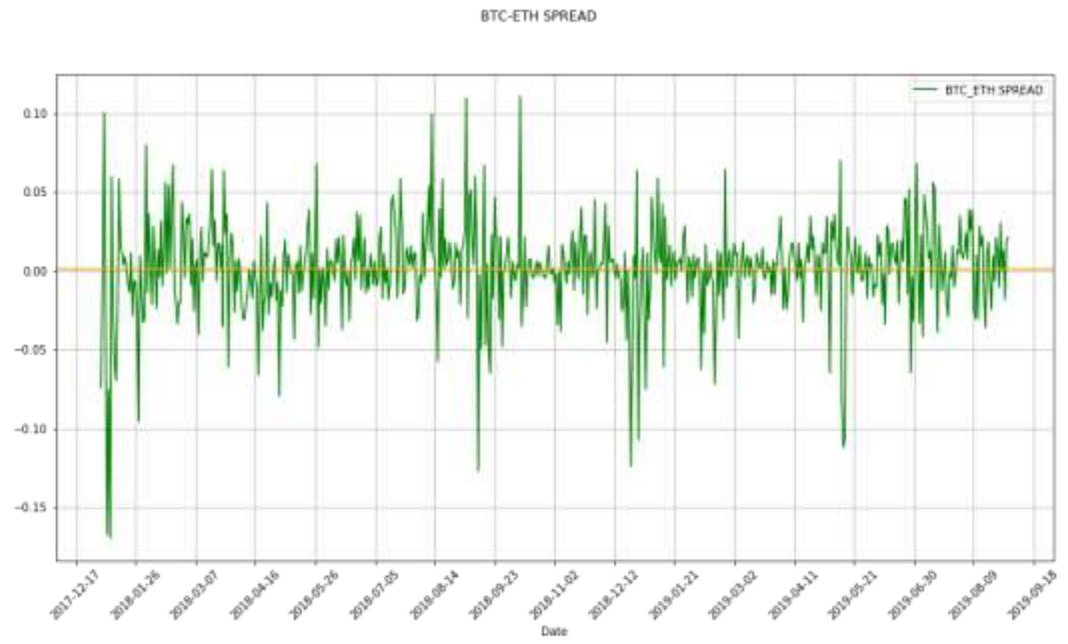
Spread is the difference between the returns.

$$Spead_{BTC\_ETH} = BTC(Returns) - ETH(Returns)$$

The mean of the spread was also considered.

Mean of spread = $\frac{\sum Spead_{BTC\_ETH}}{No.of\ Observations} = \frac{\sum BTC(Returns) - ETH(Returns)}{No.of\ Observations}$

After computing the values, the spread was visualized as shown below. With the mean represented with the orange colored line.


BTC-ETH SPREAD
BTC_ETH SPREAD
0.10
0.05
0.00
−0.05
−0.10
−0.15
2017-12-17
2018-01-26
2018-03-07
2018-04-16
2018-05-26
2018-07-05
2018-08-14
2018-09-23
2018-11-02
2018-12-12
2019-01-21
2019-03-02
2019-04-11
2019-05-21
2019-06-30
2019-08-09
2019-09-18
Date

## B. Forecast Adjusted price for selected pair

Commonly, an analyst may use various predictive tools in other to predict or forecast the trend of price movement. For this assignment we will be using the Long Short-Term Memory model (LSTM), the Moving Average as well as the ARIMA (Auto Regressive Integrated Moving Average).

First, we will consider our accuracy/performance measures

As errors cannot be completed eradicated and fully measured, we will be working with a few on a case to case base to measure the performance of our various predictive models.

In our search for a key performance indicator (KPI), we will be considering the following

- RMSE (Root Mean Square Error)
- Mean Absolute Percentage Error (MAPE)
- Mean Percentage Error (MPE)

Some formulas to help our understanding of our error terms

$$Residual\ Error\ (in\ time)\ , e_t = Forecasted\ value(FV_t) - Actual\ Value(AV_t)$$

$$Absolute\ Error\ (in\ time)\ , |e_t|\ = |Forecasted\ value(FV_t) - Actual\ Value(AV_t)|$$

$$Mean\ Percentage\ Error\ (MPE) = \left(\frac{1}{number\ of\ Observations}\sum \frac{e_t}{AV_t}\right) * 100\%$$

$$Mean\ Absolute\ Percentage\ Error\ (MPE) = \left(\frac{1}{number\ of\ Observations}\sum \frac{|e_t|}{AV_t}\right) * 100\%$$

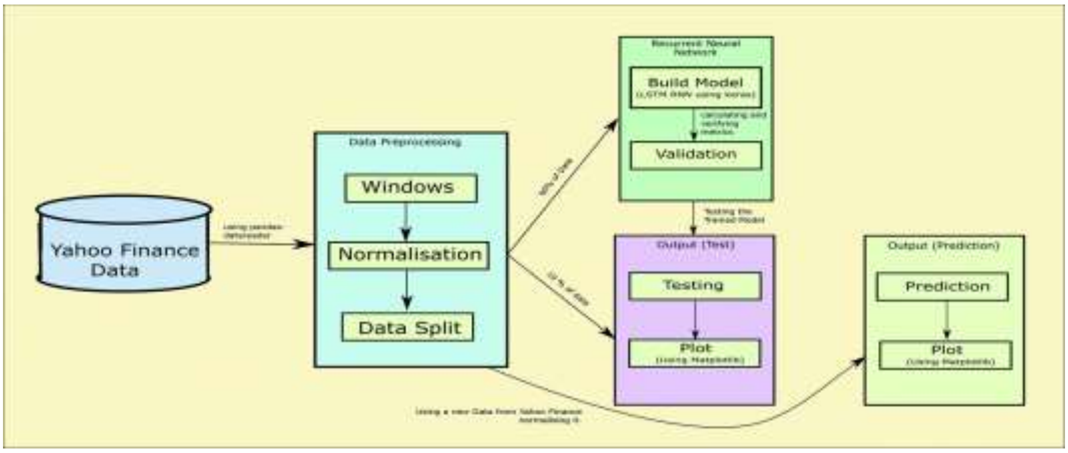
$$Root\ Mean\ Square\ Error(\ RMSE) = \sqrt{\frac{1}{number\ of\ Observations}\sum e_t^2}$$

Data Split;

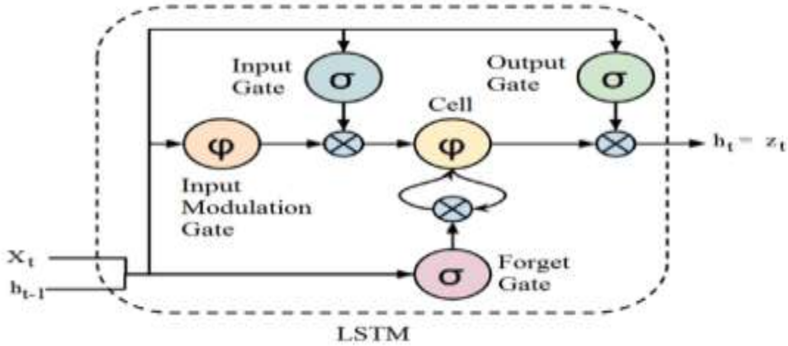
our total number of observations for the time in view is 610, thus our data will be split into training (80%) and testing (20%) and then implemented in the codes as illustrated in the Jupyter notebook. Each having 488 and 122 observations respectively.

**Long Short-term Memory (LSTM)**

LSTMs ("long short-term memory" units) is one of the most remarkable Recurrent neural networks. The LSTM is designed to identify data sequence patterns. Long Short-Term Memory networks are the best quality level to building RNN. and they have demonstrated to be exceptionally powerful for succession forecast problems. By training the models via back-propagation, the LSTM stores useful data while forgetting any that is deemed otherwise. Examples of such data that can be forecasted via LSTM are stock markets, sensors, genomes, words and handwriting.




Sourced from: Kriti Pawar,Raj Srujan Jalem,Vivek Tiwari;" Stock Market Price Prediction Using LSTM RNN (20 November 2018)"[21]

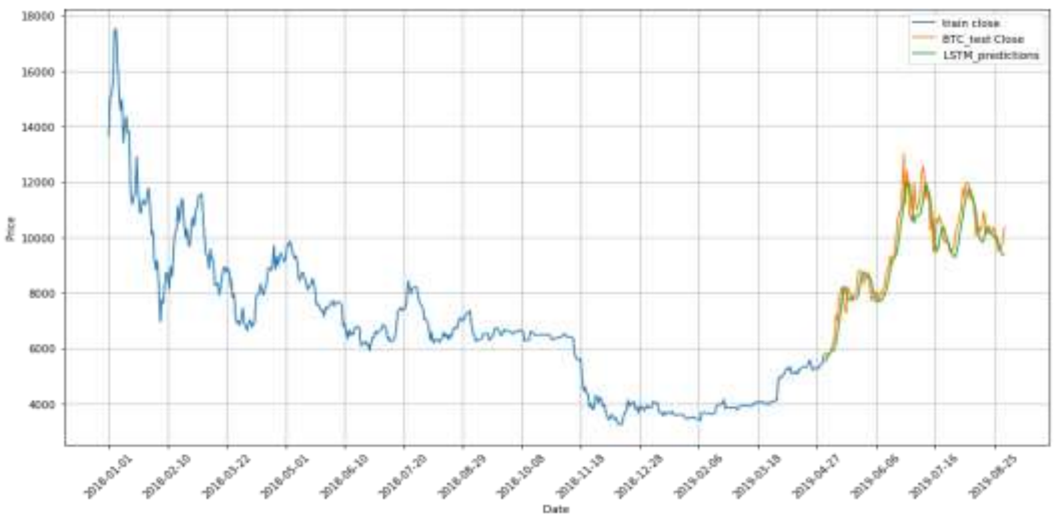



Sourced from: "Eugine Kang "Long Short-Term Memory (LSTM): Concept" Sep 2, 2017"[20]

The LSTM typically consist of the following components,

- The input gate: this point adds data/ information to the cell
- The forget gate: this point eliminates information/ data that is no more no longer essential
- The output gate:  this point picks the information/data that should be presented as output

Below is the outcome of our LSTM forecast



The above seen plot is a product of many trials. A few parameters where adjusted and the accuracy of the model was observed. Our choice LSTM model is called Vanilla LSTM model

It is characterised with its layers as thus:

1. Input layer.
2. Wholly linked LSTM hidden layer.
3. Wholly linked output layer

```
# create and fit the LSTM network
model = Sequential()
model.add(LSTM(units=13, return_sequences=True, input_shape=(x_train.shape[1],1)))
model.add(LSTM(units=3))
model.add(Dense(1))

model.compile(loss='mean_squared_error', optimizer='adam',metrics=['accuracy'])
model.fit(x_train, y_train, epochs=13, batch_size=1, verbose=2)
```

As seen in the snapshot of the code, we finally settled for an inner layer of 13 and out layer of 3. This is an approximate value of 80% of the layers set as inner and the balance 20% set as the outer, following our 80:20 "Training to Test ratio". We choose the "Adam" optimaizer because although it's a one-to-one sequence prediction model, we needed a stochastic optimaizer to check all posible outcomes and learn the pattern. The Epoch was set to 13 to allow for the inner layer to learn the pattern enough to give an acceptable result.

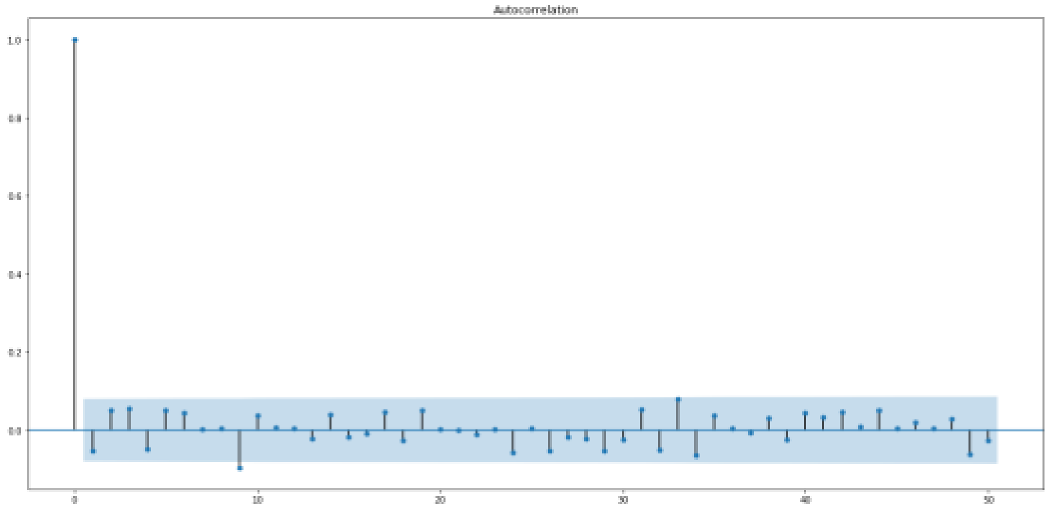
```
print(model.summary())
```

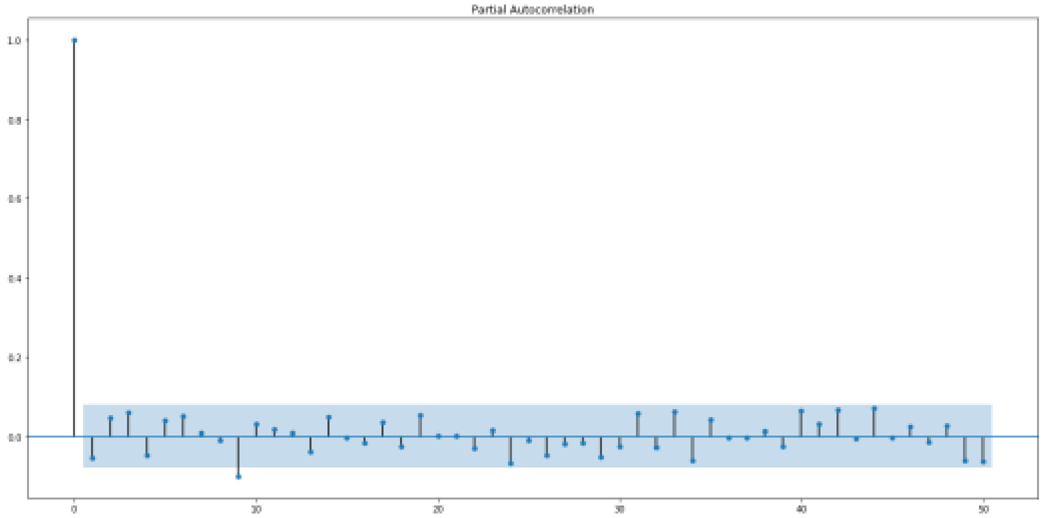
```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_33 (LSTM)               (None, 80, 13)            780
_________________________________________________________________
lstm_34 (LSTM)               (None, 3)                 204
_________________________________________________________________
dense_41 (Dense)             (None, 1)                 4
=================================================================
Total params: 988
Trainable params: 988
Non-trainable params: 0
_________________________________________________________________
None
```

**Preliminary tests for the (Moving-Average and ARIMA models)**

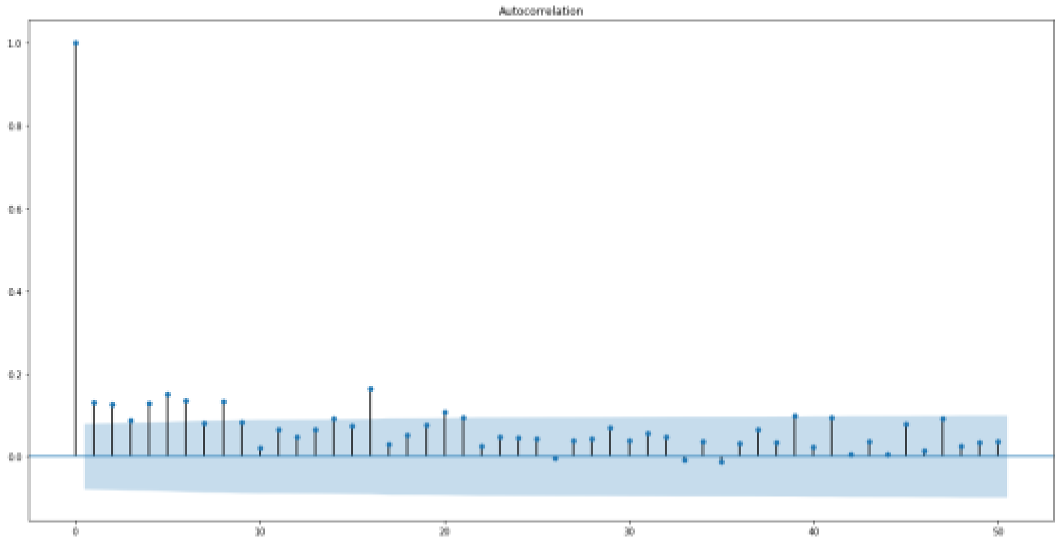
For the next two models (Moving-Average and ARIMA) we have to investigate the ACF and the PACF of the returns and the square of the returns of our time series data. ACF stands for Auto Correlation function and PACF stands for Partial Auto correlation function. Both are meant to measure the level of correlation or similarity of a particular time series with a previous time. This often helps us determine the most relevant part of the Moving-Average and ARIMA model

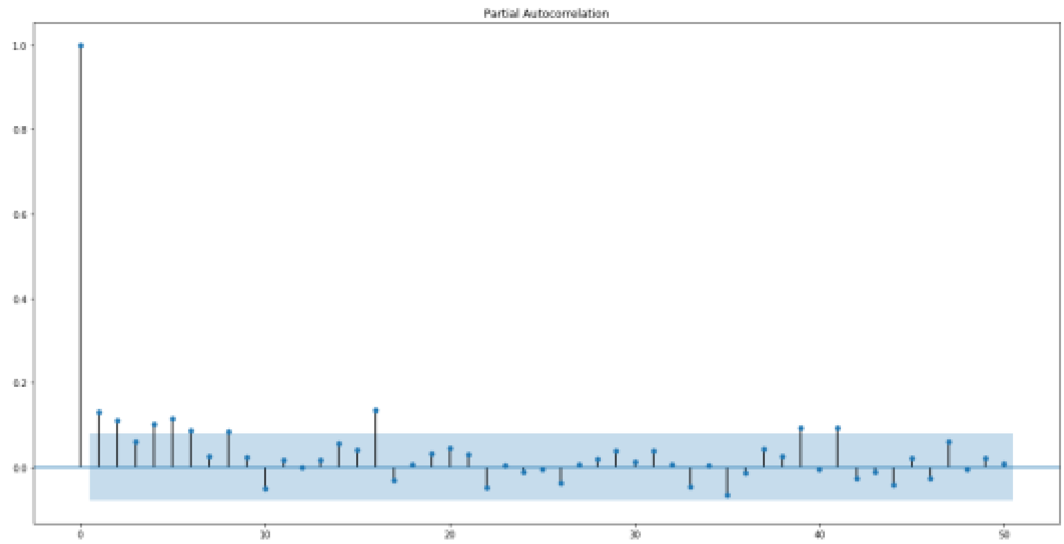
For our time series data; Return on Bitcoin, we had the following plot for the returns




And then we decided to square the returns on Bitcoin and check the ACF and PACF again






**Intuition for the Bitcoin correlograms**

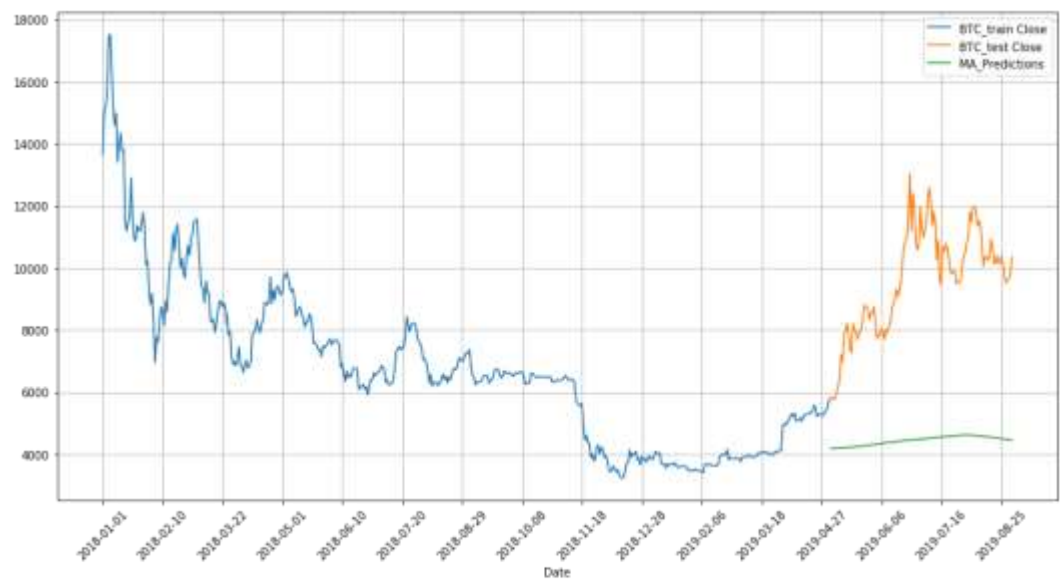
Considering the increase observed in the ACF and PACF of the returns squared, it shows a level of correlation although all still under 0.2, we observed that the data in the first half of the series are better correlated than the later end , where we noticed the indicators are more sparse and shorter in length.
Considering the negative skewness of our results, it suggests an asymmetric model, this could be as a result of product news or seasonal perceptions in the market.

However, ACF and the PACF suggests an AR model.

Thus, we will investigate with the Moving Average Model and then the ARIMA Model.

**MOVING AVERAGE**

In budgetary applications, a Moving Average is the unweighted mean of the past "n" data. In any case, in other applications, the mean is typically taken from an equivalent number of information on either side of a focal worth. This guarantees variations in the mean are lined up with the variations in the information as opposed to being moved in time.



The Moving average by itself may not be able to give a good prediction just as the ACF and PACF had originally suggested, however, we took it a notch higher and decided to try the ARIMA model

**ARIMA**

"ARIMA: Auto-Regressive Integrated Moving Average. It is a class of model that gets a collection of different standard progressive structures in data (within a particular time series)" [18]

ARIMA is an incredibly common genuine methodology for time series analysis. ARIMA models studies the past characteristics to envision the future characteristics. There are three critical parameters in ARIMA:

p (past qualities utilized for forecasting the following worth)

q (past gauge mistakes used to anticipate the future qualities)

d (differencing order)

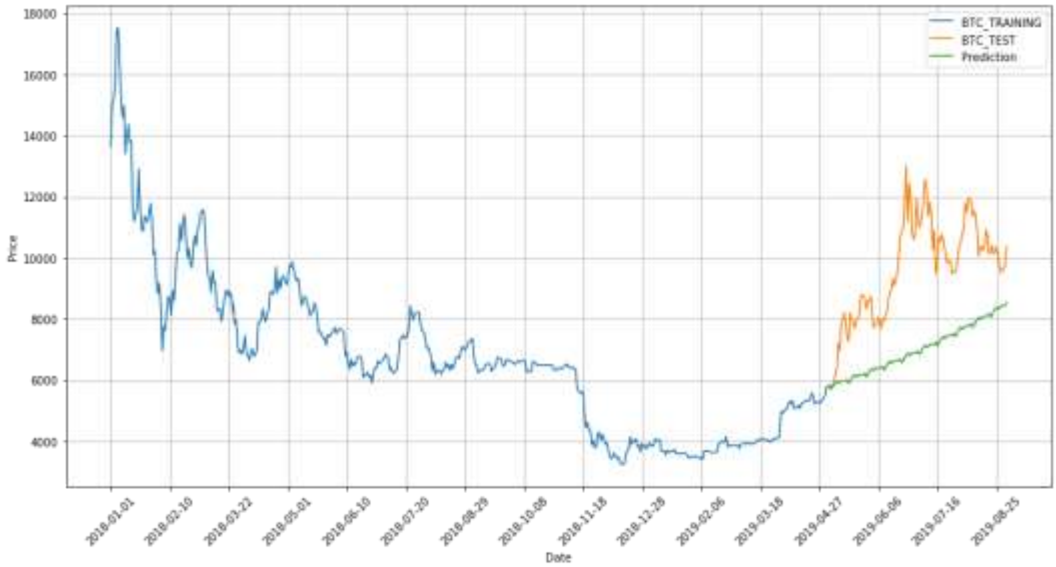
Parameter tuning for ARIMA consumes a huge amount of time. So we will use auto ARIMA which thus picks the best mix of (p,q,d) that gives the least error

Below is a summary of our ARIMA Model

```
print(B_model.summary())
```

```
                                      Statespace Model Results
==========================================================================================
Dep. Variable:                                 y   No. Observations:                  488
Model:             SARIMAX(0, 1, 0)x(1, 1, 1, 12)   Log Likelihood               -3464.411
Date:                            Mon, 16 Mar 2020   AIC                           6936.823
Time:                                    17:07:04   BIC                           6953.476
Sample:                                         0   HQIC                          6943.372
                                            - 488
Covariance Type:                              opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
intercept      1.6581      3.034      0.546      0.585      -4.289       7.605
ar.S.L12      -0.0870      0.027     -3.235      0.001      -0.140      -0.034
ma.S.L12      -0.9391      0.055    -16.948      0.000      -1.048      -0.830
sigma2      1.144e+05   6504.793     17.588      0.000    1.02e+05    1.27e+05
===================================================================================
Ljung-Box (Q):                       68.62   Jarque-Bera (JB):               847.35
Prob(Q):                              0.00   Prob(JB):                         0.00
Heteroskedasticity (H):               0.07   Skew:                            -0.50
Prob(H) (two-sided):                  0.00   Kurtosis:                         9.46
===================================================================================

Warnings:
```

We equally visualized our results in a plot as seen below.

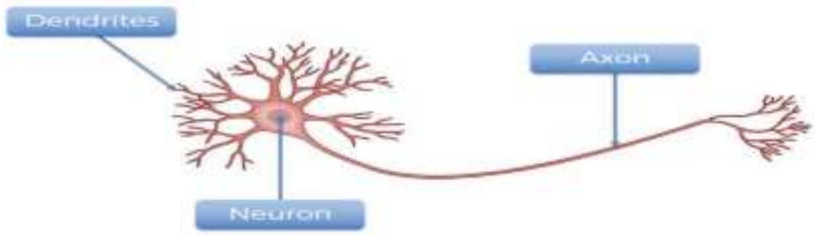


The ARIMA model was time consuming and still was not as productive as expected. The high AIC and BIC values of 6936.823 and 6953.476 respectively with the Log Likelihood of -3464.411 reflects a poorly fitted model

**The Artificial Neural Network (ANN)**

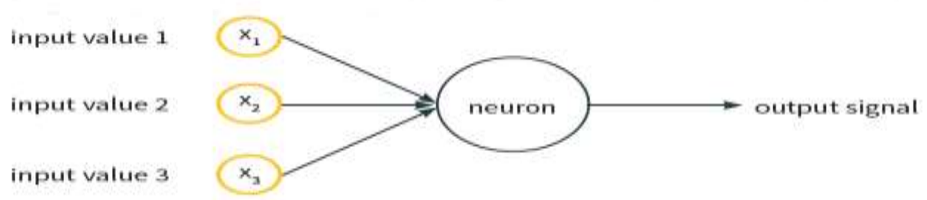
Artificial Neural Network is a data handling model which is utilized to contemplate the conduct of an intricate framework by computer replication. It is incited by the organic method which the human mind processes its data. The key component of the model is the novel structure of the data preparing framework. The objective of artificial neural network is to analyze information in a similar manner a human mind would

With this in mind, lets consider the neutrons found in the human brain and nervous system,




***Sourced from: Quantra E-book, "How to trade using machine learning" [19]***

We can now compare it to the actual ANN cells, as shown below

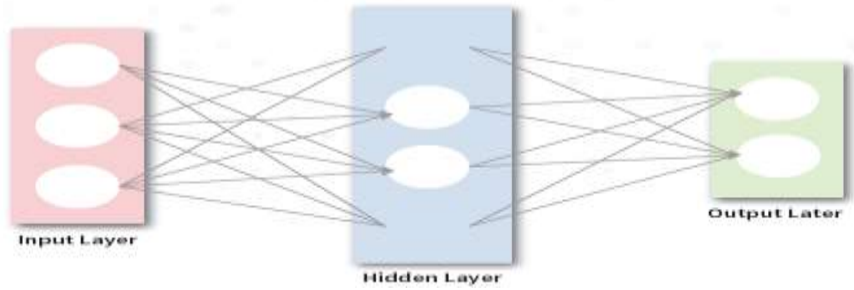



***Sourced from: Quantra E-book, "How to trade using machine learning" [19]***

Neural Networks (ANN's) utilize three neuron layers these are:

1. input layer,
2. hidden or concealed layer, and
3. the Output or yield layers.

These help to get the anticipated forecasts. The artificial nodes are the fundamentally prepared components of artificial neural network




***Sourced from: Quantra E-book, "How to trade using machine learning" [19]***

Similar to the LSTM, we tried tuned the parameters of our ANN a couple of times before we were certain that the results were ok and there was no overfitting. We still maintained our training to testing ratio of 80:20 while we employed the following code

```

modelB = Sequential()
modelB.add(Dense(units = 13, kernel_initializer = 'uniform', activation = 'relu', input_dim = 80))
np.random.seed(0)
modelB.add(Dense(units =13, kernel_initializer = 'uniform', activation      = 'relu'))
modelB.add(Dense(units = 1, kernel_initializer = 'uniform', activation       = 'sigmoid'))
modelB.compile(optimizer = 'adam', loss = 'mean_squared_error', metrics
          = ['accuracy'])
modelB.fit(x_trainB, y_trainB, batch_size = 1, epochs = 13,verbose=2)
```

As seen in the snapshot of the code, we finally settled for an inner layer of 13 and out layer of 1and  a batch size of 1. Unlike the LSTM we couldn't increase the outer layer to 3 as our ANN can only work with an outer Dense(unit) of 1.

We choose the "Adam" optimaizer because we still needed a stochastic optimaizer to check all posible outcomes and learn the pattern. The Epoch was set to 13 to allow for the inner layer to learn the pattern enough to give an acceptable result.

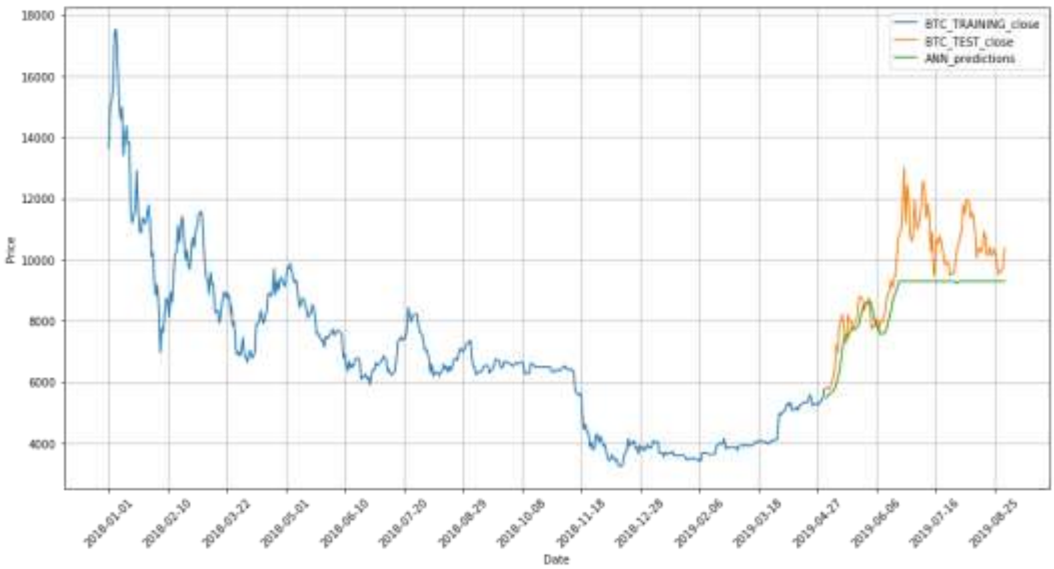
With the ANN model, we had the following results

ANN_ Model Summary:

```
print(modelB.summary())


```

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense_23 (Dense)             (None, 13)                1053
_________________________________________________________________
dense_24 (Dense)             (None, 13)                182
_________________________________________________________________
dense_25 (Dense)             (None, 1)                 14
=================================================================
Total params: 1,249
Trainable params: 1,249
Non-trainable params: 0
_________________________________________________________________
None
```

And visualized our model




The ANN model has some good success at the beginning but at the later part of the forecasting seam to overestimate. This suggests more epoch or inner Dense(unit) might give a better result but it was time consuming.

**Comparing Our Models**

For better technical analysis of our four models, below is a table of performance reflecting and estimated value for each function

| Model | MPE | MAPE | RMS | Correlation | Relevance on plot |
|---|---|---|---|---|---|
| LSTM | 0.011241 | 0.2046566 | 2323.89 | 94.7% | High |
| Moving Average | -0.525507 | 0.5255070 | 5408.40 | 81.70% | Low |
| ARIMA | -0.24349 | 0.24349 | 2764.407 | 65.7% | Mid |
| ANN | -0.08885 | 0.18821 | 2228.55 | 83.9% | High |

The ANN has the lowest MAPE (Mean Absolute Percentage Error) of all four models (≈ 0.19), the lowest RMS (Root mean Square Error) of all four models (≈ 2228.6) but not the highest percentage for correlation (≈ 84%)

The LSTM on the other hand save for the ANN model had the least MAPE and RMS values of 0.011241 and 2323.89 respectively. However, it had the highest correlation with the sample to be predicted (≈ 95%).

With more epoch and inner Dense(unit) the ANN model would perform better and give a higher correlation, but that will result in higher time consumption and a higher risk of over-fitting .

For our project we would go with the LSTM model

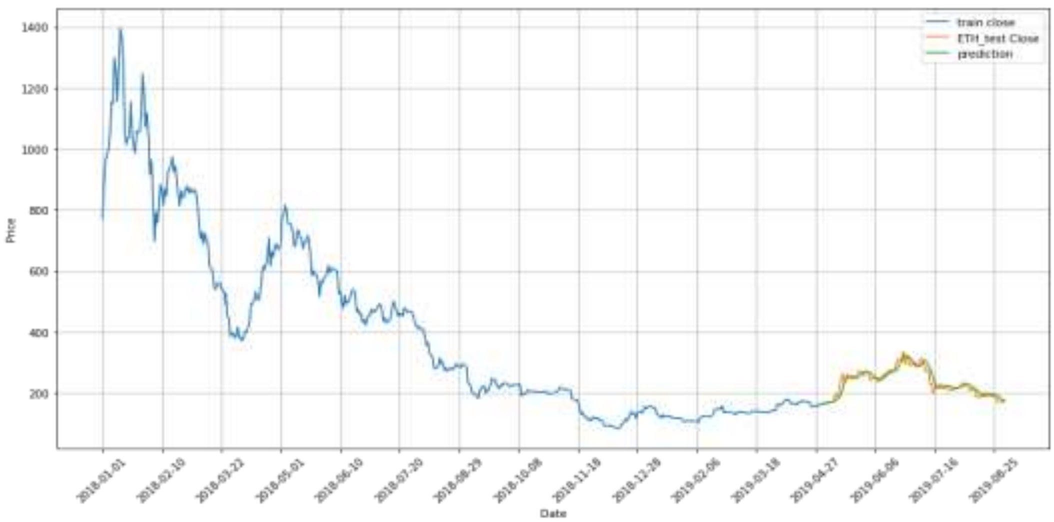
Hence, we ran LSTM for Ethereum, maintaining the same parameters we used for the Bitcoin.

Model summary

```
print(modele.summary())
```

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_9 (LSTM)                (None, 80, 13)            780
_________________________________________________________________
lstm_10 (LSTM)               (None, 3)                 204
_________________________________________________________________
dense_26 (Dense)             (None, 1)                 4
=================================================================
Total params: 988
Trainable params: 988
Non-trainable params: 0
_________________________________________________________________
None
```

Our LSTM model plot for Ethereum



Below is a summary of the estimated value for the error term and the correlation

| Model | MPE | MAPE | RMS | Correlation | Relevance on plot |
|---|---|---|---|---|---|
| LSTM on ETH | 0.04123 | 0.2127 | 59.69 | 93.2% | High |

**Cryptocurrency Price Forecasting Conclusion**

Evidently, we will adopt the LSTM as it births the best result for our assignment

Machine learning and deep learning techniques can possibly facilitate the entire procedures of investigating enormous pieces of information, spotting critical examples and producing a solitary yield that explores investors towards a specific choice dependent on anticipated resource costs. We perceive that numerous Financial Engineers and Researchers are spending restless evenings attempting to improve the presentation of Machine Learning and Deep learning models.

Kindly note that crypto currency price can be influenced by numerous other profoundly interrelated elements. these elements include but not limited to;

- Government Policies,
- Exchange/ Trading Platform utilized for the Trade,
- Demand and supply for Bitcoin and/or Ethereum,
- Availability of other (Competing) Crypto-Currencies

News on Common components like environmental change, dry season, tropical storms, floods, tremors and different elements like demonetization or merger/demerger of the associations could likewise influence the price of Bitcoin and/or Ethereum.

Huge numbers of these variables can be hard to envision and may require extra instruments to appropriately estimate them.

As at the time of this findings (March 2020) the following are news to consider as it pertains to our Choice Cryptocurrencies (Bitcoin and Ethereum)

- Dan Caplinger (a publisher for NASDAQ) on March 13, 2020, published an article that reads “Bitcoin Just Failed the Coronavirus Test”. He said in his article; “Bitcoin has taken speculators on a thrill ride, and in spite of the fact that its’ worth has risen significantly since its commencement over 10 years prior, through the years, it's had a good share of its downward movements. The most recent dive in bitcoin, however, comes when many would've figured the digital currency would be well on the way to take off: in the wake of the worldwide COVID-19(Coronavirus) pandemic” [22]
- According to a Blockchain Analyst and Writer, Vladislav Sopov, of u.today “With six new instances of coronavirus among EthCC meeting members affirmed yesterday, suspending major crypto events doesn't appear to be an overcompensation any longer.” [23]

This news will definitely affect the prices of both Bitcoin and Ethereum but the degree to which the change will happen cannot be forecasted by our LSTM model or any other model we have considered.

## C. **Developing a proposed pair trading strategy**

**The strategy**

The strategy adopted for this project is a pair- trading strategy. This involves buying and selling of two Cryptocurrencies that are highly correlated.

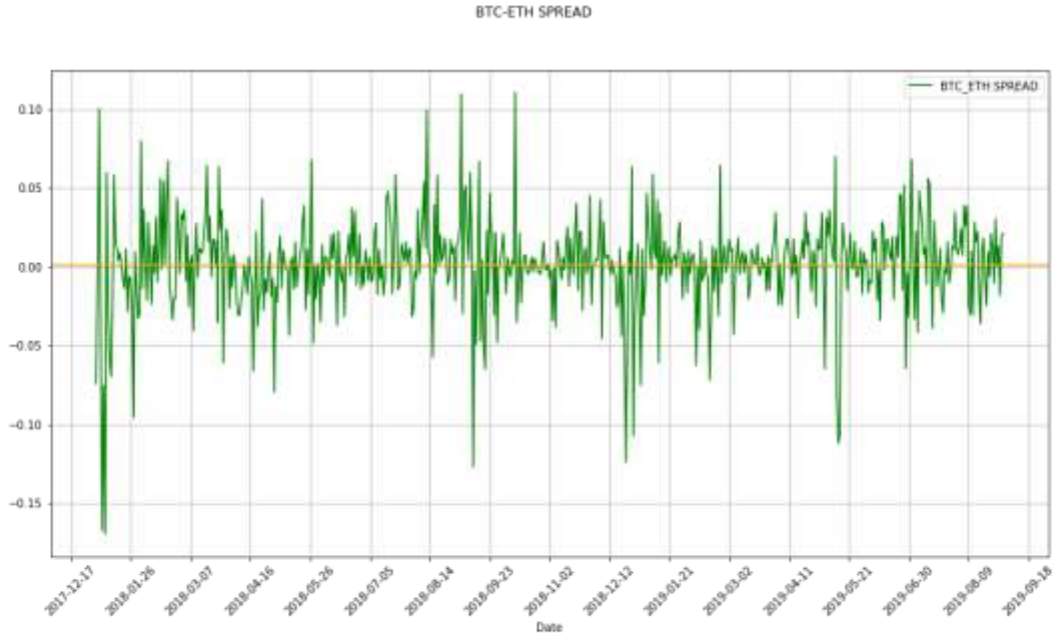
We have already ascertained the best pair for our choice cryptocurrency (Bitcoin) to be Ethereum because it had the highest correlation from the set of Cryptocurrencies considered.

This strategy is a mean-reverting, market neutral strategy. This means that prices of the paired cryptocurrencies will oscillate about a mean position. So, whether there is an upward trend or a downward trend, it doesn't really affect the strategy as both currencies are highly correlated and would trend in similar direction. This helps the retailer to focus more on just the pair and not on the entire market noise.

To continue on this project, we need to run the Engle-Granger Co-integration test. Using the Static regression, the Engle-Granger technique first develops residuals (errors) or spreads. The residuals are tried for the nearness of unit roots utilizing ADF or a comparative test. On the off chance that time series is cointegrated, at that point the residuals will be basically stationary. The Engle-Granger Co-integration test has two main components; Augmented Dickey-Fuller Test and the Phillips-Perron Test.

Augmented Dickey-Fuller Test
This test consists of evaluating whether paired assets prices spread, individual assets prices and individual assets prices differences time series are mean stationary with null hypothesis that they have a unit root and are not stationary
Phillips-Perron Test
This test builds on the Augmented Dickey-Fuller Test, it tests for null hypothesis in a series of integrated order of 1

The Spread was computed and plotted as shown below:



For this exercise we split our data; into training (80%) and testing (20%) and then implemented in the codes as illustrated in the Jupyter notebook. Each having 488 and 122 observations respectively

From our plot we observed some spikes that reveals a great difference from the mean of zero
The results of the test are as shown below

```
== BTC Prices Augmented Dickey-Fuller Test ==

   Augmented Dickey-Fuller Results
=====================================
Test Statistic                -13.495
P-value                         0.000
Lags                                2
-------------------------------------

Trend: Constant and Linear Time Trend
Critical Values: -3.97 (1%), -3.42 (5%), -3.13 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.

== ETH Prices Augmented Dickey-Fuller Test ==

   Augmented Dickey-Fuller Results
=====================================
Test Statistic                -12.928
P-value                         0.000
Lags                                2
-------------------------------------

Trend: Constant and Linear Time Trend
Critical Values: -3.97 (1%), -3.42 (5%), -3.13 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.
```

```
== BTC Prices  Differences Augmented Dickey-Fuller Test ==

   Augmented Dickey-Fuller Results
=====================================
Test Statistic                -10.050
P-value                         0.000
Lags                               17
-------------------------------------

Trend: Constant and Linear Time Trend
Critical Values: -3.97 (1%), -3.42 (5%), -3.13 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.

== ETH Prices Differences Augmented Dickey-Fuller Test ==

   Augmented Dickey-Fuller Results
=====================================
Test Statistic                -11.049
P-value                         0.000
Lags                               15
-------------------------------------

Trend: Constant and Linear Time Trend
Critical Values: -3.97 (1%), -3.42 (5%), -3.13 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.
```

```
== BTC-ETH Spread Augmented Dickey-Fuller Co-Integration Test ==

   Augmented Dickey-Fuller Results
=====================================
Test Statistic                 -9.570
P-value                         0.000
Lags                                4
-------------------------------------

Trend: Constant and Linear Time Trend
Critical Values: -3.97 (1%), -3.42 (5%), -3.13 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.

== BTC-ETH Spread Phillips-Perron Co-Integration Test ==

     Phillips-Perron Test (Z-rho)
=====================================
Test Statistic               -650.396
P-value                         0.000
Lags                               19
-------------------------------------

Trend: Constant and Linear Time Trend
Critical Values: -28.98 (1%), -21.50 (5%), -18.09 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.
```

The above results tell us that there is a unit root in our time series (Spread)

We had earlier taken the spread to simply be the difference of the returns of the two cryptocurrencies. However, considering the actual price of the Bitcoin and the Ethereum, it has become imperative to take a more technical approach to account for the difference of weight attached to each Cryptocurrency in our pair-trading portfolio.

The table below is an extract of the prices of the Bitcoin and the Ethereum.

| Date | btc | eth |
|---|---|---|
| 2018-01-01 | 13657.200195 | 772.640991 |
| 2018-01-02 | 14982.099609 | 884.443970 |
| 2018-01-03 | 15201.000000 | 962.719971 |
| 2018-01-04 | 15599.200195 | 980.921997 |
| 2018-01-05 | 17429.500000 | 997.719971 |

The pairs spread consists of pairs cryptocurrency prices linear regression residual or forecasting errors. It corresponds to being long (buying one) crypto currency and being short (selling) the other. Depending on their co-integration relationship.

$$S_t = y_t - \beta x_t$$

$$S_t = spread\ at\ time\ t$$
$$y_t = bitcoin\ price\ at\ time\ t$$
$$\beta = zero\ intercept\ linear\ regression\ coefficient\ at\ time\ t$$
$$x_t = Ethereum\ price\ at\ time\ t$$

$we\ \ got\ the\ value\ of\ \beta\ to\ be\ approximatly$ 0.61 this is also know as the hedge

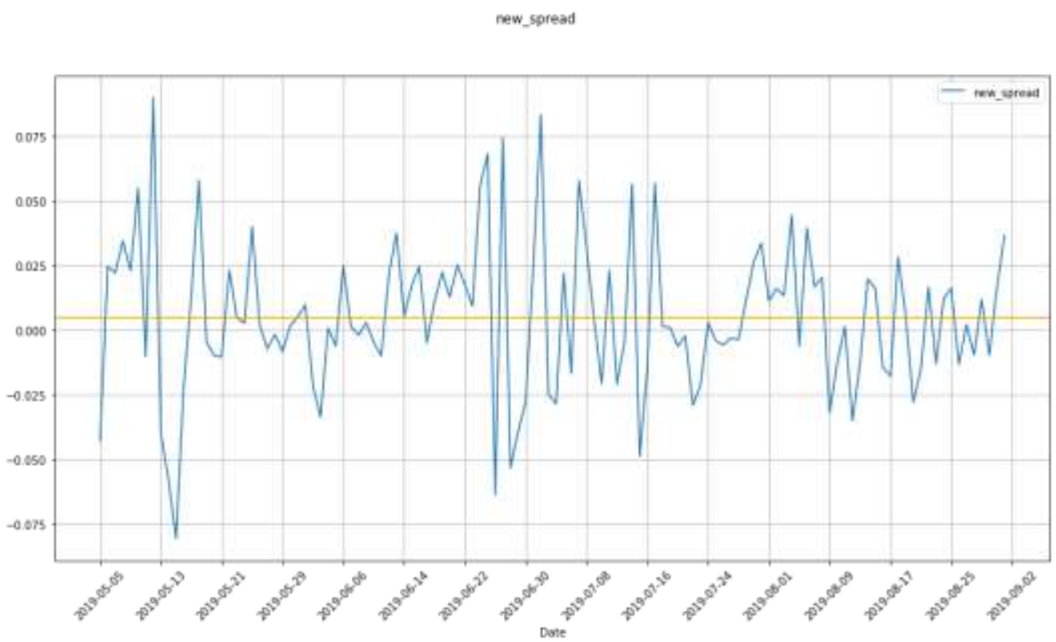
We did the “BTC-ETH” Augmented Dickey-Fuller Co-Integration Test afresh for our new Spread and obtained the following results

```
== BTC-ETH Spread Augmented Dickey-Fuller Co-Integration Test ==

   Augmented Dickey-Fuller Results
=====================================
Test Statistic                -10.822
P-value                         0.000
Lags                                0
-------------------------------------

Trend: Constant and Linear Time Trend
Critical Values: -4.04 (1%), -3.45 (5%), -3.15 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.

== BTC-ETH new_spread Phillips-Perron Co-Integration Test ==

     Phillips-Perron Test (Z-rho)
=====================================
Test Statistic                -97.898
P-value                         0.000
Lags                               13
-------------------------------------
```

Again, the process (Spread) is still weekly stationary with a unit root. This means it is not completely mean reverting

We took it a step further and calculated for the mean of the Spread and plotted the spread representing the mean with an orange line

Below is our Z-score for the spread visualized.




As observed above, the spread is seen to oscillate around its arithmetic mean (0.0048).

Considering our signal generation, we will then take the Z-score of the Spread.

$$Z(S)_t = \frac{S_t - \mu S_t}{\sigma S_t}$$

$$Z(S)_t = Z - score\ of\ the\ spread\ in\ time$$

$$S_t = spread\ in\ time$$

$$\mu S_t = mean\ of\ the\ spread\ in\ time$$

$$\sigma S_t = standard\ deviation\ of\ the\ Spread\ in\ time$$

This Pairs Trading Strategies consists of paired Cryptocurrency prices rolling spread normalized time series or z-score, associated trading signals and corresponding trading positions.

To better capture the essence of the Bitcoin-Ethereum price trends, we needed to pick a relatively large size window. we picked the tenth term in the Fibonacci Sequence, (55)

We used a 55-day window to generate our rolling spread. The reason for a Rolling Spread is to avoid bias. We deemed it necessary to use a relatively large window because of the high volatility associated with Cryptocurrencies (Bitcoin in particular)

Again, for the rolling Z-score,

$$Z(S)_{(wind0w=55)} = \frac{S_t - \mu S_{(wind0w=55)}}{\sigma S_{(wind0w=55)}}$$

$$Z(S)_{(wind0w=55)} = Z - score\ of\ the\ spread\ (with\ 55 - day\ window)$$

$$S_t = spread\ in\ time$$

$$\mu S_{(wind0w=55)} = mean\ of\ the\ spread\ (with\ 55 - day\ window)$$

$$\sigma S_{(wind0w=55)} = standard\ deviation\ of\ the\ Spread\ (with\ 55 - day\ window)$$

Just as additional precaution, we did the Co-Integration Tests on the rolling Z-score $(with\ 55 - day\ window)$

The result was as shown

```
== BTC-ETH Spread Augmented Dickey-Fuller Co-Integration Test ==

   Augmented Dickey-Fuller Results
=====================================
Test Statistic                 -5.130
P-value                         0.000
Lags                                2
-------------------------------------

Trend: Constant and Linear Time Trend
Critical Values: -4.11 (1%), -3.48 (5%), -3.17 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.

== BTC-ETH new_spread Phillips-Perron Co-Integration Test ==

     Phillips-Perron Test (Z-rho)
=====================================
Test Statistic                -50.193
P-value                         0.000
Lags                               11
-------------------------------------

Trend: Constant and Linear Time Trend
Critical Values: -26.09 (1%), -19.85 (5%), -16.90 (10%)
Null Hypothesis: The process contains a unit root.
Alternative Hypothesis: The process is weakly stationary.
```

Again, the results show a unit root is still present. It is however weakly stationary. We then calculated the mean of the rolling Z-score ( -0.0895) and then visualizes the (55-day window) rolling Z-score


We then observed from the chart that the spread tends to have oscillations between +1 and -1 and more closely between +0.5 and -0.5.

The little orange line under zero marks the mean position of the rolling z-score (-0.0895). This informs our decision to choose our inner band at +0.5 and -0.5 while we take +1 and -1 as our outer band.

With the threshold bands ascertained, and the chart above we went ahead to design our strategy signals;

| S/N | Pair Trade | Operation Description | Conditions for **Z**-score $Z(S)_t$ | Assigned signal |
|---|---|---|---|---|
| 1 | Long Pair Spread Trade | Enter the Long/Buy Signal | $Z(S)_{t-2} > -1 \rightarrow Z(S)_{t-1} < -1$ | -2 |
| 2 | Long Pair Spread Trade | Exit the Long/Buy Signal | $Z(S)_{t-2} < -0.5 \rightarrow Z(S)_{t-1} > -0.5$ | -1 |
| 3 | Short Pair Spread Trade | Enter the Short/Sell Signal | $Z(S)_{t-2} < +1 \rightarrow Z(S)_{t-1} > +1$ | 2 |
| 4 | Short Pair Spread Trade | Exit the Short/Sell Signal | $Z(S)_{t-2} > +0.5 \rightarrow Z(S)_{t-1} < +0.5$ | 1 |

BTC_ETH Trading Strategy Positions

| S/N | Signal | Position | Directions |
|---|---|---|---|
| 1 | -2 | +1 | Long Spread (Buy Bitcoin and sell Ethereum) |
| 2 | -1 | 0 | Do nothing, Maintain previous position |
| 3 | +2 | -1 | Short Spread (Buy Ethereum and sell Bitcoin) |
| 4 | +1 | 0 | Do nothing, Maintain previous position |

We ran this logic and observed the following as our last 15 trades

| Date | BTC_ETH_ROLL_RET | ftestsig2 | ftestpos2 |
|---|---|---|---|
| 2019-08-18 | 0.743537 | 0.0 | 0.0 |
| 2019-08-19 | 0.081729 | -1.0 | 0.0 |
| 2019-08-20 | -1.064928 | 1.0 | 0.0 |
| 2019-08-21 | -0.611179 | -2.0 | 1.0 |
| 2019-08-22 | 0.482117 | 0.0 | 1.0 |
| 2019-08-23 | -0.633690 | -1.0 | 0.0 |
| 2019-08-24 | 0.282996 | 0.0 | 0.0 |
| 2019-08-25 | 0.442265 | -1.0 | 0.0 |
| 2019-08-26 | -0.662026 | 0.0 | 0.0 |
| 2019-08-27 | -0.047222 | 0.0 | 0.0 |
| 2019-08-28 | -0.561268 | -1.0 | 0.0 |
| 2019-08-29 | 0.358123 | 0.0 | 0.0 |
| 2019-08-30 | -0.567011 | -1.0 | 0.0 |
| 2019-08-31 | 0.609149 | 0.0 | 0.0 |
| 2019-09-01 | 1.511466 | -1.0 | 0.0 |

We noticed that at 2019-08-19, the signal reads -1 and the corresponding position was 0, same was observed on the 2019-08-20 entry, with +1 the position remained 0. We see a difference in the position when we have a -2 on the 2019-08-21 entry; the position changed to 1 and remained so till 2019-08-23 recorded a -1 in its signal column reverting the position back to 0.With the corresponding signals, we can conclude that our trading logic is working well.

### Trading Strategy Implementation

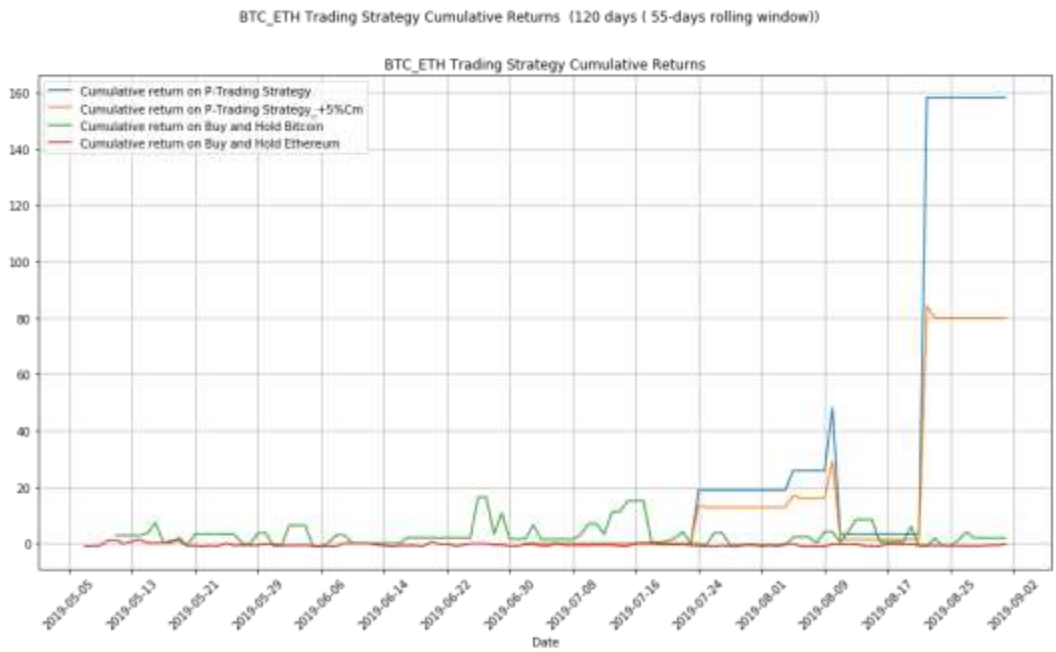
The nature of our asset in view (Cryptocurrency) is very volatile. We tested our strategy on a 120-day period. Bearing in mind that the Cryptocurrency exchangers are open and can be traded all week long (7-days a week)

Four scenarios were considered:

1. Trading the developed pair-trading strategy without any commission or fee ('P-Trading Strategy')
2. Trading the developed pair trading strategy with 5% commission/fee on every trade ('P-Trading Strategy_+5%Cm')
3. Buy and hold Bitcoin ('Buy&Hold Bitcoin')
4. Buy and hold Ethereum ('Buy&Hold Ethereum')

"The fees / commission varies from 0.1% to as much as 5.9% when using credit card". [14] The percentage allotted for fees / commission for our assignment is 5%.

The Chart below shows the cumulative returns of all for scenarios for our test period (120 days)




**Discussion**

The accompanying results that measured the performance of all scenarios are as follows:

| BTC_ETH Strategy Performance Summary | | | | |
|---|---|---|---|---|
| Test-Period | P-Trading Strategy | P-Trading Strategy_+5%Cm | Buy&Hold Bitcoin | Buy&Hold Ethereum |
| Return | 158.133 | 79.9924 | 2.0216 | -0.309 |
| Standard Deviation | 57.2392 | 57.2327 | 116.55 | 1069.04 |
| Sharpe Ratio (Rf = 0%) | 2.7627 | 1.3977 | 0.0173 | -0.0003 |

The following were also deduced from our Pair-Strategy performance.

- Highest positive daily return is 158.13
- Number of positive daily return is 61
- The Percentage of positive returns with this strategy for the period is 50.83 %
- The 120 Trade Day Return of 15813.32 %

First, we observe our Sharpe Ratio.

The Sharpe ratio is another means of comparing investments.

"The Sharpe ratio represents how much the return on a particular asset rewards the investor for the level of risk involved in the investment" [15].

Thus, the higher the ratio, the better the reward.

$$Sharpe\ Ratio = \frac{R_p - R_f}{\sigma_p}$$

$R_p$ = Expected Portfolio Return
$R_f$ =Risk Free Rate (we assumed a 0% risk free rate)
$\sigma_P$ = Portfolio Standard Deviation"

This is $\approx 2.8$ for the P-Trading Strategy (considered as very good) [16], and $\approx 1.4$ ( optimal) for P-Trading Strategy_+5%Cm (the Pair Strategy with 5% Commission / fees).

Clearly, we can see the impact of the 5% commission / fee charged on the second scenario, this births a lower Sharpe ratio and a lower return of about 8,000% as against the scenario with no commission of 15,813% returns.

The buy and hold Bitcoin had a Sharpe ratio of 0.0173 which is considered as below optimal however, it gave a return of about 202% which is a lot more profitable than Buy and Hold Ethereum.

Buy and Hold Ethereum had a Sharpe ratio of -0.0003 and a return of $\approx -31\%$. Evidently is not profitable to buy and hold Ethereum for the period in view.

We used 55 as our Rolling z-score window which is approximately the value of the standard deviation for our modeled strategy.

The Percentage of positive returns (Profitability)

$$\%Profitability = \frac{winning\ Trades}{Total\ Trades} * 100$$

%Profitability with this strategy for the period is 50.83 %

To serve as a guide for trading platforms, as updated by Andrew Munro and Tim Falk

In their 17 September 2019 update of "Best cryptocurrency exchanges" on Finder.com [17], here are some exchangers that can be considered for cryptocurrency trade.

| S/n | Exchange | Fiat currencies | Bitcoin & Ethereum |
|---|---|---|---|
| **1** | **eToro** | **USD** | **Both** |
| 2 | Kraken | USD, EUR, GBP,JPY, CAD,CHF | Both |

| | | | |
|---|---|---|---|
| 3 | Bitmex | - | Both |
| 4 | Coinmama | USD, EUR | Both |
| 5 | Paxful | DZD, XCD, ARS, AMD, AWG, AUD, AZN, BSD, BHD, BDT & 140+ more | Bitcoin only |
| 6 | Paybis | USD, EUR, GBP, CAD, JPY | Both |
| 7 | CoinSwitch | USD | Both |
| 8 | Coinbase | USD, EUR, GBP | Both |
| 9 | CEX.IO | USD, EUR, GBP, RUB | Both |
| 10 | Binance.US | USD | Both |
| 11 | Cointree | - | Both |
| 12 | bitFlyer | USD, EUR, JPY | Both |
| 13 | Changelly | USD, EUR, GBP | Both |
| 14 | KuCoin | - | Both |
| 15 | YoBit | USD, RUB | Both |
| 16 | xCoins | USD | Bitcoin Only |
| 17 | Huobi | GBP, USD, AUD, EUR, INR, BRL, KHR, CAD,CNY, NZD & 15+MORE | Both |
| 18 | SatoshiTango | ARS, USD, EUR | Both |
| 19 | Coinbase Pro | USD, EUR, GBP | Both |
| 20 | Bitstamp | USD, EUR | Both |

# Project Conclusion

The Correlation of the pair is the major basis on which the Pair trading is founded, thus we tend to observe prices reverting to their mean positions. There is no limit to the time at which the strategy should be implemented. Correlations should be studied carefully to guide against wrong assumptions.

Pair-trading with cryptocurrency increases the trader’s advantage. First with the fact that he/she has all the benefits of the cryptocurrency:

- Bitcoin and Ethereum transactions are irreversible
- Bitcoin and Ethereum are liquid
- Bitcoin and Ethereum are finite
- Bitcoin and Ethereum payments are decentralized
- Bitcoin and Ethereum transactions are inexpensive
- Bitcoin and Ethereum transactions are secure and anonymous

And is well able to hedged against loss with the pair trading strategy.

The objective of this task is to build up a model at anticipating Cryptocurrency price, utilizing AI and deep learning strategies that can be utilized by individual financial specialists or merchants who don't have the fundamental innovation to assemble such systems.

These tools introduced here could be excellent, but there is no model that fits all trading situation. Every model has only a percentage of profitability. It is never 100%. In the case studies we looked it our profitability was only about 51% but yet had a fantastic return on investment.

The tools cannot replace the individual’s trading skills, intuition and perception of the market. The tool at its best can forecast based on historical data and cannot forecast any new pattern in the market that is post-data. It is then imperative that the trader learns to train himself and in turn find the guide illustrated in this report useful.

As mentioned earlier in our Price forecast section, news on Common segments like natural change, dry season, hurricanes, floods, tremors, demonetization or merger/demerger of the affiliations, news of analysts and expert views on the performance of Bitcoin and Ethereum under a pandemic situation such as Coronavirus (which has disrupted the world economic activities as at the time of the report, March 2020) could moreover impact the cost of Bitcoin or potentially Ethereum.

The immense quantities of these factors can be difficult to imagine and may require additional instruments to suitably gauge them.

# Recommendations

While on this project we discovered a few limitations of our scope and thus give some recommendations for further work

- Cryptocurrencies are very volatile assets and could be traded in fractions of seconds. Thus, an intra-day data would better serve a retailer that is “interested in minutes-trade”
- The Cryptocurrency universe is growing and while considering pair trading, multiple pairs could be considered to create a trading portfolio.

We propose to open source our design to offer a chance to any financial analysts or Financial Engineers with similar interest to improve this work.

# References


1. https://medium.com/@onlyincrypto/how-cryptocurrencies-influence-the-global-financial-market-c726b1ef09ae accessed 11 February 2020
2. https://www.investors.com/news/technology/algorithmic-trading-algo-trading/ accessed 11 February 2020
3. Do, B., Faff, R. and Hamza, K., A New Approach to Modeling and Estimation for Pairs Trading. Monash University. 2006
4. Elliott, R., van der Hoek, J. and Malcolm, W., *Pairs Trading*. Quantitative Finance, Vol. 5, pp. 271-276. 2005.
5. Gatev, E., Goetzmann, W. and Rouwenhorst, K. G., Pairs Trading:
6. Performance of a Relative Value Arbitrage Rule. Working papers, Yale School of Management. 1999.
7. Gatev, E., Goetzmann, W. and Rouwenhorst, K. G., Pairs Trading:
8. Performance of a Relative Value Arbitrage Rule. Review of Financial Studies, Vol. 19, pp. 797-827. 2007.
9. Nath, P., *High Frequency Pairs Trading with U.S Treasury Securities: Risks and Rewards for Hedge Funds*. Working paper, London Business School. 2003
10. Vidyamurthy, G., *Pairs Trading, Quantitative Methods and Analysis*. John Wiley & Sons. Canada. 2004.
11. Shleifer, A., and R. Vishny. (1997). "The limits of arbitrage." *Journal of Finance* 52:35-55.
12. https://steemit.com/blockchain/@jsherm/the-advantages-to-cryptocurrency-over-fiat-money accessed 11 February 2020
13. Amihud, Y. (2002). "Illiquidity and stock returns: Cross-section and timeseries effects." Journal of Financial Markets 5 (1): 31-35.
14. https://www.buybitcoinworldwide.com/exchanges/coinmama/ accessed 14 March 2020
15. https://www.wallstreetmojo.com/risk-adjusted-returns/ accessed 15 March 2020
16. https://www.investopedia.com/ask/answers/010815/what-good-sharpe-ratio.asp/ accessed 15 March 2020
17. https://www.finder.com/cryptocurrency/exchanges/ accessed 15 March 2020
18. https://machinelearningmastery.com/arima-for-time-series-forecasting-with-python/ accessed 15 March 2020
19. https://www.quantinsti.com/machine-learning-for-trading-ebook/ accessed 15 March 2020

20. https://medium.com/@kangeugine/long-short-term-memory-lstm-concept-cb3283934359/ accessed 15 March 2020
21. https://images.app.goo.gl/PBwUKTPHjKTN1Ut86/ accessed 15 March 2020
22. https://www.nasdaq.com/articles/bitcoin-just-failed-the-coronavirus-test-2020-03-13/ accessed 17 March 2020
23. https://u.today/seven-ethereum-eth-paris-conference-participants-tested-positive-for-coronavirus/ accessed 17 March 2020