

Search in the blog...



Pairs Trading Basics: Correlation, Cointegration And Strategy

[Mean Reversion & Statistical Arbitrage](#)[Excel & R For Trading](#)

Oct 23, 2019

By [Anupriya Gupta](#)

Pairs trading is supposedly one of the most popular types of trading strategy. In this strategy, usually a pair of stocks are traded in a market-neutral strategy, i.e. it doesn't matter whether the market is trending upwards or downwards, the two open positions for each stock hedge against each other. The key challenges in pairs trading are to:

- Choose a pair which will give you good statistical arbitrage opportunities over time
- Choose the entry/exit points

In this article on pairs trading, we will cover the following topics:

- [Correlation](#)
- [Cointegration](#)
- [How to choose stocks for pairs trading?](#)
- [What is z-score?](#)
- [Defining Entry points](#)
- [Defining Exit points](#)
- [A simple Pairs trading strategy in Excel](#)
- [Explanation of the model](#)

Statistics play a crucial role in the first challenge of deciding the pair to trade. The pair is commonly chosen from the same basket of stocks, for instance, Microsoft and Google (technology domain) or ICICI & Axis (Indian Banking) or Nifty Index and MSCI index (market indices). Among each domain, there are



thousands of pairs are possible. The best ones are those which are based on mathematical or statistical tests. We will learn about two statistical methods in the next section of pairs trading.

Correlation

Though not common, a few Pairs Trading strategies look at correlation to find a suitable pair to trade.

Correlation is quantified by the correlation coefficient ρ , which ranges from -1 to +1. The correlation coefficient indicates the degree of correlation between the two variables. The value of +1 means there exists a perfect positive correlation between the two variables, -1 means there is a perfect negative correlation and 0 means there is no correlation.

A perfect positive correlation is when one variable moves in either up or down direction, the other variable also moves in the same direction with the same magnitude while a perfect negative correlation is when one variable moves in the upward direction, the other variable moves in the downward (i.e. opposite) direction with the same magnitude.

The correlation coefficient for the two variables is given by

$$\text{Correlation}(X,Y) = \rho = \text{COV}(X,Y) / \text{SD}(X).\text{SD}(Y)$$

where, $\text{cov}(X, Y)$ is the covariance between X & Y while $\text{SD}(X)$ and $\text{SD}(Y)$ denotes the standard deviation of the respective variables.

If the correlation is high, say 0.8, traders may choose that pair for pairs trading. This high number represents a strong relationship between the two stocks. So if A goes up, the chances of B going up are also quite high. Based on this assumption a market neutral strategy is played where A is bought and B is sold; bought and sold decisions are made based on their individual patterns.

Just looking at correlation might give you spurious results. For instance, if your pairs trading strategy is based on the spread between the prices of the two stocks, it is possible that the prices of the two stocks keep on increasing without ever mean-reverting.

$\text{Spread} = \log(a) - n\log(b)$, where 'a' and 'b' are prices of stocks A and B respectively.

For each stock of A bought, you have sold n stocks of B.

Now, both 'a' and 'b' increases in such a way that the value of spread decreases. This will result in a loss since stock A is increasing at a rate lower than stock B and you are short on stock B.

Thus, one should be careful of using only correlation for pairs trading.

Let us now move to the next section in pairs trading basics, ie Cointegration.

Cointegration

The most common test for Pairs Trading is the cointegration test. [Cointegration](#) is a statistical property of two or more time-series variables which indicates if a linear combination of the variables is stationary.

Let us understand this statement above. The two-time series variables, in this case, are the log of prices of stocks A and B. Linear combination of these variables can be a linear equation defining the spread:

As you know, $\text{Spread} = \log(a) - n\log(b)$, where 'a' and 'b' are prices of stocks A and B respectively.

For each stock of A bought, you have sold n stocks of B.

If A and B are cointegrated then it implies that this equation above is stationary. A stationary process has very valuable features which are required to model Pairs Trading strategies. For instance, in this case, if the equation above is stationary, that suggests that the mean and variance of this equation remains constant over time. So if we start with 'n', which is called the hedge ratio, so that $\text{spread} = 0$, the property of stationary implies that the expected value of spread will remain as 0. Any deviation from this expected value is a case for statistical abnormality, hence a case for pairs trading!

With the theory in mind, let us try to answer the question which you might be thinking of, in the next section of Pairs trading basics.

How to choose stocks for pairs trading?

For any pair of stocks, define the spread as below:

$\text{Spread} = \log(a) - n\log(b)$, where 'a' and 'b' are prices of stocks A and B respectively.

Assumption: n, the hedge ratio is constant.

Calculate 'n' using regression so that spread is as close to 0 as possible. Hence, we regress the stock prices to calculate the hedge ratio.

Theory: In [regression](#), we get a term called the residuals which represents the distance of observed value from the curve fitting line or estimated value. These residuals tell us how much the actual value of 'spread' deviates from 0 for the calculated 'n'. These residuals are studied so that we understand whether or not they form a trend. If they do not form a trend, that means the spread moves around 0 randomly and is stationary.

Run the Dicky Fuller test on the spread (more complicated and popular version is called [Augmented Dicky Fuller Test](#) or ADF) values inserting the value of 'n'.

[Dickey Fuller](#) test is a hypothesis test which gives pValue as the result. If this value is less than 0.05 or

So far, we have discussed the challenges and statistics involved in selecting a pair of stocks for statistical arbitrage. We understood that by using the cointegration tests, we can say within a certain level of confidence interval that the spread between the two stocks is a stationary signal. In other words, this signal is mean-reverting. The spread is defined as:

Spread = $\log(a) - n\log(b)$, where 'a' and 'b' are prices of stocks A and B respectively. For each stock of A bought, you have sold n stocks of B. n is calculated by regressing prices of stocks A and B.

Having already established that the equation above is mean reverting, we now need to identify the extreme points or threshold levels which when crossed by this signal, we trigger trading orders for pairs trading.

To be able to identify these threshold levels, a statistical construct called z-score is widely used in Pairs Trading. In the next section, along with the z-score, we will also do a brief dive in Moving averages which is another important component in Pairs trading.

What is z-score?

Simply put, given a normal distribution of raw data points z-score is calculated so that the new distribution is a normal distribution with mean 0 and standard deviation of 1. Having such a distribution $\sim N(0, 1)$ is very useful for creating threshold levels. For example, in pairs trading, we have a distribution of spread between the prices of stocks A and B. We can convert these raw scores of spread into z-scores as explained below. This new distribution will have mean 0 and standard deviation of 1. It is easy to create threshold levels for this distribution such as 1.5 sigma, 2 sigma, 2.5 sigma, and so on.

How to calculate z-score?

$z = (x - \text{mean}) / \text{standard deviation}$, where x is a raw data point and z is the z-score.

Mean and standard deviation can be rolling statistics for a period of 't' days or minutes or time intervals.

Moving average

We divide the data into subsets of size 't', where 't' specifies a fixed time period for which the average is to be calculated. For example, to calculate the moving average of the prices of stock A where 't' is 10 days, we start by calculating average after the first 10 days in the dataset. So we calculate moving average at 10th day, 11th day, 12th day and so on. The average is moving or rolling. [Moving average](#) and the standard deviation is calculated for 't' as 10 days in the table below.

The moving average for 1-08-2001 or 11th entry would not take into account the first data point, that is, stock A prices on 18-07-2001.

Using these concepts of moving averages and z-score we create the entry points for Pairs Trading.

Defining Entry points

Let us denote the Spread as s . Thus,

$$\text{Spread} = s = \log(a) - n\log(b)$$

Calculate z-score of ' s ', using rolling mean and standard deviation for a time period of ' t ' intervals. Save this as z .

Define threshold as anything 1.5-sigma, 2-sigma. This parameter will change as per the backtesting results without risking overfitting to data.

When Z-score crosses upper threshold, go SHORT:

Sell stock A

Buy stock B

When z-score crosses the lower threshold, go LONG:

Buy stock A

Sell stock B

Maintain the hedge ratio to calculate the stock quantity

We have now understood Entry points in Pairs trading. Now we will move on to the other end, exit points.

Defining Exit points

STOP LOSS

Stop loss is defined for scenarios when the expected do not happen. For example, if we chose entry signals at 2-sigma, we are expecting that the spread will revert back to mean from this threshold. However, it is possible that spread continues to blow up. Say it reaches 2.5-sigma and you incurred losses. To prevent further losses, you place Stop Loss at say 3-sigma.

In addition to placing a pre-defined stop-loss criterion such as 3-sigma or extreme variation from the mean, you can check on the co-integration value. If the co-integration is broken during the pair is ON, the strategy warrants cutting the positions since the basic hypothesis is nullified.

TAKE PROFIT

It is defined as scenarios where you take profit before the prices move in the other direction. For instance, say you are LONG on the spread, that is, you have bought stock A and sold stock B as per the definition of spread in the article. The expectation is that spread will revert back to mean or 0. In a profitable situation, the mean would be approaching to zero or very close to it. You can keep Take Profit scenario as when the mean crosses zero for the first time after reverting from threshold levels.

There can be many ways of defining take profits depending on your risk appetite and backtesting results.

What often works is your experience and a broad range of potent skillsets that allow you to grasp a hold of the complete scenario before jumping to conclusions and help you understand practically. Like we mentioned, your appetite for risk and backtesting results will work for you. Automation and practical applications are the keys here. Anto, who had been trading for 10 years, evolved his skillsets and adapted to the growing markets with the Executive Programme in Algorithmic Trading ([EPAT](#)) and is happily trading in this domain.

Let us try to recap what we have understood so far. Pairs Trading is a trading strategy that matches a long position in one stock/asset with an offsetting position in another stock/asset that is statistically related. Pairs Trading can be called a [mean reversion strategy](#) where we bet that the prices will revert to their historical trends.

So far, we have gone through the concepts and now let us try to create a simple Pairs Trading strategy in Excel.

A simple Pairs trading strategy in Excel

This excel model will help you to:

- Learn the application of mean reversion
- Understand of Pairs Trading
- Optimize trading parameters
- Understand significant returns of statistical arbitrage

Why should you download the trading model?

As the trading logic is coded in the cells of the sheet, you can improve the understanding by downloading and analyzing the files at your own convenience. Not just that, you can play around the numbers to obtain better results. You might find suitable parameters that provide higher profits than specified in the article.

Explanation of the model

In this example, we consider the MSCI and Nifty pair as both of them are stock market indexes. We implement mean reversion strategy on this pair. Mean reversion is a property of stationary time series. Since we claim that the pair we have chosen is mean reverting we should test whether it follows stationarity.

Plotting of the logarithmic ratio of Nifty to MSCI makes it appear to be mean reverting with a mean value of 2.088 but we use Dicky Fuller Test to test whether it is stationary with a statistical significance. The results under Cointegration output table shows that the price series is stationary and hence mean-reverting. Dicky Fuller Test statistic and a significantly low p-value (<0.05) confirms our assumption. Having determined that the mean reversion holds true for the chosen pair we proceed with specifying assumptions and input parameters.

Assumptions

- For simplification purpose, we ignore bid-ask spreads.
- Prices are available at 5 minutes intervals and we trade at the 5-minute closing price only.
- Since this is discrete data, squaring off of the position happens at the end of the candle i.e. at the price available at the end of 5 minutes.
- Only the regular session (T) is traded
- Transaction costs are \$0.375 for Nifty and \$1.10 for MSCI.
- The margin for each trade is \$990 (approximated to \$1000).

Input parameters

Please note that all the values for the input parameters mentioned below are configurable.

- Average of 10 candles (one candle is equal to every 5-minute price) is considered
- A "z" score of +2 is considered for buy and -2 for selling
- A stop loss of \$100 and profit limit of \$200 is set
- The order size for trading MSCI is 50 (1 lot) and for Nifty is 6 (3 lots)

The market data and trading parameters are included in the spreadsheet from the 12th row onwards. So when the reference is made to column D, it should be obvious that the reference commences from D12 onwards.

Explanation of the columns in the Excel Model

	A	B	C	D
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

Input Parameters	
Average	10
Buy Threshold	-2
Sell Threshold	2
Stop Loss	-100 USD
Take Profit	200 USD
MSCI Order Size	50 i.e. 1 lot
Nifty Order Size	6 i.e. 3 lots

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	S No.	Date & Time	MSCI	NIFTY	I(0)	Avg	Std Dev	z-Score	Signal	Buy price	Sell Price	MTM	Status	Profit/Loss	Cumulative
13	1	17-02-2015 10:00	1092.4	8822.5	2.088928									0	0
14	2	17-02-2015 10:05	1092.4	8822.5	2.088928									0	0
15	3	17-02-2015 10:10	1092.4	8822.5	2.088928									0	0
16	4	17-02-2015 10:15	1092.4	8822.5	2.088928									0	0
17	5	17-02-2015 10:20	1092.4	8822.5	2.088928									0	0
18	6	17-02-2015 10:25	1092.4	8822.5	2.088928									0	0
19	7	17-02-2015 10:30	1092.4	8830	2.089778									0	0
20	8	17-02-2015 10:35	1092.4	8830	2.089778									0	0
21	9	17-02-2015 10:40	1092.4	8830	2.089778									0	0
22	10	17-02-2015 10:45	1092.4	8830	2.089778									0	0
23	11	17-02-2015 10:50	1092.4	8829	2.089665	2.0892681	0.000438803	0.903791						0	0
24	12	17-02-2015 10:55	1092.4	8821	2.088758	2.0893417	0.000437219	-1.33475						0	0
25	13	17-02-2015 11:00	1092.4	8821	2.088758	2.0893247	0.000457905	-1.23732						0	0
26	14	17-02-2015 11:05	1092.4	8825.5	2.089268	2.0893077	0.000477024	-0.08292						0	0
27	15	17-02-2015 11:10	1092.4	8825	2.089211	2.0893417	0.000458733	-0.28384						0	0
28	16	17-02-2015 11:15	1092.4	8824.5	2.089155	2.08937	0.000438665	-0.49058						0	0
29	17	17-02-2015 11:20	1092.4	8824.5	2.089155	2.0893927	0.000418697	-0.56811						0	0
30	18	17-02-2015 11:25	1092.4	8823.5	2.089042	2.0893304	0.000400991	-0.72043						0	0
31	19	17-02-2015 11:30	1092.4	8808.5	2.08734	2.0892568	0.00037655	-5.09017	Buy	8808.5	1092.4		Buy	0	0
32	20	17-02-2015 11:35	1092.4	8817.5	2.088361	2.089013	0.000673625	-0.96745		8808.5	1092.4	54	Buy	0	0
33	21	17-02-2015 11:40	1092.4	8822.5	2.088928	2.0888713	0.000643154	0.08841		8808.5	1092.4	84	Buy	0	0
34	22	17-02-2015 11:45	1092.4	8818.5	2.088475	2.0887977	0.000581419	-0.5555		8808.5	1092.4	60	Buy	0	0

Column C represents the price for MSCI.

Column D represents Nifty price.

Column E is the logarithmic ratio of Nifty to MSCI.

Column F calculates 10 candle average. Since 10 values are needed for average calculations, there are no values from F12 to F22.

The formula =IF(A23>\$C\$3, AVERAGE(INDEX(\$E\$13:\$E\$1358, A23-\$C\$3):E22), "") means that the average should be calculated only if the data sample available is more than 10 (i.e. the value specified in cell C3), otherwise the cell should be blank.

Consider cell F22. Its corresponding cell A22 has a value of 10. Since A22>\$C\$3 fails, the entry in that cell is blank. The next cell F23 has a value since A23>\$C\$3 is true. Let's move to the next column.

In column G, the formula, `AVERAGE(INDEX(E13:E1358, A23-C3):E22)` calculates the average value of last 10 (as mentioned in cell C3) candles of column E data. Similar logic holds for column G where the standard deviation is calculated.

The “z” score is calculated in the column H. Formula for calculating “z” score is $z = (x - \mu) / (\sigma)$. Here x is the sample (Column E), μ is the mean value (Column F) and σ is the standard deviation (Column G).

Column I represents the trading signal. As mentioned in the input parameters, if “z” score goes below -2 we buy and if it goes above +2 we sell. When we say buy, we have a long position in 3 lots of Nifty and have a short position in 1 lot of MSCI. Similarly, when we say sell, we have a long position in 1 lot of MSCI and have a short position in 3 lots of Nifty thus squaring off the position. We have one open position all the time.

To understand what this means, consider two trading signals “buy” and “sell”. For the “buy” signal, as explained before, we buy 3 lots of Nifty future and short 1 lot of MSCI future. Once the position is taken, we track the position using the Status column, i.e. column M. In each new row while the position is continuing, we check whether the stop loss (as mentioned in cell C6) or take profit (as mentioned in cell C7) is hit. The stop loss is given the value of USD -100, i.e. loss of USD 100 and take profit is given the value of USD 200 in the cells C6 and C7 respectively.

While the position does not hit either stop loss or take profit, we continue with that trade and ignore all signals that are appearing in column I. Once the trade hits either the stop loss or take profit, we again start looking at the signals in column I and open a new [trading position](#) as soon as we have a Buy or Sell signal in column I.

Column M represents the trading signals based on the input parameters specified. Column I already has trading signals and M tells us about the status of our trading position i.e. are we long or short or booked the profits or exited at the stop loss. If the trade is not exited, we carry forward the position to the next candle by repeating the value of the status column in the previous candle. If the price movement occurs in such a way that it breaches the given TP or SL then we square off our position thus denoting it by “TP” and “SL” respectively.

Column L represents Mark to Market. It specifies the portfolio position at the end of time period. As specified in the input parameters we trade 1 lot of MSCI and 3 lots of Nifty. So when we trade our position is the appropriate price difference (depending on whether we are bought or sold) multiplied by the number of lots.

Column N represents the profit/loss status of the trade. P/L is calculated only when we have squared off our position. Column O calculates the cumulative profit.

Outputs

The output table has some performance metrics tabulated. Loss from all loss-making trades is \$3699 and profit from trades that hit TP is \$9280. So the total P/L is $\$9280 - \$3699 = \$5581$. Loss trades are the trades that resulted in losing money on the trading positions. Profitable trades are the successful trades ending in gaining cause. Average profit is the ratio of total profit to the total number of trades. Net average profit is calculated after subtracting the transaction costs which amounts to \$91.77.

Now it is your turn!

- First, download the model
- Modify the parameters and study the backtesting results
- Run the model for other historical prices
- Modify the formula and strategy to add new parameters and indicators! Play with logic! Explore and study!

Comment below with your results and suggestions

Summary

Thus, we have understood the concept behind Pairs trading strategy, including correlation and cointegration. We also took a look at Z-score and defined the entry and exit points when we are executing a pairs trading strategy. We also created an Excel model for our Pairs Trading strategy!

Learn how to implement pairs trading/statistical arbitrage strategy in [FX markets](#) through a project work including live examples. If you want to dig deeper and try to find suitable pairs to apply the strategy, you can go through the blog on [K-Means algorithm](#).

If you want to learn various aspects of Algorithmic trading then check out our [Executive Programme in Algorithmic Trading](#) (EPAT). The course covers training modules like Statistics & Econometrics, Financial Computing & Technology, and Algorithmic & Quantitative Trading. EPAT is designed to equip you with the right skill sets to be a successful trader. Enroll now!

 **Login to Download**

Disclaimer: All data and information provided in this article are for informational purposes only. QuantInsti® makes no representations as to accuracy, completeness, currentness, suitability, or validity of any information in this article and will not be liable for any errors, omissions, or delays in this information or any losses, injuries, or damages arising from its display or use. All information is provided on an as-is basis.

 [Algo Trading Course](#)

Share Article:     



Nov 28, 2017

Arbitrage Strategies: Understanding Working of Statistical Arbitrage

Feb 15, 2021

Mean Reversion Strategies: Introduction, Trading, Strategies and More

40 Comments

[Sign Up To Comment](#)



Laura Simoes • 2 Years Ago

Hi! I have a doubt: how did you find the order size? The reversion is on daily basis or in 5-min basis?



Laura Simoes • 2 Years Ago

The "n" parameter.



Jan Tarnogrocki • 2 Years Ago

Hi There, In the excel file you are trading 50 contracts MSCI and 6 Nifty. The average net profit per trade is 91.77 USD. But the spread is not taken into consideration. If it is above 1.64 USD per contract the strategy is at break even. In addition slippage on real markets has to be outperformed too.

I do really like this piece of work you are presenting here especially how complex statistic is explained in an easy and understandable way.

But I am afraid the results in the excel file shown are far away from reality.



(I would be happy if I were wrong. But I coded the strategy in Java using tick data for back testing including spread and there is no penny left at the end of the day)



QuantInsti • 2 Years Ago

Hi! Thanks for your query. Just to understand your query better, could you please elaborate on "But the spread is not taken into consideration"?



Jan Tarnogrocki • 2 Years Ago

Hello,

I am referring to the spread between Bid and Ask you have to pay every time a trade is opened and closed (not the spread between the cointegrated pairs used for generating trade signals).

The assumption is also posted within the excel file: "Bid Ask Spread is ignored."



Dirk Damman • 2 Years Ago

Hello, If you change the exit parameters based on $I(0)$ 1or -1. You get the following results. This looks more profitable ?

Output

Loss from SL -893

Profit from TP 7865

Total Profit 6972

Loss Trades 8

Profitable Trades 72

Average Profit (USD) 87.15

Net Average Profit (USD) 82.70

Return per month 17%



QuantInsti • 2 Years Ago

Hi! Thank you for the comment. As we said in the article, this sheet was only for illustration. You can play around the numbers to obtain better results. And you can definitely find parameters that profit higher profits than specified in the article. Hope that answers your query?



QuantInsti • 2 Years Ago

Hi! Thank you for the comment. As we said in the article, this sheet was only for illustration. You can play around the numbers to obtain better results. And you can definitely find parameters that profit higher profits than specified in the article.Hope that answers your query?



Miguel de Lazio • 2 Years Ago

Why is the last value of "I(0)" not used in the 10-period moving average? MA is calculated from previous periods only and row 22 has no average.



QuantInsti • Last Year

Thanks for your comment. The only reason to exclude "I(0)" is to avoid the look-ahead bias. If our trading hypothesis says that we can take a position at the close price, we can include row 22 in our calculations and start from there. Hope this helps?



Miguel de Lazio • Last Year

I don't understand why there would be a look-ahead, at least not when I assume that the time column shows the opening time of the bar. The closing prices are available at the end of the bars when the signal is calculated. In the Excel sheet at row 22, closing prices of 10 bars are available. (closing time of this bar is one tick before 10:50, in my program I set this to 10:49:59 and 999 milliseconds which becomes the timestamp of the signal) Then you make a trading decision, e.g. place buy order. The profit in the next bar can be calculated using the closing price and 'closing signal' of the previous bar and closing price of the next bar.



Miguel de Lazio • Last Year

Hi again, I think I can see what you mean. The VLOOKUP is doing a lookup using the Date & Time of the next row. So my assumption above is incorrect.



QuantInsti • Last Year

Thank you for your comments Miguel. We hope this blog is helpful to you. :)



Omar Batista Iamas • Last Year

Hello, I am surprised when you change the order of the data in the "Raw Data" sheet, the results change because the Z-Core is inverted.



QuantInsti • Last Year

Yes, you are right. When you change the order of the raw data, keeping all the parameters same, the value of spread changes and thus z-score.Note: In most cases, there would be a minor difference in the spread so it won't make much difference.



Miguel de Lazio • Last Year

I do not understand how to calculate the relative order size using the 'n' ratio that is found using " $\log(a) - n\log(b)$ ". If 'n' would be calculated from " $a - nb$ " then the ratio will be directly related to how much the two prices will have to change (relatively speaking) before the spread reverts to the mean. But " $\log(a) - n\log(b)$ " gives a completely different 'n' value. So how is it used to manage the risk and order size?



QuantInsti • Last Year

Thank you for the comment, Miguel and apologies for the delay in the response. When you are regressing raw prices to calculate spread, then you need to manage quantity. With log prices, you need to maintain value. For example, $\text{Spread} = a - n * b$ Here, for 1 quantity of 'a', you need to sell n quantity of 'b'. $\text{Spread} = \log(a) - m\log(b)$ Here, for the investment of \$100 in 'a', you need to invest \$100m in 'b'. We hope this helps with your query?



Filippo Pacinelli • Last Year

Hi! Nice analysis! Can you explain the theoretic assumption behind using rolling mean in the computation of the z-score? Since it modifies the structure of the vector, making it more stationary, why is it a reliable measure to build the strategy on? Maybe my reasoning is wrong, I hope I have explained myself. Thanks, Bye!



QuantInsti • Last Year

Thank you for your comment. We'd be glad to help you out.

Z-score describes a value's relationship to the mean of a group of values. It is measured in terms of standard deviation from the mean.

$$\text{Z-score} = (x - \text{Mean}) / \text{Standard deviation}$$

Mean reversion strategy involves the purchase or sale of financial instruments whose recent performance differs from their historical average. So, we use z-score values to find how much the prices vary from their historical average.

We hope this helps with your query?



Khalil Elhanafi • Last Year

Can you explain what do you mean by "The margin for each trade is \$990 (approximated to \$1000)." and what does that add to this model. Thanks.



QuantInsti • Last Year

Hi! Thank you for the comment. This is one of the assumptions of the strategy. What it tries to say is that if you are taking a position in Nifty futures, you need to keep around \$990 or around INR 75,000 with the

broker to initiate the trade. This margin varies from broker to broker. We hope this helps with your query.



Khalil Elhanafi • Last Year

Thank You for your reply. We need 990\$ * 2 for the strategy because in each sell or buy we are taking two position. is that true ?



Khalil Elhanafi • Last Year

Thank you for your reply, can we use the same strategy with cryptocurrencies ?



QuantInsti • Last Year

Hi! Yes, you can use the same strategy with cryptocurrencies.



QuantInsti • Last Year

Yes, you need to place two orders simultaneously. So, the margin required is twice.



David Rooker • Last Year

I like this strategy, but with little to no understanding on programming I have no idea how to implement it on a trading account :/



QuantInsti • Last Year

We'd be glad to help you, David. Our self-paced specialization courses from Quantra can help you learn how to do it. Feel free to connect with our team here: <https://bit.ly/31sTxoG>



masoud shirin nasab • Last Year

HI, CAN MAKE A EX : FOR FOREX PAIRS : EURUSD AND USDCHF?



QuantInsti • Last Year

Hi! Thanks for your comment. As suggested in the article, there are tests of correlation, and cointegration can be performed to check the feasibility of pair trading between the two instruments. The correlation can be verified using 1. online tools like <https://www1.oanda.com/lang/en/forex-trading/analysis/currency-correlation>. 2. through excel "=CORREL(A2:A7,B2:B7)" 3. through python (df['EURUSDClose'].pctchange()).corr(df['USDCHFCLose'].pctchange()) And Cointegration can be performed by: 1. through excel using regression 2. through python using below snippet: import statsmodels.tsa.stattools as ts cadf = ts.adfuller(df1.spread) We hope this helps.



Vedant Chinmaya Khandelwal • Last Year

The excel sheet isnt getting downloaded? Can someone mail it to me instead?



QuantInsti • 9 Months Ago

Hi Vedant, apologies for the late reply. We have fixed the download button. Do check.



?? ? • 10 Months Ago
hi what does mtm and status mean?



Vibhu • 9 Months Ago
MTM, also known as mark-to-market, is the concept of valuing assets and liabilities at the last known price is known as mark-to-market. Consider you are long 1 contract in August corn futures trading at \$500. At the end of the day, the price falls to \$490. You were long, and the price fell, so you have an unrealised loss of \$10. This \$10 will be deducted from your account at the end of the day. And the procedure is called daily mark-to-market settlements.

Status is simply the daily position.



M G • 9 Months Ago
hello , i am unable to download the xls . can you please provide it ?



QuantInsti • 9 Months Ago
Thanks for bringing this to our notice, and for sharing the correct link. We have fixed the same. :)



Simone Bosio • 7 Months Ago
Hi There, great article and superb explanation. The only thing I'm missing is the "Return per month" calculation (cell: K10). Why are you multiplying the average net profit times 2 and dividing the result by 1000? Thank you in advance



QuantInsti • 7 Months Ago
Hi Simone, thank you for your comment. Since we are taking a trade on two instruments, each utilising a margin of \$1000, we divide the "Net Average Profit" with \$2000, to get the percentage of profit per trade.

However, there is a discrepancy in this. You need to apply the formula as $=K9/(1000*2)$

☐ We will rectify this in the downloads as soon as possible. Hope this helps.

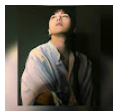


Simone Bosio • 7 Months Ago
Thanks for the prompt reply, yes this makes sense. Have a nice day, Simone



QuantInsti • 7 Months Ago
Thank you, Simone. We're happy to help. :)





Jiawei Feng • 6 Months Ago

Hi! I think this is a great blog so far and I did learn a lot from it. However, i just have one small question. Why is the order size of Nifty set to 6 and the order size of MSCI set to 50. I am curious how do we decide the order size. By the way, what does it mean by "1 lot" and "3 lots"?

Search in the blog...



Career Growth

[Industry](#)

[Jobs & Skills](#)

[Trading Desk Setup](#)

[Success Stories](#)

[EPAT Trading Projects](#)

Tutorials

[Automated Trading](#)

[Excel & R For Trading](#)

[Forex & Crypto Trading](#)

[Machine Learning](#)

[Maths and Econometrics](#)

[Mean Reversion & Statistical Arbitrage](#)

[Momentum Trading](#)

[More Trading Strategies](#)

[Options Trading](#)

[Portfolio & Risk Management](#)

[Python For Trading](#)

[Sentiment Trading](#)

[Technical Indicators](#)

LEARN ALGO TRADING

- HANDS-ON TRAINING
- ONLINE & LIVE LECTURES
- USE PYTHON FOR TRADING
- JOB ASSISTANCE
- FINANCIAL AID AVAILABLE

15+ World-class Faculty

KNOW MORE



About

QuantInsti
Quantra
Blueshift

Events & Announcements

Announcements
Webinars

Quick Links

Contact Us
Privacy Policy
Blog Contribution

Copyright © 2021 QuantInsti.com All Rights Reserved.

