# Algorithmic Statistical Arbitrage Case

Rotman International Trading Competition (RITC)x

## OVERVIEW

The Algorithmic Statistical Arbitrage Case challenges participants to develop data-driven trading algorithms using the RIT API to identify and exploit short-term deviations in statistical relationships among correlated stocks and a benchmark index. Using historical data to estimate expected volatilities, correlations, and betas, participants must design systems that can detect when realized values diverge from expectations and execute trades to capture these mispricings.

Throughout the case, markets will be fast-paced and highly liquid, requiring algorithms that can adapt quickly to changing conditions while maintaining disciplined risk control. Success will depend on building robust, adaptive strategies that translate quantitative insights into real-time, automated decision-making to generate consistent arbitrage profits.

## DESCRIPTION

The Algorithmic Statistical Arbitrage Case will be 5 minutes long and represent 3 months of trading.

All trades must be executed by a trading algorithm, so participants will not be allowed to trade manually through the RIT Client once the case begins. However, participants are allowed to modify their algorithms. A base template algorithm will be provided for participants and can be modified for use in the competition. However, participants are encouraged to create their own algorithms using the RIT API.

## MARKET DYNAMICS

In this case, participants will be trading three stocks from the same industry that are also included in the same index. This index is composed of 1,000 publicly traded companies. The three stocks will also be correlated with the index. Hence, there will be four securities of which the details are shown below.

| Ticker | NGN | WHEL | GEAR | RSM1000 |
|---|---|---|---|---|
| Fee/share (Market orders) | \$0.01 | \$0.01 | \$0.01 | N/A |
| Max order size | 10,000 | 10,000 | 10,000 | N/A |
| Type | Stock | Stock | Stock | Non-tradable index |

At the start of each iteration, historical data (representing 3 months of trading) will be made available to the participants on the RIT User Application. At the beginning of each iteration, participants will need to calculate the historical volatilities, correlations, and other relevant statistical measures according to the formulas provided in the section "Statistical Formulas" below. These statistical measures generated from the historical values can be interpreted as the expected value of these measures.

The realized correlations in each iteration will be very close to the historical correlations. For example, if the historical correlation between stock NGN and the RSM1000 Index (calculated at the beginning of the case using the historical data) is 0.5, participants should expect that at the end of the case, the realized correlation between stock NGN and RSM1000 (calculated over the entire iteration, that is, from tick 0 to tick 300) should be very close to 0.5[1].

Participants are encouraged to build an algorithm that can identify short-term deviations of the realized statistical measures from their expected values and exploit statistical arbitrage opportunities throughout the case. During the 300 ticks of trading, the market will be very liquid and participants can expect to be price takers in the market.

## STATISTICAL FORMULAS

1. The rate of return for security $x$ at time $t$ is given by

$$r_{x,t} = \frac{Price_{x,t}}{Price_{x,t-1}} - 1 \,,$$

2. where $Price_{x,t}$ and $Price_{x,t-1}$ are the price of the security at time $t$ and $t-1$. The mean return on security $x$ over $N$ observations, with one observation per tick. We'll refer to this mean return as "mean tick return", which can be expressed as

$$\mu_x = \frac{1}{N}\sum_{t=1}^{N} r_{x,t}$$

3. The tick variance on security $x$ over $N$ observations is given by

$$\sigma_x^2 = \frac{1}{N}\sum_{t=1}^{N}\left(r_{x,t} - \mu_x\right)^2$$

4. The tick standard deviation (tick volatility) of security $x$ over $N$ observations:

$$\sigma_x = \sqrt{\sigma_x^2}$$

5. The covariance between securities $x$ and $y$ over $N$ observations is

$$\sigma_{xy} = \frac{1}{N-1}\sum_{t=1}^{N}\left(r_{x,t} - \mu_x\right)\left(r_{y,t} - \mu_y\right)$$

6. The correlation between securities $x$ and $y$ over $N$ observations can be expressed as

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

[1] Deviations are possible due to the unpredictable behaviour of the participants' algorithms in the market.

7. Volatility conversion (300 ticks represent three months of trading):

$$\sigma_{annual} = \sigma_{quarterly} \times \sqrt{4} = \sigma_{tick} \times \sqrt{4 \times 300}$$

where $\sigma_{annual}$ is the annual volatility, $\sigma_{quarterly}$ is the quarterly volatility and $\sigma_{tick}$ is the tick volatility.

8. Beta ($\beta_x$) calculation between a single stock (security $x$) and the index ($RSM1000$) is

$$\beta_x = \rho_{x,RSM1000} \times \frac{\sigma_x}{\sigma_{RSM1000}}$$

where $\rho_{x,RSM1000}$ is the correlation between stock $x$ and $RSM1000$ returns.

9. Beta ($\beta_{port}$) for a portfolio of stocks ($port$)

$$\beta_{port} = \sum_{x=1}^{N} w_x \beta_x$$

where $w_x$ is the weight of stock $x$ in the portfolio.

## STATISTICAL ARBITRAGE EXAMPLE

At the beginning of the case, using the historical data provided and the formulas provided in the section "Statistical Formulas", the following statistics were calculated:

**Observed Historical Volatility**

| RSM1000 | NGN | WHEL |
|---|---|---|
| 0.0220 | 0.0197 | 0.0193 |

**Observed Historical Correlation**

| | RSM1000 | NGN | WHEL |
|---|---|---|---|
| **RSM1000** | 1 | 0.5360 | 0.6462 |
| **NGN** | – | 1 | 0.3278 |
| **WHEL** | – | – | 1 |

**Historical Beta**

| NGN | WHEL |
|---|---|
| 0.4810 | 0.5659 |

We can calculate:

$$PTD\ Return_{x,t} = \frac{Price_{t=currrent\ tick}}{Price_{t=1}} - 1$$

where $PTD\ Return$ is the Period to Date return at time $t$ from the beginning of the iteration for stock $x$.

$$Expected\ PTD\ Return_{x,t} = PTD\ Return_{RSM1000,t} * \beta_x$$

where $Expected\ PTD\ Return_{x,t}$ is the Expected Period to Date return at time $t$ from the beginning for stock $x$, $PTD\ Return_{RSM1000,t}$ is the Period to Date return at time $t$ from the beginning for $RSM1000$ and $\beta_x$ is the $Beta$ between stock $x$ and the index ($RSM1000$).

$$Divergence_{x,t} = PTD\ Return_{x,t} - \ Expected\ PTD\ Return_{x,t}$$

where $Divergence_{x,t}$ is the divergence at time $t$ for stock $x$. We define the $divergence$ at time $t$ for stock $x$ as the difference between the realized $PTD\ Return$ and the $Expected\ PTD\ Return$ for stock $x$ at time $t$.

In the table below, NGN starts to diverge after trading begins, and by tick 46, the divergence is at +8.00%. At the same time, WHEL shows a slightly positive divergence. In order to take advantage of this, one can short-sell NGN and buy WHEL in order to create a market neutral position in terms of beta (i.e. you would like the beta of the portfolio to be as close as possible to zero). Since the two stocks have different betas, $\beta_{NGN} = 0.4810$ and $\beta_{WHEL} = 0.5659$, one may decide to short $10,000 worth of NGN shares (approximately 338 shares) and buy $8,500 worth of WHEL shares (285 shares). The resulting portfolio's beta ($\beta_{Port}$) will be very close to zero since the ratio is approximately the following:

$$|(\$value\ of\ NGN\ position)/(\$\ value\ of\ WHEL\ position)|\ =\ |\ \beta_{WHEL}\ /\ \beta_{NGN}\ |$$

| | $Price$ | | | $PTD\ Return_{x,t}$ | | | $Expected\ PTD\ Return_{x,t}$ | | $Divergence_{x,t}$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| Tick | RSM 1000 | NGN | WHEL | RSM 1000 | NGN | WHEL | NGN | WHEL | NGN | WHEL |
| 1 | 35.96 | 27.04 | 29.31 | – | – | – | – | – | – | – |
| … | … | … | … | … | … | … | … | … | … | … |
| 46 | 36.78 | 29.50 | 29.77 | 0.0228 | 0.0910 | 0.0157 | 0.0110 | 0.0129 | 0.080 | 0.0028 |
| … | … | … | … | … | … | … | … | … | … | … |
| 59 | 36.18 | 26.84 | 29.36 | 0.0061 | −0.0074 | 0.0017 | 0.0029 | 0.0035 | −0.0103 | −0.0018 |

When the divergence converges towards zero at tick 59, one can close the position, and the realized profits from the statistical arbitrage will be equal to the arbitrage spread minus commissions:

$$(\$29.50 - \$26.84)\ \times\ 338\ +\ (\$29.36\ -\ \$29.77) \times\ 285\ -\ (338\ +\ 285) \times\ \$0.01 \times\ 2 \\ =\ \$769.77$$

## TRADING LIMITS AND TRANSACTION COSTS

| Case time (tick) | Gross/Net |
|---|---|
| $1 \sim 299$ | $500{,}000/100{,}000$ |
| $300$ | $0/0$ |

Each participant will be subject to gross and net trading limits during trading in each iteration. The gross trading limit reflects the sum of the absolute values of the long and short positions across all securities and the net trading limit reflects the sum of long and short positions such that short positions negate any long positions. Trading limits will be strictly enforced and participants will not be able to exceed them. Each position in stock will be counted towards trading limits with a multiplier of 1 (i.e. if you long 100 shares of any stock, your gross and the net trading limits will increase by 100).

The maximum trade size will be 10,000 shares per order for stocks. Transaction fees will be set at $0.01 per share.

## POSITION CLOSE-OUT

Any non-zero position will be closed out at the end of trading at either the last traded price or the correct price if the stock is experiencing a divergence due to market shock. As such, participants are not required to close statistical arbitrage positions as the iteration ends.

## KEY OBJECTIVES

**Objective 1**

Create a model using the provided template to calculate relevant statistical measures between the 3 stocks and non-tradable index.

**Objective 2**

Build a trading algorithm and optimize the trading parameters such that the algorithm efficiently takes advantage of statistical arbitrage opportunities while managing risk exposure and order flow.