INTRODUCTION TO QUANTITATIVE TRADING

Semester Project

July 29,2023

Overview

To dive into Trading strategies by this project..

Understand and Implement from scratch…

Milestones

Week - 1:

Basics of Python
Python libraries(Numpy,Pandas,Matplotlib)
Extracting historical stock data from yahoo finance
Plot candlestick and line graphs
Analyze bullish/bearish periods
Plotting various ratios:

I. Open/Close
II. Open/Low
III. High/Low
IV. High/Close

Week - 2:

Correlation and Cointegration
Moving Averages (SMA and EMA)

Golden Cross:When SMA(15 days) overpowers SMA(50 days)
Death Cross:When SMA(50 days) overpowers SMA(15 days)
Sharpe Ratio
Pairwise-Trading
Mean Reversion Strategy

Week - 3 & 4:

Useful Tests, Ratios and statistics to analyze stock prices:
Hypothesis Testing
ADF test
Stationary and Non stationary stocks
Z score
Long Position vs Short Position
Signals
Returns
Bottom line
Indicators
Generating Sell/Buy signals using Z-Score graph on Stationary pair of stocks

Correlation : Correlation refers to a statistical measure that describes the relationship between two or more variables. Correlation does not imply causation, but it helps in understanding the strength and direction of the relationship between variables. Correlation refers to the statistical relationship between the price movements of two or more stocks or other financial instruments. It helps investors and analysts understand how the prices of different stocks move relative to each other.

Cointegration: It refers to a statistical property that indicates a long-term equilibrium relationship between the prices or returns of two or more financial assets. It suggests that while the individual prices or returns may wander independently in the short term, they tend to move together over the long term.

Difference between correlation and cointegration:
Correlation measures the short-term relationship between variables, whereas cointegration focuses on the long-term relationship. Correlation can be a temporary or spurious relationship, whereas cointegration implies a more meaningful and persistent connection.

Sharpe ratio - it is a measure of a stock's returns compared to the return of a risk-free asset. It also compares the returns to its risk. Sharpe ratio is given by the formula - $\frac{returns - risk\ free\ return\ rate}{standard\ deviation\ of\ returns}$

Pairwise trading strategy - this strategy is used on a pair of highly correlated stocks. It works by matching the long/short position on one stock with an opposite position on the other stock. This is a mean reversion strategy, because we predict that the ratio of the correlated two stocks will always revert back to its mean.

Mean reversion strategy - in this strategy, we predict that the value of a stock will eventually revert to its long time mean. In this strategy, we go long when the value of the asset reaches a low below its long time mean, and short when it reaches a high.

Hypothesis Testing :- An analyst performs hypothesis testing on a statistical sample to present evidence of the plausibility of the null hypothesis. Measurements and analyses are conducted on a random sample of the data to test a theory. Analysts use a random data sample to test two hypotheses: the null and alternative hypotheses. The Null Hypothesis is the assumption that the event will not occur. A null hypothesis has no bearing on the study's outcome unless it is rejected. H0 is the symbol for it.
The Alternate Hypothesis is the logical opposite of the null hypothesis. The acceptance of the alternative hypothesis follows the rejection of the null hypothesis. H1 is the symbol for it.

Stationary and non stationary stocks :A stationary stock refers to one whose price or return series displays statistical properties that remain constant over time.

The test statistic is compared to the critical values at different significance levels, usually 1%, 5%, and 10%. If the test statistic is more negative than the critical value, you can reject the null hypothesis and conclude that the time series is stationary. If the test statistic is less negative than the critical value, you cannot reject the null hypothesis and conclude that the time series has a unit root. The p-value is the probability of obtaining a test statistic as extreme or more extreme than the observed one under the null hypothesis. If the p-value is less than the significance level, you can reject the null hypothesis and conclude that the time series is stationary. If the p-value is greater than the significance level, you cannot reject the null hypothesis and conclude that the time series has a unit root.

A non-stationary stock refers to one whose price or return series exhibits trends, changes in mean or variance, or other irregular patterns over time. Non-stationary stock series lack a consistent behavior and are characterized by unpredictable movements. Non-stationarity can arise due to various factors such as market trends, economic conditions, news events, or company-specific factors.

ADF Test :- It is one of the most common statistical tool to analyse the stationarity of a series. This is a hypothesis testing involving null and alternate hypothesis and a test statistic is computed and p-values reported, using which you infer whether a series is stationary.The ADF test results will give you a test statistic and a p-value. The test statistic is compared to the critical values at different significance levels, usually 1%, 5%, and 10%. If the test statistic is more negative than the critical value, you can reject the null hypothesis and conclude that the time series is stationary. If the test statistic is less negative than the critical value, you cannot reject the null hypothesis and conclude that the time series has a unit root. The p-value is the probability of obtaining a test statistic as extreme or more extreme than the observed one under the null hypothesis. If the p-value is less than the significance level, you can reject the null hypothesis and conclude that the time series is stationary. If the p-value is greater than the significance level, you cannot reject the null hypothesis and conclude that the time series has a unit root.

Z-Score :- It describes a value's relationship to the mean of a group of values. Z-score is measured in terms of standard deviations from the mean. If a Z-score is 0, it indicates that the data point's score is identical to the mean score. Z-scores may be positive or negative, with a positive value indicating the score is above the mean and a negative score indicating it is below the mean.

Z-scores are measures of an instrument's variability and can be used by traders to help determine volatility. The statistical formula for a value's z-score is calculated using the following formula:

$$z = (x - \mu) / \sigma$$

Where:

- z = Z-score
- x = the value being evaluated
- μ = the mean
- σ = the standard deviation

Moving Average - It is a statistical calculation used to analyze the trend and smooth out price fluctuations over a specific period. It provides a clearer picture of the stock's price direction by calculating the average closing prices over a set number of periods. The moving average is often used by traders and analysts to identify potential buy or sell signals and to gauge support and resistance levels in the stock's price chart.

Week 5:

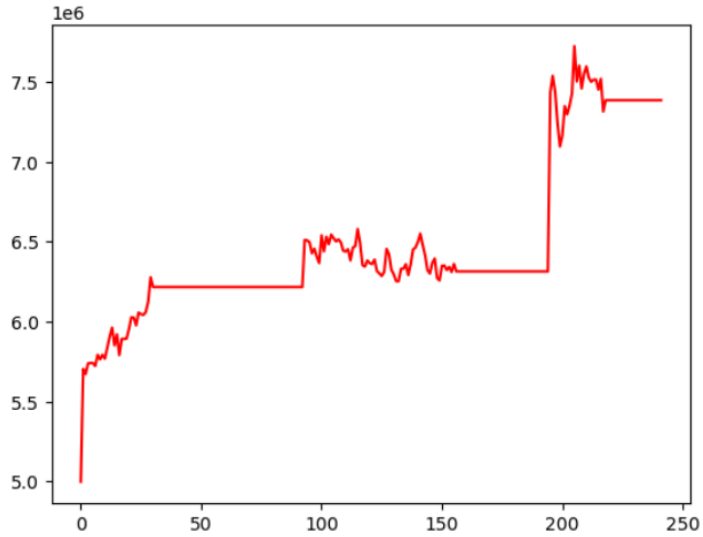
- In week 5, we improved our week 4 code by adding trailing stop loss
- Stop Loss:
    - Stoploss is defined as a certain price a trader sets where he would square off his trade if its price moves in the opposite direction that he expects it to. A stoploss is set to minimize one's losses in case the price does not move as the trader expects it to. In the mean reversion strategy, the stoploss can be set at anywhere between 90-99% of the price. This can be decided depending upon the risk one is willing to take and the returns generated in each case.
- Trailing Stop Loss:
    - The concept of trailing stoploss is that as the price moves in the desired direction the stoploss is also moved in that direction. The stoploss is continuously changed as a certain percentage of the price as long as it is moving in the desired direction. Otherwise, it is kept unchanged. Trailing stoploss helps limit losses and also lock in profits. Trailing stoploss also maximizes the profits as it allows the trader to profit as long as the price moves in his favour.
- So our code had 2 functions, with stop loss(week 5) and without stop loss(week 4). The results obtained were almost similar indicating that the mean reversion strategy fused with pairs trading strategy is decent.

In [12]:
```
without_stop_loss(idbi,axis,z_score2,1.3)
```

```
% returns:  47.685065762719994  %
capital:  7384253.288136
non investment capital:  208.65199600020424
```

Out[12]:

	trade	number	square_off
0	253	1	283
1	345	2	409
2	447	2	471

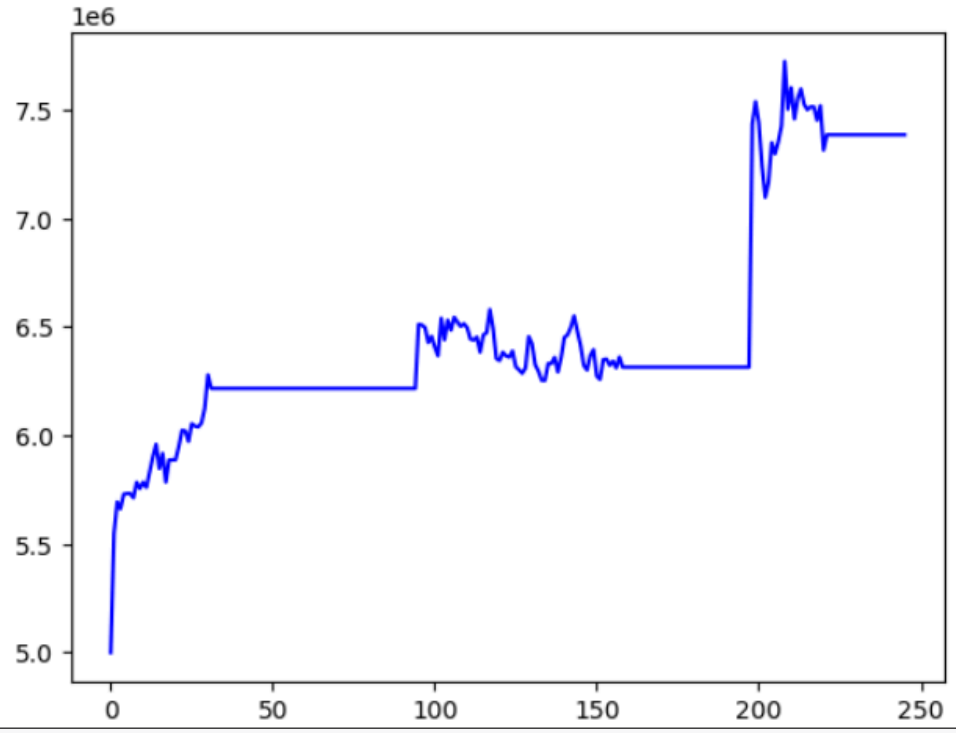


In [13]:
```
with_stop_loss(idbi,axis,z_score2,1.3)
```

```
   trade  number  square_off
0    252       1         283
1    345       2         409
2    447       2         471
% returns:  47.7179836053  %
capital:  7385899.180265
non investment capital:  629.7321040001698
```

Out[13]:

	trade	number	square_off
0	252	1	283
1	345	2	409
2	447	2	471



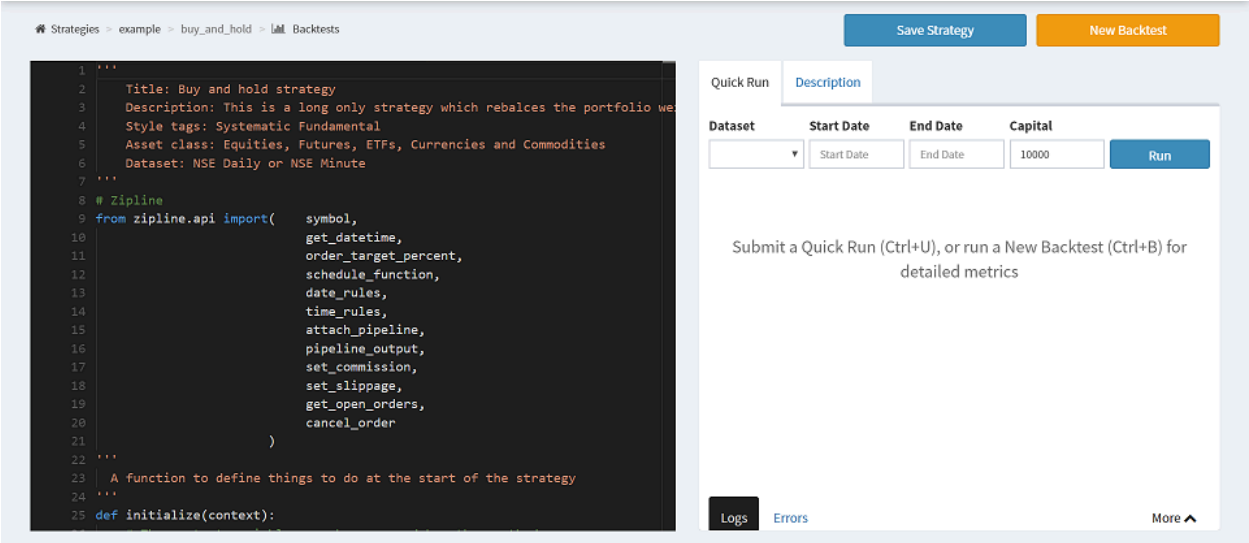
- As the strategy was decent we backtested it on blueshift platform to obtain the results.

Week 6:

Blueshift Platform-Blueshift® is a platform to research and trade systematic investment strategies in Python. It is fast (real-time), flexible and reliable - anything you can code, it can run. It is asset-class and instruments agnostic - we support multiple asset classes and instruments.

First We have go through this blueshift doc and understand its platform and API like using basic functions -context.securities,schedule_function,date.history,square_off,lookback ,initialize, `def handle_data(context,data) etc..`

Implemented strategy and testing that is last part using our Portfolio of stocks.



Some of the basic functions which we have used are:

- schedule_function() - used to schedule the functions involved in the strategy. eg: enter() was scheduled daily after the market opens, and close_out() before it closes.
- square_off() - used to square off the trade
- set_stoploss() and set_takeprofit() are used to initialize the stop loss and take profit parameters

- order() - used to place order for the stock, along with the quantity

  Using the Blueshift platform was quite easy to backtest and modify the parameters to improve the strategy.

- From Week 5 results it was clear that the mean reversion strategy implemented in Week 4 was decent, as the trailing stop loss was encountered rarely and the strategy gave almost similar results with or without stop loss.
- Hence we decided to backtest this strategy on the Blueshift platform. We backtested the strategy from 1st Jan 2023 to 28th July 2023. Before backtesting we ensured that the pair of stocks (IDBI and Axis bank) are stationary in the time period.
- Here are the results of ADF test on the pair of stocks which gave p-value less than 0.05.

CODE

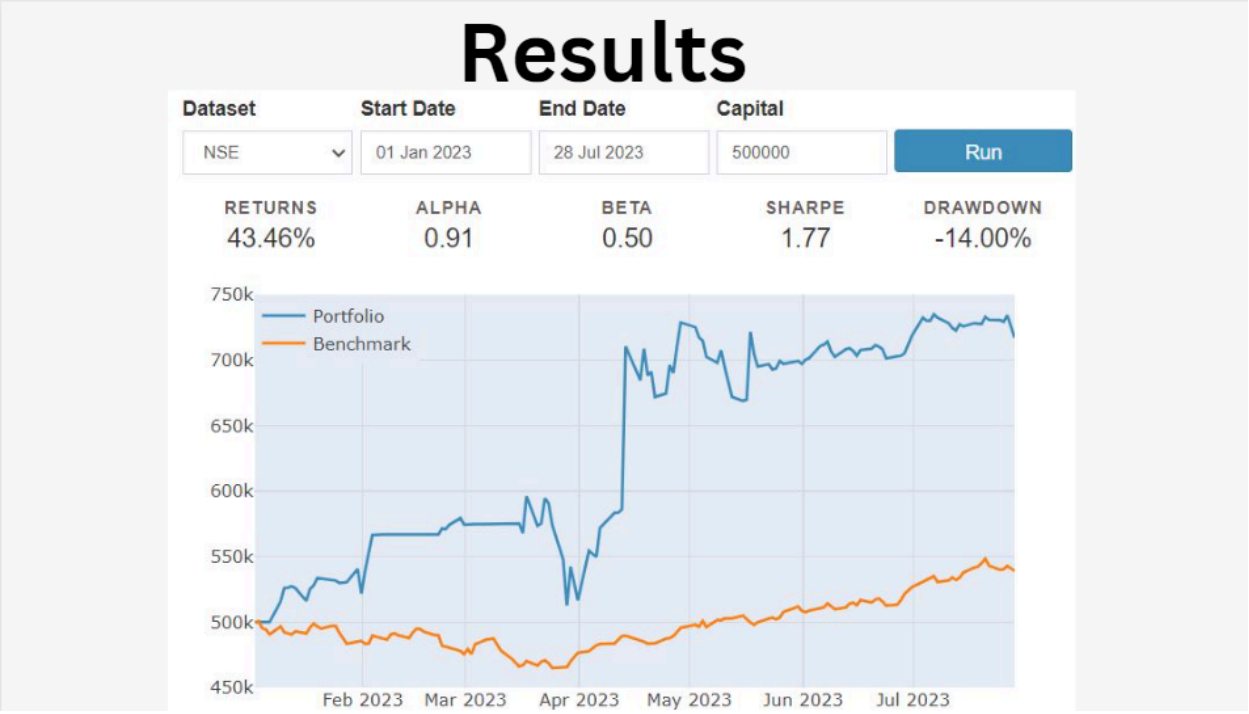
```
    k = 0.75
    if z_score > k and z_score_prev < k:
        print('detected 1')
        # order(context.securities[0], -size)
        order(context.securities[0], -0.5 * context.portfolio.cash / s1_close[-1])
        # order(context.securities[1], size)
        order(context.securities[1], 0.5 * context.portfolio.cash / s2_close[-1])
    if z_score < -k and z_score_prev > -k:
        print('detected 2')
        # order(context.securities[0], size)
        order(context.securities[0], 0.5 * context.portfolio.cash / s1_close[-1])
        # order(context.securities[1], -size)
        order(context.securities[1], -0.5 * context.portfolio.cash / s2_close[-1])
    if 0.1 > z_score > -0.1:
        print('detected 0')
        square_off(context.securities)

    context.traded = True
    schedule_once(set_targets)

def close_out(context, data):
    # square_off(context.securities)
    for oid in context.open_orders:
        cancel_order(oid)

    for asset in context.portfolio.positions:
        order(asset, 0)

def set_targets(context, data):
    for asset in context.portfolio.positions:
        if asset in context.entered:
            continue
        set_stoploss(asset, 'PERCENT', context.params['stoploss'])
        set_takeprofit(asset, 'PERCENT', context.params['takeprofit'])
        context.entered.add(asset)
```



Conclusion- We have used the Blueshift platform and its API call to backtest strategies on Portfolio and wrote down returns .