CSI Project- Tanay Sheth

This project is designed to test candidate's problem solving capacity in a real-world situation using python.

1) Data download
Please download the index composition of CSI500 index at date = '2021-01-01'. This is available on various free, open source securities data platforms.
tip1: Google translate add-on can be used in case the website is not in English.

2) Download 30min bar data from 2022-04-01 to 2022-07-31 for all 500 stocks of the CSI500 index

3) Design a mean-reverting strategy:
- We are trying to design a long only strategy where you hold a long position of 100 every day (constant notional) across the stocks that you think will rebound up.
- We are trying to do better than just buy and hold. We want to beat the 500 stocks equal-weighted.
- Design 3-5 features that would describe how much a stock is oversold or overbought · Test the features from 2022-04-01 to 2022-06-30
- What would be your idea to improve the performance of the strategy?
- How you quantify how good is your strategy?
- Test the out-sample from 2022-07-01 to 2022-07-31

Requirements:
1) Send the project via github or textfile .py
2) If you use any third part package or library please include the lib install in the code as well 3) Explain clearly the objectif of each section of your code

Backtesting Report

Returns (Without leverage)

	Raw Total Return	Annualized Return
In-Sample	46.4917%	53.683%
Out-Sample	11.3012%	39.1485%

Annualized Return = monthly returns * (12)^0.5

Strategy Performance

	Hit ratio	Sharpe Ratio	Treynor ratio	Alpha
In-Sample	0.390	0.505	0.039	16.4%
Out-Sample	0.344	0.276	0.578	10.4%

Alpha is simply (returns - market returns) here.

Drawdowns

	Maximum Drawdown	Maximum Drawdown Duration
In-Sample	3.5%	4 Trading sessions
Out-Sample	3.4%	3 Trading sessions

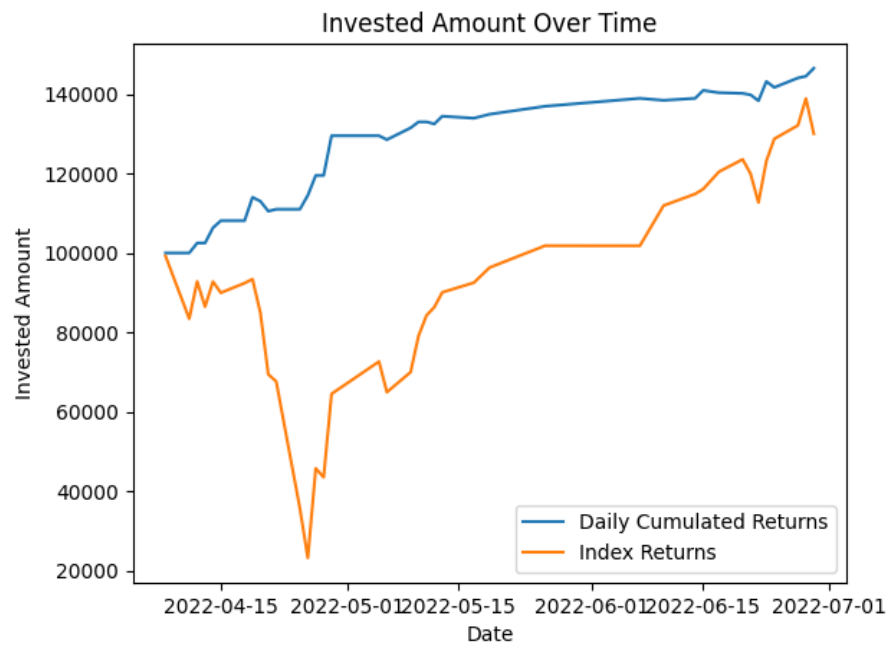### Other Expenses

	Brokerage	Slippage PnL
In-Sample	6150 (6.15%)	34.157%
Out-Sample	3050 (3.05%)	5.196%

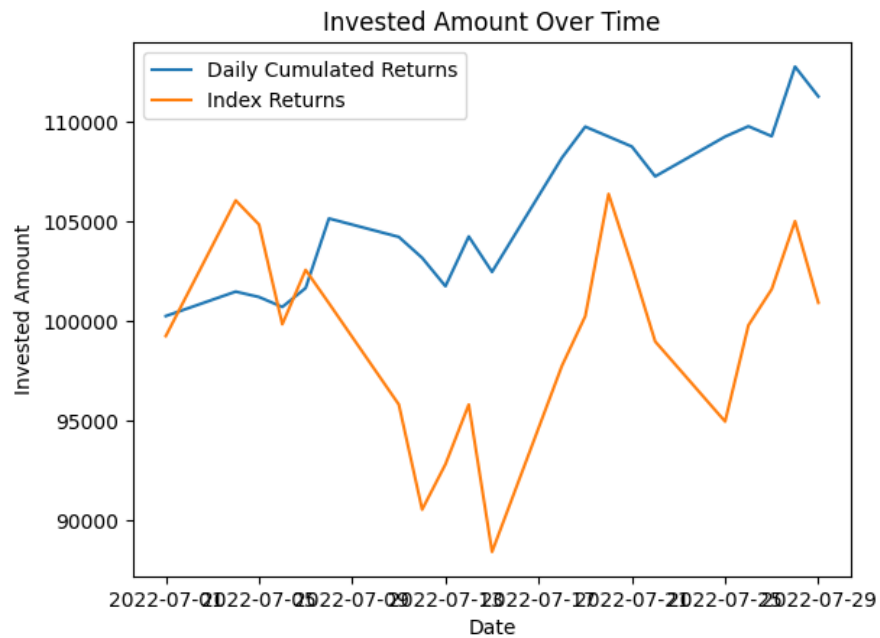Intraday Brokerage is taken as Transaction value * 0.05 (.05%)
Slippage is considered as the PnL of strategy if entry position is entered 0.1% price late because of execution time

Plots of Invested Amount over time

In-Sample Plot :

Out-Sample Plot:

Strategy

Entry conditions : (Everyday at 9:30)
Filter stocks for : RSI indicator < 40 and MACD Line > Signal Line
Rank stocks based on K% (Stochastic Oscillator) and choose top 5 from the filtered stocks

Exit Conditions:
Target 2% , stop loss 0.5%
Else square off at end of day

Ways we can improve this strategy:
More granular (minute-level) and expansive data (volume could provide scope for more oversold indicators) to optimize the strategy
Increase the testing period both in-sample and out-sample
Minimize the universe, once a month, based on fundamentals as we are dealing with mid cap and small cap stocks

Data Challenges:
Only daily EOD OHLC data is available for CSI500 data at open-source data providers and APIs , 2022.
Sources Tested: Finance, Alphavantage, TwelveData,Official Stock Exchange Sites,
Unable to connect to chinese data API TuShare and Sina Finance

To perform backtesting **synthetic data production** technique is used:
Chinese have 4 hour trading sessions. So 8 ; 30 min candlesticks are formed daily.
-   Open from OHLC is used as close for first candle
-   6 completely random values between high and low of OHLC are taken for 6 candles
-   Close from OHLC is is used as final closing price for the day

By this method 8 close prices for daily 30 min candles are produced to prepare and backtest the strategy. Same backtesting code can be used to test the strategy on real data