# **Statistical-Arbitrage-ML**

For a family of stocks, generally belonging to the same sector or industry, there exists a correlation between prices of each of the stocks. There, though, exist anomalous times when for a small period of time, the correlation is broken. But the market self corrects in some time and the correlation is re-established.

Develop Machine Learning Algorithm to predict statistical arbitrage opportunities in NSE based on the 2016 data. Test this algorithm on 2017 data.

Statistical arbitrage comprises a set of quantitatively driven algorithmic trading strategies. These strategies look to exploit the relative price movements across thousands of financial instruments by analyzing the price patterns and the price differences between financial instruments. It can be categorized as a medium-frequency strategy where the trading period occurs over the course of a few hours to a few days.

**Steps we need to follow to build a model:**

* Pre-Processing
* Time series Analysis
* Classification
* Prediction

**Initially:**

1)Read Data

2)Iterates over the DataFrame columns, returning a tuple with the column name and the content as a Series. The index of the row.

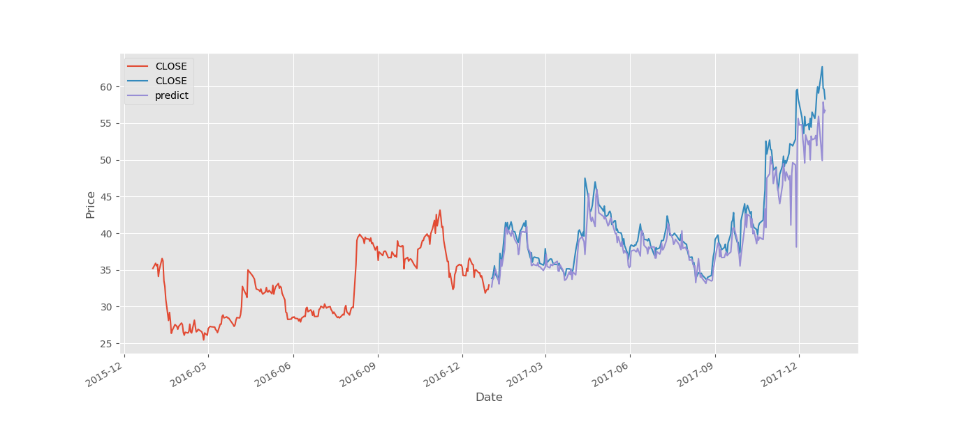
3)Converting 20MICRONS dataframe to CSV

**Selecting Features and labels:**

For the NSE stock exchange data,

Features are ['OPEN', 'HIGH', 'LOW’, 'TOTTRDQTY', 'Date', 'PREVCLOSE', 'TOTTRDVAL', 'TOTALTRADES’]

And the target variable is [‘CLOSE’] and the plot is



We introduce a new parameter which calculate High low %change i.e., [‘HC\_PCT’] which acts as common feature to [‘HIGH’] and [‘LOW’].

**Data Pre-Processing:**

***TIME SERIES ANALYSIS*** :- Separate Train and test data

* Splitting on basis of time series analysis, 2016 reports as training data and 2017 reports as testing data.

**CLASSIFIER:**

TheilSenRegressor() is used.

* Accuracy score of our model is 98.2% which is decent for a yearly predictions.

**PREDICTION:**

predict = clf.predict(x)

test['predict'] = predict

Plot the prediction results

