Quant Club Recruitment Task 2025

Backtesting a Trading Strategy in Python

Quant Club, BITS Pilani

Statement

You will **backtest a simple trading strategy** using historical stock price data. Backtesting means applying your strategy to past data and seeing how it would have performed.

You will:

- Work with real stock price data (open, high, low, close).
- Understand what it means to generate a trading *signal* or *alpha*.
- Calculate the day-to-day profit and loss (PnL) of your strategy.
- Evaluate performance using standard metrics such as Sharpe ratio, drawdown, and cumulative return.

Step 1: Get the Data

Use the Python library `yfinance` to download daily stock data. Choose any stock (e.g., Apple, Amazon) or index (e.g., S&P 500).

The data you get will typically include:

- `Open`: Price at the start of the day.
- `High`: Highest price of the day.
- `Low`: Lowest price of the day.
- `Close`: Price at the end of the day.

For our strategy, we will mainly use the `Close` price.

Points to Ponder:

1. Why do you think traders often use the `Close` price instead of intraday prices?
2. What are the limitations of using freely available Yahoo Finance data?

Step 2: Define the Strategy

We will use a very simple moving-average based trading strategy.

- Compute a **moving average (MA)** of the close price:

$$MA_t = \frac{1}{n} \sum_{i=0}^{n-1} Close_{t-i}$$

  where n is the lookback window (e.g., 20 days).

- The trading rule:

$$\begin{aligned} \text{Position}_t &= +1 \quad \text{if } Close_t > MA_t \quad \text{(go long, buy 1 share)} \\ \text{Position}_t &= -1 \quad \text{if } Close_t < MA_t \quad \text{(go short, sell 1 share)} \end{aligned}$$

  Thus, each day you either hold $+1$ share (long) or -1 share (short).
  Points to Ponder:

1. Why might moving averages capture trends in stock prices?
2. What could happen if your window size n is too small or too large?

Step 3: Compute Daily PnL

Your profit and loss (PnL) on each day depends on your position and the price change.

$$PnL_t = \text{Position}_{t-1} \times (Close_t - Close_{t-1})$$

- If you are long ($+1$) and price goes up, you gain. If it goes down, you lose.
- If you are short (-1) and price goes down, you gain. If it goes up, you lose.

The cumulative PnL is simply the running total:

$$CumulativePnL_t = \sum_{i=1}^{t} PnL_i$$

1. Why is the position from *yesterday* used to calculate today's PnL?
2. What are some assumptions we are making here.

Step 4: Evaluate the Strategy

Now you must evaluate how good (or bad) your strategy is. Some standard metrics:

1. Sharpe Ratio

The Sharpe ratio measures returns relative to risk:

$$Sharpe = \frac{E[R]}{\sigma(R)}$$

where R is the series of daily returns, $E[R]$ is the average daily return, and $\sigma(R)$ is the standard deviation of daily returns.

2. Maximum Drawdown

Drawdown measures the fall from a peak in your cumulative PnL.

$$Drawdown_t = \frac{CumulativePnL_t - \max_{s \le t} CumulativePnL_s}{\max_{s \le t} CumulativePnL_s}$$

The **maximum drawdown** is the largest such drop over the entire time.

3. Total Return

This is just the final cumulative PnL after the backtest period.

1. Why is the Sharpe ratio more informative than just looking at total return?
2. Why is maximum drawdown an important risk measure for traders?

Step 5: Interpreting Results

At the end, you should be able to produce:

- A time series of your daily positions (+1 or -1).
- A time series of daily PnL.
- Cumulative PnL curve.
- Sharpe ratio, maximum drawdown, and total return.

Additional Requirements for 2024 Batch

In addition to the moving average strategy described earlier, the 2024 batch is required to go one step further:

- **Develop Your Own Strategy:** You must design and implement a new trading strategy of your choice. This could be based on technical indicators (e.g., RSI, Bollinger Bands), price patterns, or any idea you can justify logically.
- **Performance Metrics:** The strategy should perform *reasonably well* when evaluated using the same metrics as before (Sharpe ratio, maximum drawdown, cumulative PnL, etc.). In particular, try to maximize the Sharpe ratio.

- **Interpretability:** It is not enough that the strategy makes money in backtests. You must also be able to explain the underlying intuition of why the strategy could work in markets.
- **Validation:** To avoid overfitting, split your data into *in-sample (training)* and *out-of-sample (validation)* periods. Develop and tune your strategy on the in-sample data, then test its robustness on the out-of-sample data.

Deliverables

Submit:

- Your Python code that downloads data, computes moving average, generates positions, calculates PnL, and computes the metrics.
- Plots of cumulative PnL.

Closing Note

The purpose of this exercise is not just to write code, but to understand the importance and meaning of backtesting trading ideas, and more importantly, to learn how to interpret the results correctly. This is meant to get you started on thinking rigorously about strategies, data, and metrics.

Additional Resource

You may refer to the following video for more information about back testing trading strategies.

Neurotrader - Backtesting trading strategies